“IT’S LIKE, YOU KNOW”: VARIATION IN THE PRODUCTION AND SOCIAL PERCEPTION OF DISCOURSE-PRAGMATIC LIKE IN PIDGIN AND HAWAI’I ENGLISH

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI’I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN LINGUISTICS SEPTEMBER 2019

by Claire Marie Stabile

Dissertation committee:

Katie Drager, co-chair
Kamil Ud Deen, co-chair
Bradley McDonnell
Rory Turnbull
Candace Fujikane
Acknowledgments

Acknowledgments go here.
Abstract

This dissertation examines variation and change in discourse-pragmatic *like*, in two languages of Hawai‘i, Pidgin (Hawai‘i Creole) and Hawai‘i English (a local variety of English). While discourse-pragmatic *like* has been the focus of robust study in global Englishes, to date, no such study has examined this phenomenon in Hawai‘i, making this dissertation the first of its kind.

Combining results from a matched-guise perception experiment with multiple corpus-based analyses, this study provides a unique perspective on how this rapid, global language change is operating within a local, multilingual context and also provides insights into what type of social work *like* is performing for different groups of speakers in Hawai‘i.

Major findings show that in Hawai‘i, discourse-pragmatic *like* is patterning similarly with other varieties studied worldwide, with a few interesting differences. Young Pidgin speaking men in Hawai‘i use discourse marker *like* at higher rates than young women. In Hawai‘i English, however, women and men pattern more similarly. This finding deviates from patterns observed in other studied varieties, where young men are using discourse marker *like* at lower rates than young women (D'Arcy, 2007). A novel approach developed in this dissertation examines how speakers are using discourse marker *like* within the context of the surrounding discourse; for example, to elaborate, clarify, or provide illustrative commentary.

By examining discourse-pragmatic *like* in Hawai‘i using both a perception experiment as well as multiple corpus analyses, the work presented here not only provides a detailed description of discourse-pragmatic like in Hawai‘i, but provides methodological recommendations for other researchers conducting future work on discourse-pragmatic elements.
# Table of Contents

ACKNOWLEDGMENTS ........................................................................................................ II
ABSTRACT .......................................................................................................................... III

1. INTRODUCTION ............................................................................................................ 1

2. BACKGROUND ................................................................................................................ 5

   2.1. Language variation and change .............................................................................. 5
       2.1.1. A brief history .............................................................................................. 5
       2.1.2. Common research methods ......................................................................... 8
           2.1.2.1. Methods of analyzing corpora ............................................................... 8
           2.1.2.2. Matched-guise perception tasks ......................................................... 9

   2.2. Discourse-pragmatic variation .............................................................................. 10
       2.2.1. What is a discourse-pragmatic variable? ..................................................... 10
           2.2.1.1. Discourse markers ............................................................................... 10
           2.2.1.2. Discourse particles ............................................................................. 11
           2.2.1.3. Quotatives .......................................................................................... 12
           2.2.1.4. Approximate adverbs ........................................................................ 12
           2.2.1.5. Sentence adverbs ............................................................................... 12
           2.2.2. Discourse-pragmatic like ......................................................................... 12
               2.2.2.1. Discourse marker like ................................................................... 13
               2.2.2.2. Discourse particle like .................................................................... 14
               2.2.2.3. Quotative like ................................................................................ 14
               2.2.2.4. Approximate adverb like ................................................................. 15
               2.2.2.5. Sentence adverb like ....................................................................... 16
           2.2.3. Ideologies around like ............................................................................... 16

   2.3. Language in Hawai‘i .............................................................................................. 18
       2.3.1. Ideologies ..................................................................................................... 19
           2.3.1.1. Sociolinguistic ideologies in Hawai‘i ..................................................... 19
           2.3.2. Discourse-pragmatic like in Hawai‘i ........................................................ 20
               2.3.2.1. Quotatives in Hawai‘i English ............................................................... 20
               2.3.2.2. Quotatives in Pidgin ........................................................................ 21

3. PERCEPTION OF QUOTATIVE LIKE IN HAWAI‘I ..................................................... 23

   3.1. Introduction ............................................................................................................. 23
   3.2. Methods .................................................................................................................. 23
       3.2.1. Conditions .................................................................................................... 23
       3.2.2. Stimuli .......................................................................................................... 23
           3.2.2.1. Critical Items ..................................................................................... 23
           3.2.2.2. Control items ...................................................................................... 24
           3.2.2.3. Distractor items .................................................................................. 24
           3.2.2.4. Voices used in the experiment .............................................................. 24
       3.2.3. Procedure ...................................................................................................... 25
           3.2.3.1. Presentation of stimuli ....................................................................... 25
       3.2.4. Participants .................................................................................................... 26

   3.3. Analysis ................................................................................................................... 26
       3.3.1. Coding participant demographics ................................................................ 26
1. Introduction

This dissertation sets out to examine variation and change in discourse-pragmatic *like*, a global phenomenon of interest to sociolinguistics, within the unique multilingual context of Hawai‘i. Sociolinguists have long been interested in the social drivers of linguistic variation and change at all levels of the grammar. Since the early 1980s, researchers have been especially intrigued by the rapid, uniform, and global expansion of the discourse-pragmatic functions of *like*. Discourse-pragmatic *like* entered the system of global Englishes and diffused rapidly, making its way into neighboring contact varieties, and giving rise to a robust literature on quotative *be like* in particular (Butters, 1982; see Tagliamonte, D’Arcy, & Louro, 2016 and D’Arcy, 2017 for a detailed overview). This dissertation represents the first effort to examine multiple types of discourse-pragmatic *like* in two languages of Hawai‘i: Pidgin (Hawai‘i Creole) and Hawai‘i English (a local variety of English).

The main goals of this study are to provide a description of discourse-pragmatic *like* in Pidgin and Hawai‘i English; who uses it, how are they using it, and how it is perceived and socially interpreted by Hawai‘i Locals. An additional aim is to compare how it patterns in both varieties, as well as to patterns in other studied Englishes worldwide. A key contribution of this dissertation is in the comparison of perception and production data, combined with an in-depth analysis of the discourse marker. In doing so, the study provides a detailed picture of discourse-pragmatic *like* in Hawai‘i made possible only by the combination of methods employed herein.

Chapter 2 provides context and background for the current study. Section 2.1 provides a brief history of research into language variation and change (2.1.1), and outlines several research methods commonly employed by sociolinguists (2.1.2). Section 2.2 presents an overview of the literature on discourse-pragmatic variation, covering the types of discourse-pragmatic variables under study (2.2.1) with a particular focus on the research conducted on discourse-pragmatic *like* (2.2.2). Section 2.2.3 details research on ideologies surrounding *like*. Finally, section 2.3 provides contextualizing information on language in Hawai‘i. An overview of language ideologies is presented in section 2.3.1, and section 2.3.2 covers the research conducted thus far on discourse-pragmatic *like* in Hawai‘i.

A matched-guise perception experiment (chapter 3) provides insight into how young Hawai‘i Locals perceive discourse-pragmatic *like* when used in Pidgin and Hawai‘i English. While perception work in other varieties has found discourse-pragmatic *like* is frequently associated with women and often evaluated negatively, no such work has yet examined how discourse-pragmatic *like* is socially evaluated in Hawai‘i. Examining how speakers perceive discourse-pragmatic *like* sets the stage for comparing these perceptions to corpus distribution in order to form a more nuanced picture of discourse-pragmatic *like* in Hawai‘i. These results are compared with perception work on *like* in other varieties, as well as popular ideologies surrounding *like* in order to
contextualize this work within the greater field of study.

Chapter 4 presents three distinct yet complimentary production-based studies of discourse-pragmatic *like*. Drawing on corpora of Pidgin and Hawai’i English, this dissertation first presents a frequentist analysis of *like*, describing how each variant is used by different demographic groups in both languages (section 4.1.4). Following this, a variationist analysis is used to provide a detailed picture of how discourse marker *like* has changed over time in the systems of Pidgin and Hawai’i English (section 4.1.5). Conducting both a frequentist as well as variationist analysis allows a more detailed understanding how discourse marker *like* is increasing in apparent time relative to other variables of *like*, as well as other discourse marker variants that speakers may choose to employ. Finally, all tokens of discourse marker *like* are analyzed within the context of the surrounding discourse, to examine the distribution of pragmatic uses of discourse marker *like* across age and gender in both varieties (section 4.1.6). Examining the type of social work that discourse marker *like* performs for different groups in both Pidgin and English may shed light on the social forces driving language variation and change within the unique context of Hawai’i. To foreshadow the results, while women and men have both increased their use of discourse marker *like* in apparent time, it appears that they are doing so for different conversational purposes.

The major findings of the corpus study show that in Hawai’i, discourse-pragmatic *like* is patterning similarly with other varieties studied worldwide. The frequentist analysis described in chapter 4 shows that in both Pidgin and Hawai’i English, discourse-pragmatic *like* has increased in use among the younger cohort of speakers. In Pidgin, young men use discourse marker *like* at higher rates than young women. The variationist analysis offers a slightly different perspective on variation in the discourse marker systems of both Pidgin and Hawai’i English. These results provide evidence that young Pidgin speaking men in Hawai’i are using discourse marker *like* at higher rates than young women. In Hawai’i English, however, women and men pattern more similarly. This finding deviates from patterns observed in other studied varieties, where young men are using discourse marker *like* at lower rates than young women (D’Arcy, 2007). A novel approach developed in this dissertation examines how speakers are using discourse marker *like* within the context of the surrounding discourse; for example, to elaborate, clarify, or provide illustrative commentary. Understanding the pragmatic uses that speakers employ specific forms for may help us to understand what social factors may be driving the changes occurring within the system.

By examining discourse-pragmatic *like* in Hawai’i using both a perception experiment as well as multiple corpus analyses, this dissertation is able to provide a unique perspective on how this rapid, global language change is operating within a local, multilingual context. This study also provides insights into what type of social work *like* is performing for different groups of speakers in Hawai’i. The results presented here not only provide detailed description of discourse-pragmatic like in Hawai’i, but provide methodological recommendations for other researchers conducting future
work on discourse-pragmatic elements.
2. Background

2.1. Language variation and change
2.1.1. A brief history
Since the emergence of sociolinguistics as a sub-discipline, researchers have focused on understanding language variation and change. Beginning with Labov's influential study of sound change on Martha's Vineyard (1963), research has focused on understanding the relationship between linguistic and social variables. The original Martha's Vineyard study looked at variation in vowels among speakers of different ages and proposed that the differences observed could be used as synchronic evidence for ongoing language change. Labov's study was groundbreaking on two fronts; it pioneered the method of apparent time and established the systematic relationship between linguistic variables and social traits.

Eckert (2011) describes the history of sociolinguistics as occurring in three “waves” of study. Research in the first wave focused on establishing links between linguistic variables and broad social categories such as geographic region, socioeconomic status, speaker age, and speaker ethnicity (Labov, 1963; Labov, 1966; Wolfram, 1969). These pioneering studies established the importance of describing variation and considering social categories within linguistic research and theoretical frameworks. Many of these studies employed the apparent time method, which has continued as a method for investigating language change through the second and third waves of study and has been central to the study of language variation (Pope, Meyerhoff, & Ladd, 2007).

Apparent time is a synchronic method of modeling language change, used as a proxy for longitudinal change. The major assumption of this method is that the majority of linguistic features are acquired during childhood and early adolescence and remain relatively stable thereafter. While it is known that individuals do change their language in adulthood to a degree, (see (Harrington, 2006) for one example) this is likely not a confounding factor for this study. Given the breadth and depth of work on discourse-pragmatic like worldwide, we would expect to see evidence occurring elsewhere. More research into discourse-pragmatic like in Hawai’i would of course be a welcome addition to this body of research. On this assumption, cross sections of the population can be used to show changes in the linguistic system over time; a speaker’s age can be used to examine the time period in which they acquired their linguistic systems. By examining multiple cohorts of speakers together, a picture of change over time can be constructed.

In the second wave of sociolinguistic study, researchers began to examine how speakers align with locally relevant categories, and how this helps to explain how change diffuses in a given region. For example, Eckert’s seminal study of high school cliques examined how girls and boys in two distinct social groups (burnouts and jocks) produced particular vowels, finding that Burnout girls participated at much higher
rates in sound change taking place in nearby Detroit (Eckert P., 2000). Studies in the second wave, such as Labov’s (1972) study of African American Vernacular English in New York, conceptualized speakers’ use of linguistic features as an agentive act on the part of the speaker to index in-group status. The Milroy’s (Milroy & Milroy, "Belfast: change and variation in an urban vernacular", 1978) Belfast study focused on sociolinguistic patterns within a localized context. The study used social network analysis to investigate relationships between 48 individuals from three working-class, urban communities in detail. The results indicated that participants’ use of “vernacular” variants was strongly influenced by how tightly integrated their social network ties were, and demonstrated the usefulness of social network analysis in dialectology.

Studies comprising the third wave of sociolinguistic research are interested in stylistic variation; how speakers employ their linguistic resources to construct contextually relevant stances, personae, and styles. This work illuminates the importance of understanding context and social goals when examining a speaker's linguistic production as well as perception, and moves beyond static social categories. Studies in the third wave focus on the ways in which speakers utilize their linguistic resources as stylistic practice to position themselves in the social landscape (Bucholtz & Hall, 2005). Podesva (2007) describes the ways in which Heath, a gay medical student, employs falsetto phonation to construct situationally dependent personae. In the medical clinic, Heath adopts a “caring doctor” persona, and his use of falsetto is shorter and less frequent. At a barbecue with friends, Heath employs falsetto to construct a “diva” persona, his falsetto is longer, more frequent, and has a wider intonational contour. By altering the way in which he employs his linguistic resources, Heath is able to construct various personae as he moves throughout his social interactions.

In addition to work on language change, sociolinguists are interested in attitudes and ideologies. Research into language ideologies is concerned with social evaluations of language, frequently as they relate to standard and non-standard varieties (Milroy, Language ideologies and the consequences of standardization, 2001; Lippi-Green, 2012). Language ideologies are particularly important to understand because they can bias social interaction (Cargile & Giles, 1997). This is particularly relevant in situations of multiculturalism and language contact, such as Hawai‘i, where language attitudes often intersect with attitudes towards education and social mobility in the public discourse (Marlow & Giles, 2008; Bayer, 2009). Thus, it is important to understand what language ideologies exist in Hawai‘i, and how local language varieties are perceived. Section 2.3 provides an overview of relevant language history and ideologies in Hawai‘i.

The focus on style and stance extends to perception as well as production, with studies focusing on listener perception of speaker style (Campbell-Kibler, 2011), gender and sexuality (Levon, 2014), and geographic region (Barnes, 2015; Bucholtz, Bermudez, Fung, Edwards, & Vargas, 2007). In a perceptual study of 97 Californians, Villarreal (2018) found that listeners’ perceptions of the California Vowel Shift (CVS) differed
from the social meanings indexed by this shift in previous production work. The California vowel shift is a series of vowel shifts which appear to be coordinated, and captured the attention of researchers in the 1980s. While production research has shown that the California Vowel Shift carries social meanings of carefreeness, femininity, and privilege, this perception work found that only listeners from Southern California rated speakers with CVS as sounding feminine, while listeners from Northern California rated CVS as more masculine. This work demonstrates that the social meanings carried by linguistic variables are not solely determined by speakers’ use of these forms, but also by listeners’ exposure and experience with the forms as well as the speakers who use them. In other words, speakers and listeners participate together in constructing social meaning out of linguistic forms. Recently, perception of sociolinguistic variation has been of increasing interest in the field of sociolinguistics. Matched guise work has shown that listeners perceive discrete social associations for different variants of (ING) (Campbell-Kibler, 2011). In her study, listeners perceived -ing guises as more intelligent/educated, and -in guises as less formal, leading Campbell-Kibler to conclude that variants are linked to specific meanings, though variants are not “social flip sides of the same coin” (Campbell-Kibler, 2011, p. 436).

Recently, there has been renewed interest in combining perception and production methods to explore nuanced questions in sociolinguistics. Recent work (Hay, Drager, & Gibson, 2018) has shown that when listeners have experience with a particular variable in the input, it does not stand out to them as much. In a production study, the authors analyzed corpus data to determine the frequency of r-sandhi in modern New Zealand English, and determined that r-sandhi is present in several contexts in modern New Zealand English. This was then followed up with a perception study, which demonstrated that New Zealand listeners who were familiar with r-sandhi in a given context expected to hear it in that context, and did not demonstrate surprisal or increased attention to its presence. Conversely, listeners from San Diego California (where r-sandhi is not present) who were not familiar with r-sandhi in the given context demonstrated surprisal and increased attention to the presence of [u] in the experimental stimuli. Taken together, these results demonstrate that when listeners have prior experience with a variable in the input, that variable ceases to be surprising to them. If listeners are not familiar with a particular variable, they will demonstrate surprise and increased attention to the variable. This body of work informs the interpretation of the results of the perception experiment detailed in chapter 3.

While many studies focus on variation on a small scale, a broader goal of variationist sociolinguistics is to integrate the study of variation into the larger understanding of language. The goal then is that linguists should not treat language as “an abstract object which can be accounted for without reference to social concerns of any kind” (Romaine, 1994, p. ix) but should instead consider how social factors can influence language use and change over time. While most variationist work has long centered on sounds, there has been increasing work at other levels of the grammar. This dissertation represents one such effort.
One case of language change that has captured the attention of the field has been the rise of discourse-pragmatic *like*, and the quotative function in particular. These features are of particular interest due in large part to the fact that they arose seemingly simultaneously in worldwide Englishes, and present an opportunity to document such a change in real time. Section 2.2.2 contains an overview of discourse-pragmatic *like*. While the quotative system of English has been of interest to linguists for several decades now (Butters, 1980; Tagliamonte, D'Arcy, & Louro, 2016) the quotative systems of Hawai‘i English and Pidgin have only recently been examined and described (Drager, Chun Comstock, Stabile, & Schutz, in prep; Schutz, Chun Comstock, Stabile, & Drager, in prep). Hawai‘i English is a variety of English spoken in Hawai‘i, which is phonologically similar to other North American varieties (Kirtley, Grama, Drager, & Simpson, 2016). Pidgin (sometimes referred to by linguists as Hawai‘i Creole, Hawai‘i Creole English, or Hawai‘i English Creole) is an English lexified-creole which developed during the plantation era of the 1800s, and is often referred to by linguists as Hawai‘i Creole (Sakoda & Siegel, 2003). (See section 2.3 for a brief overview of language in Hawai‘i).

Quotative *like* is one of 12 varieties of *like* described by D'Arcy (2017). *like* is of particular interest to researchers in that while it has been present in the system of English for hundreds of years, its usage is morphing and expanding to include discourse-pragmatic functions (D'Arcy, 2017). Additionally, *like* is socially salient\(^1\), and frequently commented upon (Daily-O'Cain, 2000; D'Arcy, 2007; D'Arcy, 2017). While a growing body of research has documented the discourse pragmatic features of *like* in world Englishes, to date no such research into the discursive functions of *like* has focused on languages in Hawai‘i, making this dissertation the first to do so in the region.

2.1.2. Common research methods
A number of methods are used in sociolinguistic research. Ones that are relevant for this dissertation are methods of analyzing corpora, and the matched-guise experimental methodology.

2.1.2.1. Methods of analyzing corpora
Corpus data is often employed in variationist sociolinguistics, since it can readily providing data from speakers of relevant demographic groups. Corpora may consist of written or audio recorded data. When analyzing corpus data, two main approaches are often used. The corpus linguistics approach, also known as the frequentist method, is a form-based approach which uses normalization (per 100 or 1000 words, depending on the size of the corpus) as the method of quantification. In contrast, the sociolinguistic, or variationist method frames the analysis within the envelope of variation; investigating the set of variants which can be used to convey the same or similar meaning. For example, calculating the proportion of all quotatives which are

\(^{1}\) While the term “salience” can mean different things in the context of different sub-fields, the meaning used herein is one of “noticeability” or “standing out”
realized as _be like_ out of all quotatives. As opposed to calculating absolute frequency, the variationist analysis calculates proportional frequency.

There are pros and cons to each type of analysis. The variationist analysis allows us to see how a variant operates in relation to other variants of the same variable and, in conjunction with the apparent time method, gives us a window into the changes that occur within a system (for example, discourse markers). With this method, we can see the change over time in how speakers choose to say the same thing. In contrast, the frequentist analysis is argued to be closer to the perception of the speaker (Dinkin, 2016). That is, a native speaker does not distinguish between discourse marker _like_ and quotative _like_, instead viewing all types of _like_ as the same thing (see section 2.2.3 for a discussion of ideologies surrounding _like_). Since the frequentist method calculates the frequency of a particular _like_ out of all _likes_ in the corpus, it treats all types of _like_ as equal, as perception research suggests speakers do.

Dinkin (2016) argues that we must look beyond the envelope of variation, looking at both the variant as well as the variable, in order to get a full picture of the sociolinguistic work a variant performs. Increasingly, researchers have begun to recognize and apply the potential for frequentist and variationist analyses to work in conjunction to improve sociolinguistic research (Bauer, 2002; Beal, Corrigan, & Moisl, 2007; Anderson, 2008; Romaine, 2008; Szmrécsanyi, 2017). This dissertation aims to contribute to this larger discussion by conducting not only a variationist analysis of discourse marker _like_, but also a frequentist analysis of all types of _like_, thus examining both the variant as well as the variable. Details of the current approach can be found in chapter 4.

2.1.2.2. **Matched-guise perception tasks**

Matched-guise methodology is often employed in perception research, and involves presenting listeners with the same stimuli in different “guises”, where the linguistic variable of interest can be manipulated while holding other factors constant. This methodology was developed by Lambert, Hodgson, Gardner, and Fillebaum in order to investigate French Canadian attitudes towards French and English, and has remained a staple of sociolinguistic work ever since (Lambert, Hodgson, Gardner, & Fillenbaum, 1960). Commonly, matched-guise work involves having participants listen to speakers assumed to be different, which in reality are the same speaker manipulated in some fashion (e.g. digitally raising or lowering pitch, fronting or backing of vowels, adding or removing discourse markers, etc.) and evaluating the “speakers” for a variety of traits.

Matched guise work on contact language varieties in Northwestern Spain has demonstrated that manipulating the variants (in this case, vowels) caused listeners’ perception of the speakers’ social attributes to shift (Barnes, 2015). While the study concluded that discrete linguistic variables in contact can index social information, not all variants have equal salience (Barnes, 2015). Crucially, Barnes observes that “the social meaning linked to different languages in a bilingual community is also indexed
by the individual linguistic features of each language” (Barnes, 2015, p. 18). In other words, linguistic features of a language can take on the social meanings associated with that language, even when these features are used in another variety. The linguistic features themselves come to take on social meaning. In a modified matched-guise task, Levon (2014) found that both cognitive and attitudinal factors influenced listeners’ perceptions of social meaning. The study found that processing constraints modulate the degree of attention that listeners dedicate to perception of speech, such that listeners attend to the variables which are most perceptually dominant. Listeners may be exposed to many socially meaningful linguistic variables in the speech stream, but may not attend to all of them.

2.2. Discourse-pragmatic variation
2.2.1. What is a discourse-pragmatic variable?
Broadly, discourse-pragmatic elements comprise a set of expressions characterized by semantic shallowness and optionality, which perform functions relating to pragmatics, such as turn taking, hedging, or mitigation (Miller & Weinert, 1995), or to the structure of the discourse itself, such as illustrating, expanding, exemplifying, hesitating, signaling politeness, or establishing solidarity (Schourup L., 1985; Underhill, 1988; Andersen, 1998; Diskin, 2017). These expressions are drawn from a variety of lexical classes, and undergo some kind of semantic bleaching in their journey towards pragmatic functionality.

Discourse-pragmatic lexical elements (sometimes referred to as markers, particles, or operators) have been an area of interest for a number of years. Much attention has been paid to their functions (Fraser B., 1988; Schourup L., 1999), grammaticalization and ability to introduce reported speech, (Tagliamonte, D'Arcy, & Louro, 2016; D'Arcy, 2017), and the demographics of the speakers who use them (Daily-O'Cain, 2000; Tagliamonte & D'Arcy, 2005; D'Arcy, 2007).

Beyond these broad outlines, many researchers disagree on the specifics of these elements. Indeed, disagreements exist in the very terminology used to refer to them. Some labels, such as ‘discourse marker’ or ‘discourse particle’ can refer to some, but not all of these expressions (notably excluding the quotative), while other researchers use the term ‘discourse-pragmatic marker’ inclusive of the quotative. Schourup (1999, p. 229) argues that the term ‘discourse particle’ is problematic since “‘particle’ has traditionally been a syntactic term, whereas DMs are more often regarded as comprising a functional class that draws on items belonging to various syntactic classes.” In an effort to be broad and inclusive, this dissertation will refer to this set of expressions as “discourse-pragmatic elements”, and will follow D'Arcy’s framework of differentiating between the discourse marker and discourse particle as described in Section 2.2.2.

2.2.1.1. Discourse markers
Discourse markers, such as oh, so, well, and like, have been an area of interest since
the early 1980s when Levinson (1983) first noted that they serve to signal a relationship between an utterance and prior discourse. As previously discussed, the literature around discourse markers can be challenging to navigate, given that not all researchers draw a distinction between discourse markers and discourse particles, and often use the two terms interchangeably. D’Arcy (2006; 2017) draws a distinction between discourse markers and discourse particles based off of their position within an utterance: the marker appears only in clause-initial position, while the particle is situated clause internally. This is the framework adopted herein.

Schourup (1985) notes that these markers serve as a type of “discourse glue”, managing overall discourse coherence and flow. One of the most characteristic features of discourse markers is the function they serve in relating utterances to the larger discourse context, in effect gluing utterances together. Many studies emphasize the fact that the discourse marker specifically serves to relate the immediately following utterance to the prior utterance or turn.

Fraser (1990) conceptualizes discourse markers as a type of “commentary pragmatic marker”, separate from other commentary markers. Fraser’s reasoning for this distinction is that discourse markers alone indicate commentary specifying the nature of the sequential discourse relationship between the current utterance and prior discourse. Fraser (1988) also notes that discourse markers are lexical adjuncts, independent of a well-formed sentence, and as such, their absence does not make a sentence ungrammatical. However, removing discourse markers does eliminate useful information about the speaker’s stance towards the relationship between the current utterance and prior discourse.

Listeners have been shown to be sensitive to the removal or substitution of discourse markers in multiple languages (Groen & Noyes, 2010). For example, Native English listeners showed faster recognition of words preceded by discourse marker *oh* than in conditions where *oh* was either elided completely, or replaced with a pause (Fox Tree & Schrock, 1999). Second language learners of English displayed better comprehension of a video-taped lecture when the lecture contained naturally occurring discourse markers, as compared to those who viewed the same lecture with the discourse markers elided (Flowerdew & Tauroza, 1995). An overview of work on discourse marker *like* specifically can be found in section 2.2.2.1.

2.2.1.2. Discourse particles
While some studies treat the marker and the particle as one and the same, D’Arcy draws a distinction between the two, noting that in addition to appearing in different positions (clause initial vs medial), the marker serves to signpost adjacent content, while the particle functions to “focus, highlight, evaluate, and otherwise draw the speaker and hearer together (2017, p. 57)”. In contrast to the marker, discourse particles occur clause medially, as shown in Example (1). Example (1) and other Pidgin examples shown herein are excerpts taken from the Pidgin and Hawai‘i English
interviews discussed in Section 4.3

Example 1: Discourse particle *like*

It wasn’t till later on the next day they found the bugger. He died. He got off was *like* over one cliff.

Lani, woman from O‘ahu, born 1958

2.2.1.3. Quotatives
Quotatives are used to introduce reported speech, thought, or gesture, and are perhaps the most studied discourse-pragmatic element. Beginning with Butters’ notable study (1980), much attention has been paid to rapid global changes in the quotative system of English, with a particular focus on quotative *be like* (see Tagliamonte, D’Arcy, & Rodríguez Louro 2016 for an overview). Innovative quotatives, such as *be like, be all*, and *go* have attracted attention (mostly negative) in popular media and discourse, due in part to their association with youth and teenagers (Buchstaller, 2013). Apart from these new quotatives, in English, *say* and *think* are the most common quotatives worldwide. In addition to these common quotatives, a variety of lower frequency quotatives are produced by speakers, such as *shout, scream, sing, whisper, ask* and others (Buchstaller, 2013). Some speakers may also produce a null, or “zero” quotative to introduce reported speech, often using prosodic cues to indicate that the utterance is not their own (D’Arcy, 2010; Buchstaller, 2013) (see section 2.3.2.2. for an example of the null quotative). A detailed discussion of quotative *be like* can be found in section 2.2.2.3.

2.2.1.4. Approximate adverbs
Approximate adverbs precede numerical expressions and quantifiers, and signify an approximative meaning (Schourup L., 1985; Underhill, 1988; Miller & Weinert, 1995; D’Arcy, 2006). In addition to the approximate adverb *like*, approximate adverbs such as *almost* and *barely* contribute to textual coherence, and have also received attention in the literature (Amaral, 2010).

2.2.1.5. Sentence adverbs
Sentence adverbs, such as *about* and *like* function to indicate the truth value of propositions. Speakers utilize sentence adverbs to assign varying degrees of reliability to sentences (Michell, 1976).

2.2.2. Discourse-pragmatic *like*
Multiple types of *like* are used in English. The grammatical functions, such as the noun, verb, and preposition have been attested in the grammar for quite some time. Others, such as the discourse-pragmatic functions (quotative, approximate adverb, sentence adverb, discourse marker, and discourse particle; described in sections 2.2.2.1–2.2.2.3) are of particular interest to sociolinguists. Despite popular ideologies to the contrary (see section 2.2.3), discourse-pragmatic *like* is not a singular lexical item that can
appear randomly at any point in a sentence (D'Arcy, 2007, p. 388). Instead, it is versatile and flexible, serving discourse-pragmatic functions such as (but not limited to) approximation, focus, signaling subjective information, introducing reported speech or thought, introducing new information, or introducing/qualifying a proposition (Miller & Weinert, 1995; Underhill, 1988; Andersen, 1998; D'Arcy, 2017). While not all work on like draws a distinction between all the discursive elements listed herein, this dissertation follows D'Arcy's analysis of treating these features separately because it allows for a more fine-grained analysis and interpretation of how these forms are used. There are five types of discourse-pragmatic like: approximate adverb, sentence adverb, discourse marker, discourse particle, and quotative. These are described in more detail below.

The rise of discourse-pragmatic like has been well documented for several decades (Butters, 1982; Underhill, 1988; Andersen, 1998; Daily-O'Cain, 2000; D'Arcy, 2005; Buchstaller, 2006; see D'Arcy 2017 for an overview). While discourse-pragmatic like is multifunctional and at times ambiguous (Buchstaller, 2006), the various functions of like are interrelated and relate to a speaker's epistemic stance.

Multiple models have been proposed for the grammaticalization of like in its multiplicity of forms (Romaine & Lange, 1991; Buchstaller, 2006; D'Arcy, 2007; D'Arcy, 2017). The rapid global expansion of like (with regard to the quotative usage in particular) has brought some scholars to advocate for a rethinking of the underlying assumptions of existing models of variation and change (Buchstaller, 2006; Tagliamonte et. al., 2016). The fact that multiple models and potential trajectories of grammaticalization and distribution exist is a testament to the fact that the rise of discourse-pragmatic like is unprecedented in the recorded history of language, and that gathering more data from diverse geographic locations will contribute to shaping theories of language variation and change.

2.2.2.1. Discourse marker like

Example 2: discourse marker like

Like, here da wada da wada so deep, ok I goin’ hold this net and da odda guy going walk em out.

Keoni, man from Hawai‘i Island born 1967

Discourse marker like appears in clause-initial position, and serves an interactional need (Fuller, 2003), particularly as a marker of focus (Romaine & Lange, 1991; Underhill, 1988; Andersen, 1998). D'Arcy (2017) demonstrates that like as a discourse marker developed from the sentence adverb and has been attested in historical materials from the eighteenth century onwards. While discourse marker like is slightly less frequent than other discourse markers among older speakers, this trend reverses with speakers born in the 1970s, with like becoming the dominant discourse marker found in the system. This recent increase in use, combined with popular ideologies surrounding like (see section 2.2.3) makes discourse marker like a particularly
interesting subject of study.

2.2.2.2. Discourse particle *like*

**Example 3: discourse particle *like***

Because a dinosaur a billion years ago *like* chewed this leaf a certain way, I’m eating Raisin Bran.

(D’Arcy, 2017, p. 228)

Discourse particle *like* is one of the most commented upon forms, by linguists and the general public alike. Structurally, it differs from the marker (section Discourse marker *like* 2.2.2.1) in that the marker occurs in clause-initial position and connects the following utterance to preceding discourse, while the particle targets clause-internal positions and serves to mark stance\(^2\) or focus.

Underhill (1988) posits that discourse particle *like* developed from the approximate adverb *like*. In contrast, D’Arcy’s work provides evidence that the particle developed from the marker, and posits that both were distributed globally by British emigrants (2017, p. 58). While neither the marker nor the particle are new, both have recently undergone a period of rapid expansion. Both forms, particularly the particle, are frequently commented upon metalinguistically, and are often associated with youth and inarticulate speech. Although the discourse particle can serve a textual function, and occasionally co-occurs with false starts and hesitations, this does not mirror the wider distribution of the particle. While popular ideologies *link* discourse-pragmatic functions of *like* with women (Daily-O’Cain, 2000; D’Arcy, 2007; D’Arcy, 2017), the discourse particle is more frequently used by men (D’Arcy, 2007). Discourse particle *like* serves to mark focus (Underhill, 1988) and epistemic stance (D’Arcy, 2017).

2.2.2.3. Quotative *like*

**Example 4: quotative *like***

They said something and I’m *like*, “Oh god, you’re from Maui aren’t you.”

Abcde, woman from O’ahu born 1988

Since the 1980s, research on quotatives has frequently focused on quotative *be like* (Butters, 1982). Quotative *like* has been identified in English varieties worldwide; in the United States (Blyth, Recktenwald, & Wang, 1999; Daily-O’Cain, 2000; Kohn & Franz, 2009; Buchstaller & D’Arcy, 2009; Barbieri, 2009), Canada (Tagliamonte & Hudson, 1999; Tagliamonte & D’Arcy, 2004; 2005; 2007), the United Kingdom (Tagliamonte & Hudson, 1999; Tagliamonte & D’Arcy, 2004; 2005; 2007), Australia and New Zealand (Buchstaller

\(^2\) Stance can be broadly defined as how an individual positions themselves towards ongoing interaction epistemically, interactionally, or socially.
and D’Arcy 2009; D’Arcy 2010, 2012), and the Caribbean (Höhn, 2012; Deuber, 2014; Bogetić, 2014) and appears to have entered these systems at a similar time worldwide (Tagliamonte et. al., 2016). This entry into the quotative system of English happened rapidly and uniformly; with be like patterning similarly across studied varieties in terms of grammatical constraints such as person, number, and content of the quote (Tagliamonte et. al., 2016).

Romaine and Lange (1991) proposed that quotative like developed from the conjunction; like can occupy the syntactic position preceding a clause, and thus can precede a quotative. From this quotative complementizer phase, like comes to be associated with quotation itself. D’Arcy (2007; 2017) asserts that the discourse marker existed long before the quotative and was in fact the source of the quotative. Historically discourse marker like co-occurred with be, eventually undergoing re-analysis as quotative be like. Buchstaller (2006) theorizes that grammaticalization does not progress unilaterally but is more accurately characterized by a semantic field of related and overlapping meanings. Under this model, the quotative and other discourse-pragmatic functions arose due to a shared core meaning of “similarity”. Regardless of the exact mechanism by which it spread, the rise of quotative like has been described as “possibly the most vigorous and widespread [documented] change in the history of human language” (Tagliamonte S. A., 2012, p. 248). While it is possible that other changes of this scale have simply gone undocumented before, the rise of be like is the fastest and most widespread change that has been documented in real time.

2.2.2.4. Approximate adverb like

Example 5: approximate adverb like

I get some wild pigs that I wen catch when we go hunt one time from baby. I catch one small laidis, and den was in da forest we catch was only like twenty pounds.

Arika, man from Hawai’i Island born 1986

In its function as an approximate adverb, like carries an approximative meaning (Underhill, 1988). While like as an approximate adverb has existed in English since the first third of the twentieth century, it has recently undergone a rapid expansion in use (D’Arcy, 2006). For example, approximate adverb like has quickly become the preferred variant for young speakers of Toronto English, almost completely replacing about (D’Arcy, 2006). Approximate adverb like alternates with about, approximately, and roughly, and precedes quantifiers and numerical expressions.
2.2.2.5. Sentence adverb *like*

**Example 6: sentence adverb *like***

And every theater I think had their own mascot *like*.

Grant, man from O'ahu born 1951

Sentence final adverb *like* has long been a part of English grammar, and serves a pragmatic function; anticipating and countering objections and assumptions (Miller & Weinert, 1995). It is likely that the sentence adverb developed from the conjunction (D'Arcy, 2005, p. 206). While other discourse-pragmatic forms of *like* are found in a wide variety of Englishes, and are undergoing expansion, sentence adverb *like* is rare, and appears to be undergoing recession (D'Arcy, 2017). Sentence adverb *like* alternates with ‘as it were’ or ‘so to speak’.

2.2.3. Ideologies around *like*

Despite the complexity of the empirical picture around *like*, multiple forms of *like*, including quotative *be like*, are often popularly associated with the speech of young women, (Daily-O'Cain, 2000; D'Arcy, 2007). The attitudes around *like* are complex, with listeners assigning both negative and positive evaluations to *like* (Daily-O'Cain, 2000).

Prevailing attitudes towards *like* are clearly negative (Daily-O'Cain, 2000; Buchstaller, 2013; D'Arcy, 2007; D'Arcy, 2017). That *like* is associated with negative characteristics is not entirely surprising, given the fact that women’s speech is more likely to be policed than men’s, and evaluated more negatively (Eckert P. M.-G., 2003; Lakoff, 2004; Bucholtz, 2014). While popular ideologies hold that women use *like* more than men, the empirical evidence does not bear this out (D'Arcy, 2007; Daily-O'Cain, 2000).

D'Arcy (2007) explores several of these ideologies with respect to *like*, finding that speakers perceive *like* to be a meaningless lexical item that can be used anywhere in a sentence, and associate it with youth, and women in particular. She notes a number of reactions to the use of *like*, including the opinion that it is a “meaningless tic” which reflects poorly on the user, causing them to appear “less educated, intelligent, or interesting” (2007, p. 388). In particular, D'Arcy examines a set of beliefs she calls “The like language myth”:

> *Like* is just *like*, that is, there is one *like* that is recycled repeatedly.  
> *Like* is meaningless; it simply signals a lack of articulacy.  
> Women say *like* more than men do.  
> *Like* began with Valley Girls.  
> Only young people, and adolescents in particular, use *like*.  
> *Like* can be used anywhere in a sentence.”  

(2007, p. 388)
D’Arcy also notes that the media is often a source of these language myths, and the association between *like* and younger speakers appears to be constant throughout the English-speaking world. The associations between *like* and women, and *like* and the United States vary in saliency worldwide (D’Arcy, 2007).

Additionally, experimental approaches provide evidence for these ideologies, demonstrating that participants rate *like* negatively in terms of status-based (i.e. educated, intelligent) criteria. A matched-guise study demonstrated that participants had “an abundance of very strong negative opinions [about *like*]” (Daily-O’Cain, 2000, p. 69), and when answering an open ended questionnaire, participants associated *like* with sounding “uneducated or lazy”, “Californian ‘Valley Speak’”, and “imitating an airhead” (p. 70). Multiple studies have demonstrated that participants associate *like* with negative traits of the California “Valley Girl” stereotype, such as being meaningless, inarticulate, and youthful (Daily-O’Cain, 2000, p. 69; D’Arcy, 2007, p. 388). However, this work also paints a more complex picture, since *like* is also evaluated positively in terms of solidarity-based criteria (i.e. cheerful, friendly) (Daily-O’Cain, 2000).

Following Daily-O’Cain’s work, Buchstaller (Buchstaller, He goes and I’m like: The new Quotatives re-visited, 2006) investigates perceptions of quotative *be like* in the British Isles. Using a combination of modified matched-guise and a social attitudes survey, Buchstaller finds that covert and overt associations towards *like* vary greatly. Participants were exposed to a written matched-guise task, containing either quotative *say*, or quotative *be like*. After viewing the text, participants were asked to provide their perceptions of the speakers’ age, gender, and social class, as well as a variety of personality traits. After completing the task, participants were then asked to complete a social attitudes questionnaire, in order to assess their overt attitudes towards *like*.

Results showed that participants strongly associate *like* with younger speakers, a result which is in line with Daily-O’Cain’s findings (2000). Buchstaller’s respondents from the British Isles held no associations between *like* and gender or social class. With regards to the social attitudes survey, *be like* is perceived to be a feature of working-class female speech. In terms of personality traits, the *be like* guise was rated as less ambitious and less educated, but also judged to be more animated, giddy, and trendy. This finding that *be like* use is associated with both positive and negative traits is similar to Daily-O’Cain’s findings. Buchstaller concludes that “we cannot assume that the same perceptual information is associated with global features in different localities.” (2006, p. 375). This finding in particular sets the stage for the investigation of these global variants in the local context of Hawai’i.

Further matched-guise work has also shown that the more recent a function of *like* is in the grammar, the more socially salient (and commented on) it is (Maddeaux & Dinkin, 2017). Participants were each exposed to six guises: a control guise, not containing *like*, three unrelated fillers, and two guises containing *like*. The *like* guises contained ten tokens of a particular type of *like*; either grammatical (verb, preposition,
or comparative) or vernacular (discourse marker, vP initial discourse particle, NP-initial discourse particle, or approximate adverb). Participants rated the speaker on eight traits (friendliness, intelligence, politeness, articulateness, youth, interestingness, confidence, and femininity) along a five point Likert scale. The results do not support the hypothesis that listeners evaluate all functions of like the same, but do provide evidence that listeners perceive differences between the socially noticeable “vernacular” and non-noticeable “grammatical” categories of like. These findings provide nuance when considered along D’Arcy (2007), who finds that in popular ideology, like is conceptualized as monolithic, and highlight the need for ongoing research into the complex ideologies surrounding like.

2.3. Language in Hawai‘i
Hawai‘i’s ethnic diversity, history of contact, and geographic isolation make it a particularly interesting location to study sociolinguistic variation and language change. Within the relatively short time frame of 150 years, Hawai‘i was forcibly changed from an independent, monarchy ruled Pacific Island nation into the still disputed 50th U.S. State (Sato, 1993; Sakoda & Siegel, 2003; Lockwood & Saft, 2016).

During the 1870s, an influx of imported laborers began to arrive from China, Japan, Portugal, and the Philippines, brought by the predominantly white, English-speaking plantation owners to work the plantation fields (Sato, 1993; Sakoda & Siegel, 2003; Lockwood & Saft, 2016). From this diverse labor force, an early, Hawaiian-based pidgin began to emerge (Sato, 1993; Sakoda & Siegel, 2003; Lockwood & Saft, 2016). Eventually, as the white, English-speaking Americans took over economic and political control of the islands (often through force), English became the dominant language, and so the pidgin became more heavily English-lexified, eventually undergoing creolization and stabilizing into what is today known as Pidgin, or Hawai‘i Creole (Sato, 1993; Sakoda & Siegel, 2003; Lockwood & Saft, 2016).

The history of education in Hawai‘i is tied together with language and ideologies (see section 2.3.1.1 for a discussion of ideologies). Prior to the arrival of missionaries in the 1820s, separate systems of education existed in Hawai‘i. The children of the ali‘i (royal families) and the children of the maka‘āinana (commoners) received separate educations (Bayer, 2009). Missionaries arriving in the 1820s were primarily concerned with educating the ali‘i, with a particular focus on literacy (Bayer, 2009). This system of separate schools for separate groups continued, with missionaries first homeschooling their children, and later establishing elite private schools, such as Punahou, for the benefit of their own children. The language of instruction for these elite schools was English, while most children received education in Hawaiian (Bayer, 2009).

By the 1920s the majority of children attending public school were the non-Caucasian, Pidgin-speaking children of plantation workers (Bayer, 2009). These ethnic and linguistic differences led to the creation of the English Standard system of schools. As
the European population of Hawai'i grew, those who could not afford to send their children to private school were hesitant to send them to public school with the Pidgin speaking children of laborers. Children who passed an admissions test and oral interview in Standard American English were admitted to English Standard schools. Children who spoke Pidgin were denied entry into the English Standard schools and instead attended public school. This system of separate schooling served to keep children of different ethnic, socioeconomic, and linguistic groups apart, and likely encouraged elitism. The last of the English standard schools were phased out by 1960, and all public schools were required to achieve similar standards (Bayer, 2009). While English Standard schools no longer operate in Hawai'i, many people educated in that system are alive today and there remains a sharp, and socially salient divide between public and private schools that is linked with classism and ideologies around Pidgin and English (see section 2.3.1.1 for further discussion of language ideologies in Hawai'i).

While Hawai'i is geographically isolated, the primary economic sectors are tourism and the military, and as a result many Locals have ongoing contact with visitors from the United States and around the world. Many Locals are bilingual in Pidgin and English, and code-switching is common (Drager, 2012). Therefore, in addition to contact with global Englishes, there is the additional factor of speaker-internal language contact between Pidgin and English in Hawai'i. This makes Hawai'i a particularly interesting place to explore language variation and change in the context of language contact.

2.3.1. **Ideologies**

2.3.1.1. **Sociolinguistic ideologies in Hawai'i**

Ideologies surrounding language and education are part of the community discourse in Hawai'i (Bayer, 2009). There is a prevailing attitude that while Pidgin is a widely used language in Hawai'i, linked with localness (Da Pidgin Coup, 1999; Marlow & Giles, 2008; Drager & Grama, 2014; Drager, 2012), “There’s a time and place for it” (Lee, 2017). That is, there is a perception that Pidgin is not a language that can be used for socioeconomic advancement, or even in certain social situations such as school or professional environments.

The popular discourse surrounding Pidgin extends into educational contexts (Da Pidgin Coup, 1999; Bayer, 2009). The predominant attitude in Hawai'i is that public schools are failing, and private schools are succeeding, and the long history of separate education for separate groups in Hawai'i may have encouraged elitism in the public discourse surrounding public/private school choice in Hawai'i (Bayer, 2009). Common ideologies link Pidgin with public schools, and English with private schools in Hawai'i (Bayer, 2009). Pidgin has even been accused of causing poor performance of Hawai'i students on standardized writing tests (cf Da Pidgin Coup, 1999). Given these perceptions, many parents feel that in order for their children to succeed in school, speak English, and go to college, private school is the logical choice (Bayer, 2009). Given these facts, we might assume an ideological relationship between perceived education and language variety.
This association was confirmed in a matched-guise study, in which University students in Hawaii rated Pidgin higher in terms of solidarity-based traits (i.e. friendliness, warmth), but lower in terms of superiority-based traits (i.e. education, class) (Ohama, Gotay, Pagano, Boles, & Craven, 2000). Additionally, participants with higher Pidgin skills were more likely to rate Pidgin higher in terms of solidarity-based traits than participants with lower Pidgin skills; no interaction of participant language ability and superiority-based traits was found.

2.3.2. Discourse-pragmatic like in Hawaii

2.3.2.1. Quotatives in Hawaii English

In previous work, we find that quotative use in Hawaii English patterns similarly with previously studied North American varieties (Schutz, Chun Comstock, Stabile, & Drager, in prep). For older speakers, quotative say is the most common quotative compared with other varieties. We find that quotative tell is more frequent in Hawaii English, found at 9% among older speakers (Drager et. al., in prep); quotative tell is generally so infrequent in English systems that most researchers include it in the “other” category when conducting analysis. That quotative tell is relatively frequent in Hawaii English is potentially explained by the high frequency (14%, among young and middle aged speakers) of tell in Pidgin (Schutz et. al. (in prep)). Examples of each are shown in Example 7 below.

Example 7: The five most frequent quotatives in Hawaii (Drager et. al., 2016)

Say: And Tutu Man would say, “Chee, this lady, what’s the matter with her?”

Be like: Some freaky things kind of started happening, and they’re like, “I think we peed on like some sacred ground or whatever.”

Go: He goes, “Cause these guys are gonna beat you when you get outside.”

Null: “Did you do your homework?” 0 “Yes.” 0 “Let me see your homework.” No homework. 0 “Stand in the corner for a while.”

Think: And he’s thinking, “She needs to shut up already.”

Also notable is the relatively high rate of the null, or zero quotative, in Hawaii English (20% in middle aged speakers). This contrasts with work on most other varieties, which report low rates (<5%) rates (Buchstaller, Rickford, Closs Traugott, Wasow, & Zwicky, 2010). The only study to report higher rates of the null quotative than what has been observed in HE was conducted on Maori English (24% among younger speakers and 42% with older speakers (D’Arcy, 2010)). As with quotative tell, the relatively high rates of the null quotative in Hawaii English may be explained via transfer from Pidgin.

---

3 the highest rate reported in the literature is 8%, among a group of Latino speakers studied by Kohn and Franz (2009)
Drager et. al. (in prep) found that quotative *be like* entered the system of Hawai’i English at the same time that it entered the systems of other studied English varieties worldwide. Quotative *be like* is common among middle aged speakers, and dominates the system for younger speakers of Hawai’i English. Like other studied varieties, quotative *be like* is most strongly associated with first person, the historical present tense, as well as internal thought. Unlike previously studied varieties however, *be like* is not favored by the present tense in Hawai’i English. Additionally, we find no gender difference in the use of *be like* in Hawai’i English.

2.3.2.2. **Quotatives in Pidgin**

The five most frequent quotatives in Pidgin are *go*, *say*, *tell*, *be like*, and the null quotative. However, no particular quotative dominates the system for either young or middle-aged speakers. Compared to previously studied varieties of English, rates of *tell* and the null quotative are relatively high (14% and 24%, respectively). Despite the fact that quotative *be like* entered the system of Hawai’i English concurrently with other varieties, it took roughly one generation later for quotative *be like* to enter the system of Pidgin, and *be like* is used only by the youngest studied speakers (born between 1983 and 1988). Indeed, it appears that *be like* is becoming the preferred quotative for young speakers of Pidgin. After the entrance of quotative *be like* into the system of Pidgin, we observed lower rates of other quotatives, with *go* and *tell* having the lowest rates among the youngest cohort of speakers (Drager et. al., in prep.). Additionally, we observed no gender-based differences in the use of quotative *be like* in Pidgin (Drager et. al., in prep.). Also of note is the fact that *be like* retains its original English morphology when used in Pidgin.

Given that quotative *like* in Hawai’i English patterns similarly to other studied varieties of English, we might expect to see a similar peak in use of the discourse marker among speakers born after the 1970s. Predictions for Pidgin are less clear, given that quotative *like* took roughly a full generation more to enter the system of Pidgin. Will the discourse-pragmatic functions of *like* pattern similarly in Pidgin as in Hawai’i English? Or, as we found with quotative be *like*, will the system of Pidgin lag a generation behind that of Hawai’i English in the expansion of discourse-pragmatic *like*? How do the users of these forms perceive them when used in a local context? This dissertation addresses these questions through employing a corpus-based analysis of archived sociolinguistic interviews conducted in Pidgin.
3. Perception of quotative *like* in Hawai‘i

3.1. Introduction
Following the work outlined in Chapter 2, this chapter presents results from a matched guise experiment designed to explore the following research questions.

1. Given that *be like* is relatively new in Pidgin, and *tell* has high rates in Pidgin (Drager et. al., in prep), do Hawai‘i Locals perceive a link between quotative verb and the language variety being voiced in the quote? Higher localness ratings are expected when Pidgin is paired with *tell*, than when paired with *like*, and English paired with *like* are expected to produce the lowest localness ratings.

2. Popular discourse generally associates *like* with women (Daily-O’Cain, 2000; D’Arcy, 2007). Is that the case in Hawai‘i? Higher femininity ratings are expected in the *like* condition than in the *tell* condition.

3. Previously, Pidgin has been rated lower in terms of status-based criteria, and higher in terms of solidarity-based criteria (Ohama, Gotay, Pagano, Boles, & Craven, 2000). Do these ideologies persist nearly two decades later, and might there be an interaction with the quotative? Since *like* is generally perceived negatively (Daily-O’Cain, 2000; D’Arcy, 2007), lower ratings for education are expected with *like*.

3.2. Methods
3.2.1. Conditions
Two quotative conditions were included in the study: *be like* and *tell*. These conditions were crossed with two language conditions: Pidgin and English.

3.2.2. Stimuli
3.2.2.1. Critical Items
Eighteen unique critical items were created which consisted of a sentence frame, quotative verb, and the following quoted material. In all critical items, the sentence frame portion of the stimuli consisted of Hawai‘i English, and the quoted portion was either in Pidgin or Hawai‘i English, depending on the language condition of each item. Example sentences are as shown in Example 8 below.

Critical items were created by splicing the quotative verb token and quoted speech portion into identical sentence frames, creating identical lead-in sentences as in 1a-1d below. The same quotative tokens were used across the different language conditions, resulting in identical sentence frames (e.g. 1a and 1c below). Additionally, the prosody and phonetics of the quoted material were consistent with the structural and lexical aspects of the language variety being voiced in the quote. Critical items were matched
for content across language variety (as displayed in Example 8 below).

Example 8: sample critical items

a. **Pidgin + like:** I saw her at the store yesterday and she was *like* "ho, sista! where you was on Friday? you neva come practice."

b. **Pidgin + tell:** I saw her at the store yesterday and she *told* me "ho, sista! where you was on Friday? you neva come practice."

c. **English + like:** I saw her at the store yesterday and she was *like* "girl, where were you on Friday? you didn't come to practice."

d. **English + tell:** I saw her at the store yesterday and she *told* me "girl, where were you on Friday? you didn't come to practice."

3.2.2.2. Control items

Nine Pidgin and nine English control sentences were used in the experiment. Control items were recorded by the same speaker who produced the critical items. Unlike the critical items, control items were not matched for content across language variety. Control sentences consisted of tokens overheard in the natural speech of native speakers of Pidgin and Hawai‘i English, and contained no quoted speech. Having items with no quotative material allows comparison between sentences with and without quotatives, while holding other factors, such as language condition and speaker, constant.

Example 9: sample control items

a. **Pidgin:** I wen try tequila, dose guys had to pick me up cause my legs no more work.

b. **English:** We were hiking last weekend and we saw three baby pigs, they were so cute!

3.2.2.3. Distractor items

Nine Pidgin and nine English distractor sentences were created for the experiment. Distractor items resemble the control items but were produced by novel speakers. Having a range of voices for the distractor items served both as a check that participants were sensitive to, and using, the full range of the rating scales, as well as to distract them from the fact that the critical items contained quotatives. As with the control items, all distractor items consisted of sentences overheard in natural conversation and did not contain any quotatives.

Example 10: sample distractor items

a. **Pidgin:** Eh, 99 percent of us is portagee right now brah! aes why no can!

b. **English:** My mom makes the best lasagna ever, you have to try it some time.

3.2.2.4. Voices used in the experiment

Critical and control sentences were recorded by a trained linguist who is a native speaker of both Hawai‘i English and Pidgin. Prior to recording the stimuli for the
experiment, a norming study was conducted.

Distractor items were recorded by three non-native speakers of Pidgin; speaker 4, a New York native who had just arrived in Hawai‘i at the time of recording, speaker 3 who had been living in Hawai‘i for 5 years at the time of recording, and speaker 2 who had arrived in Hawai‘i at age 8, and maintained residence for 19 years.

3.2.3. Procedure
From the bank of sentences created for the experiment, four running lists were created using a Latin square design. Items were balanced between running lists so that individual participants were not exposed to the same sentences in more than one guise. Each participant rated items from a single running list, consisting of eight critical items (containing quotatives), five control items, and thirteen distractor items, for a total of 26 items. Each participant was exposed to two each of the four possible combinations of quotative verb and language variety (Pidgin+like, Pidgin+tell, English+like, English+tell). Additionally, each running list contained only one production of each distractor item (i.e. participants were not exposed to the same item produced by different speakers). Each participant was exposed to thirteen items produced by speaker 1 (8 critical items and 5 controls), and thirteen distractor items produced by the novel speakers (speakers 2, 3, and 4, described in section 3.2.2.4). The statistical models used for analysis are described in more detail in section 3.4.2.

3.2.3.1. Presentation of stimuli
The experiment was implemented via PsychoPy version 1.84.2 (Peirce, 2017). Participants used a laptop computer and noise cancelling headphones to complete the task. Participants were instructed to listen to each sentence and click on the rating scale using their first intuition. Participants were also instructed that they could not listen to any sentence more than once. Participants rated each token on six social traits: education, friendliness, femininity, localness, authenticity, and naturalness. These traits were chosen to follow similar matched-guise studies investigating ideologies towards quotatives and language in Hawai‘i (Daily-O‘Cain, 2000; Ohama, et al., 2000).

Participants heard each sentence one at a time, while viewing a screen containing six equal sized, five-point rating scales, arranged in two rows on the screen. Placement of the scales and left-right orientation did not vary across trials; this was done in order to facilitate participants moving more quickly through the trials (see Figure 1 for an example of scale placement on-screen). Traits for rating were shown on either side of each rating scale and participants used the mouse to click on the scale to record their rating. Each scale disappeared from the screen once clicked, ensuring that participants did not second-guess their intuition, and were aware of the items they had completed before moving on to the next token. After listening and rating, participants clicked the space bar in order to hear the next sentence.
3.2.4. Participants
Participants were recruited through the University of Hawai‘i’s Linguistics Beyond the Classroom program, as well as through personal contacts of the researcher. All participants were born and raised in Hawai‘i, and had not spent more than three months living outside of Hawai‘i. Participants were born between 1983 and 1999; with a mean age of 24.7; similar to the mean age of 24 reported by Ohama et. al. (2000). Participants from this age range are consistent with the ages of the participants who produced quotative be like in the Drager et. al. (in prep) study. After excluding 2 participants who answered five for all items (i.e., reporting that even the New Yorker sounded Local), as well as 3 whose participation was impacted by a software error, the final data set for this analysis contains forty-two speakers.

After participants completed the task, demographic information was collected via a short survey. Participants provided their birth year, gender identity, the neighborhood they grew up in, the high school(s) they attended, any languages spoken, any time spent outside of Hawai‘i, their self-reported identity as a speaker or comprehender of Pidgin, as well as their occupation.

3.3. Analysis
3.3.1. Coding participant demographics
In order to conduct statistical analysis, the demographic data collected from
participants was organized into several meaningful categories.

While participants were asked to provide their gender, pronouns, and sex assigned at birth (in line with recommendations by Zimman, 2018), all participants indicated they were cisgender, 24 of whom identified as women, and 18 identifying as men.

Participants were categorized into two groups based on their self-reported Pidgin ability; all participants reported having some ability to understand Pidgin, although many participants hedged when it came to talking about their speaking ability, saying for example; “oh yeah, I understand Pidgin, but...I don’t really speak it.” Participants who hedged in this manner were categorized as comprehenders rather than speakers.

As discussed in chapter 2, in Hawai’i, there is a prevailing community-wide belief that private schools are better than public schools, and that private schools emphasize “proper speech” (Bayer, 2009). In order to investigate this association, participants were categorized into four groups, based on the type of high school they attended; public, private, charter, and GED. Since private schools are ideologically linked with a more “standard” speech variety, participants who attended these schools may well see themselves as the arbiters of the standard language. During analysis, it was determined that Pidgin speaker identity co-varied with high school type; most participants who self-identified as Pidgin speakers graduated from public school, rather than private or charter school. Because of this, high school type was removed from analysis.

Participants were categorized into eight moku, or land districts, based on the neighborhood they grew up in. The moku used in the study are the eight used by Drager and Grama (2014) to analyze beliefs about links between language use and region; these consist of the seven original moku made on the island of O’ahu during ancient times, plus the addition of the modern land division of Central O’ahu for the purposes of analysis. In addition to the moku described in Drager and Grama (2014), the current study added two more moku for the purpose of participant categorization: Neighbor Island, to include participants from neighboring islands, and Other, to include a single participant who moved so frequently during their youth that they were unable to identify a single hometown when asked. However, the distribution of participants across the moku groups was uneven (due to the small sample size), so this factor was removed from analysis.

3.3.2. Data preparation and coding
All predictors for the models were coded using deviation coding. Deviation coding functions by comparing the mean of each level of a variable to the grand mean of that variable (UCLA). This is a good choice for this type of sociolinguistic work, in that it essentially compares a subset of the population to the general population. For example, we can compare participants who attended public schools to participants from all school types using this method. When interpreting the results of a model
using deviation coding, the intercept is the grand mean, rather than the baseline condition (Politzer-Ahles, 2015). In this particular study, treating the data with deviation coding did not produce meaningfully different results from treating the data with dummy coding.

3.4. Results
Statistical analyses were computed via R (R Core Team, 2017), using the clmm function of the ordinal package (Christiansen, 2015). Data tidying was performed using the dplyr package (Wickham, 2017). Data visualization was plotted using the Likert package (Bryer & Speerschneider, 2016) with supplemental visualization via the ggplot2 package (Wickham, 2009).

3.4.1. Analysis of all stimulus voices
For all graphs, the percentage of responses was plotted along the x axis, and each condition is shown on the y axis. The right side of the graph shows the percentage of positive responses (4s and 5s on the rating scale), while the left side shows the percentage of negative responses (1s and 2s). The graphs are centered on the percentage of neutral responses (3s, shown in light grey).

Results of naturalness and authenticity ratings were plotted for all four voices used in the study. The results (Figure 2) show the percentage of positive (on the right) and negative responses (on the left). It is evident from the graphs that participants rate each voice as being more local relative to the length of time the speaker has spent in Hawai‘i. These results demonstrate that as a group, participants are sensitive to, and make use of, the entire range of the rating scales. This verifies that the speaker who produced the critical items (s1, shown at the bottom of the graph) has appropriate vocal qualities for the task.

Figure 2: plot comparing localness ratings across all stimulus voices for all participants in Pidgin and English conditions
3.4.2. Social ratings
Multilevel ordinal logistic regression models were fit to the subset of the data containing the critical and control items. The response to each of the social factors (localness, awkwardness, naturalness, femininity, friendliness, and education) was set as the dependent measure for each model, in separate analyses for each question. In each regression model, social factor ratings were specified as ordinal variables, which consisted of the five response categories shown on the rating scale in PsychoPy (Figure 1). All models included language condition, quotative type (like, tell, or control), participant gender, and Pidgin speaker identity as fixed effects. All models were designed with a maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013) which included by-participant slopes of the interaction between quotative type and language variety as well as random intercepts of item nested inside running list; many items were unique to specific running lists. In all models, the levels for the variable were set the same, as displayed in Table 1. All models tested for an interaction between language condition and item type. Models were built using simultaneous entry of all planned predictors, and predictors which did not reach significance were not trimmed from the models, since the aim of these models is confirmatory (following Winter, Rentz, & Roettger, in prep). Furthermore, removing non-significant predictors from the models did not result in a change in significance of predictors. No factors reached significance in either the model of naturalness or the model of education, so these models will not be discussed further.

Table 1: ordering of levels for factors included in ordinal regression models

<table>
<thead>
<tr>
<th>Factor: Language condition</th>
<th>Item type</th>
<th>Gender</th>
<th>Pidgin identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pidgin</td>
<td>like</td>
<td>women</td>
<td>speaker</td>
</tr>
<tr>
<td>English</td>
<td>tell</td>
<td>men</td>
<td>comprehender</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.4.2.1. Authenticity
Figure 3 compares ratings for authenticity in both the Pidgin and English language conditions. It is apparent from the graph that participants find the speaker to be equally authentic in both language conditions, and in the statistical model of authenticity ratings, no factors reached significance, so authenticity will not be discussed further.
3.4.2.2. **Localness**

Figure 4 compares localness ratings for all three conditions (like, tell, and control items) in both Pidgin and English. While the percentage ratings for localness are similar in Pidgin and English, it is evident from the graph that Pidgin receives more ‘5’ ratings for localness than English. This is also reflected in a statistical model of localness ratings (Table 2).

Table 2 shows the model output for localness ratings. The Pidgin language condition was found to be a significant predictor of higher localness ratings (p<0.001), as was self-reported Pidgin speaker identity (p<0.05). There was no interaction effect of language condition and quotative verb.
Table 2: model fit to localness ratings:

local.clmm <- clmm(localness.response ~ lang.cond * quot.type + identity + gender + (1 + quot.type / lang.cond | participant) + (1 | list/item), Hess = T, threshold = "flexible", link="logit", data = data)

|                         | Estimate | Std. error | Z value | Pr (>|z|) |
|-------------------------|----------|------------|---------|----------|
| Pidgin condition        | 1.65179  | 0.39102    | 4.224   | p<0.001  |
| like                    | -0.28487 | 0.37494    | -0.76   | 0.4474   |
| tell                    | -0.0146  | 0.34838    | -0.42   | 0.672    |
| Participant gender - f  | -0.69955 | 0.38143    | -1.834  | 0.0667   |
| Identity - speaker      | 0.90142  | 0.38566    | 2.337   | p<0.05   |
| Pidgin : like           | -0.32128 | 0.46915    | -0.685  | 0.4935   |
| Pidgin : tell           | 0.05953  | 0.34832    | 0.171   | 0.8643   |

3.4.2.3. Femininity

In a plot comparing femininity ratings in Pidgin and English (Figure 5), it is apparent that participants rate *be like* as feminine in both language varieties. Quotative *be like* is rated as less feminine when followed by Pidgin as compared with English, although this effect does not reach significance in the statistical model (Table 3).

Figure 5: plot comparing femininity ratings in Pidgin and English across quotative and control conditions

Evident in the model of femininity (Table 3) is a significant interaction between the Pidgin condition and quotative *tell* (p <0.01), an effect which appears to mainly be carried by the difference in ‘4’ and ‘5’ ratings (as shown in Figure 5). Participants assign higher ratings when Pidgin is paired with quotative *tell* than when it is paired
with quotative *be like*. We also see a significant effect of Pidgin speaker identity; participants who self-identified as speakers of Pidgin were more likely to assign higher ratings for femininity. No main effects of language condition or quotative verb reached significance in the model.

**Table 3: model fit to femininity ratings**

```r
c Feminine.clmm <- clmm(femininity.response ~ lang.cond * quot.type + identity + gender +
(1 + quot.type / lang.cond | participant) + (1 | list/item), Hess = T, threshold = "flexible", link="logit",
data = data)
```

|                              | Estimate | Std. error | Z value | Pr (>|z|) |
|------------------------------|----------|------------|---------|----------|
| **Pidgin condition**         | -0.01987 | 0.26641    | -0.075  | 0.94055  |
| *like*                       | -0.37496 | 0.30818    | -1.217  | 0.22373  |
| *tell*                       | -0.16229 | 0.30426    | -0.533  | 0.59376  |
| **Identity - speaker**       | 1.12908  | 0.64663    | 1.746   | 0.08079  |
| **Gender - women**           | -0.57355 | 0.55231    | -1.038  | 0.29906  |
| **Pidgin : like**            | -0.03436 | 0.3277     | -0.105  | 0.91649  |
| **Pidgin : tell**            | 0.84625  | 0.31291    | 2.704   | p<0.01   |
3.5. Interim Discussion
Regarding quotatives, the results show that young people rate quotative *be like* and quotative *tell* as equally local when paired with Pidgin. This result is somewhat surprising, given the fact that quotative *be like* is a recent addition to the system of Pidgin, having arrived a full generation after being adopted into the system of Hawai‘i English (present only in the speech of those born between 1983-1988 (Drager et. al., in prep). The fact that it took a full generation for quotative *be like* to show up in the system of Pidgin could suggest some resistance to adopting its use on the part of speakers, or perhaps this delay resulted from lack of exposure. Since *be like* is recent in Pidgin, we might not expect it to be perceived as local, particularly since quotative *tell* is less rare in Hawai‘i English and Pidgin than in mainland varieties. In this study however, participants rate these two quotatives as equally local when paired with Pidgin.

The results of perceived localness, naturalness, and authenticity ratings show that participants do not rate *be like* differently from *tell* in terms of these factors. This result suggests that, at least for the generation of speakers who are using it in Pidgin, it appears to have been fully accepted as a part of the system. It is clear from the results that young people view Pidgin as more local than English, so perhaps what is driving this result is the strong salience of Pidgin as being local (Marlow & Giles, 2008), so much so that it overrides any non-localness that *like* may have.

In the model of femininity ratings (Table 3), there is a significant interaction effect between quotative *tell* and the Pidgin condition; participants assign higher ratings when Pidgin is paired with *tell* than when it is paired with *like*. This attention to the form of the quotative indicates that listeners may be sensitive to the role shift signaled by reported speech and suggests social salience around the role of the talker versus the speaker being quoted. Since *like* is linked with femininity in popular ideologies (Daily-O’Cain, 2000; D’Arcy, 2007; Buchstaller, 2013; D’Arcy, 2017), higher femininity ratings are expected with quotative *like* than with quotative *tell*. However, quotative *like* is perceived as less feminine than quotative *tell* when paired with Pidgin. One possibility is that the perceptual salience (Levon, 2014; Barnes, 2015) of Pidgin is playing a mediating role (see chapter 5 for more detailed discussion).

Additionally, participants who self-identified as speakers of Pidgin were more likely to rate the speaker highly in terms of localness, potentially due to intergroup solidarity, similar to the results reported by Ohama et. al. (2000).

3.6. Conclusion of perception study
Although the current study is constrained by several limitations such as small sample size, imperfect balance across conditions, and co-linearity between high school type and Pidgin speaker identity, these results still manage to highlight some surprising similarities with respect to perceptions of the quotative system. Listeners rate *be like*
and *tell* as equally local, suggesting that *be like* has been fully accepted by younger speakers as a part of the system of Pidgin. The results also suggest that while listeners are sensitive to the role shift indicated by reported speech, the perceptual dominance (Levon, 2014) of Pidgin may play a mediating role in how listeners attend to the speech signal.

Taken together, these results indicate a need for further research into language use and ideologies in a local context and raise several questions for future study. Perhaps with more balanced participant demographics and a larger sample size, it will be possible to further tease apart some of the ideologies surrounding quotatives and language varieties in Hawai‘i. While these questions cannot be answered at present, the work presented herein suggests many avenues of research worth pursuing.
4. Corpus Analysis

4.1. Corpus data analysis
While the matched-guise study presented in Chapter 3 gives us a picture of how young Hawai’i Locals perceive quotative *like*, we do not yet have a full picture of how the discourse-pragmatic *like* patterns in Hawai’i. Much work on discourse-pragmatic *like* has examined it within the framework of apparent time (see chapter 2 for an overview), looking at variation across age and gender⁴. Except for the quotative, none of this work has yet been conducted in the context of Hawai’i, where discourse-pragmatic *like* can be observed in both Pidgin and Hawai’i English. In order to examine the distribution of discourse-pragmatic *like* within the context of apparent time, this dissertation uses corpus data from both Pidgin and Hawai’i English, and compares this empirical evidence to the social perception of *like* examined in the matched-guise study. The results from the matched-guise experiment presented in chapter 3 shows that quotative *like* and *tell* are not perceived differently with regards to localness, potentially due to either the overwhelming perceptual salience of Pidgin, or due to the fact that discourse-pragmatic *like* has become a regular part of the grammar of Pidgin and Hawai’i English. Since these features are a regular part of the grammar, Local speakers do not perceive them as being non-local. The corpus-based analyses presented in this chapter were conducted to examine the distribution of discourse-pragmatic *like* in both languages.

4.2. Corpus analysis methods
As D’Arcy (2017, p. 163) explains, there are two methodological approaches that could be taken with this kind of data: the frequentist method (also known as the corpus linguistics method) and the variationist method. The frequentist method uses absolute frequency via normalization to quantify the data (e.g. tokens of interest per 100 words) whereas the variationist method is a proportional analysis, assessing the number of attested tokens out of all potential uses of the variable which results in proportional frequency (e.g. tokens of *like* out of all discourse markers produced). Both methods account for social context and are informative with regard to the behavior of a particular form within the larger system. While D’Arcy opts for the variationist approach in her 2017 book, she notes that *like* is “ideally suited to corpus linguistics methods, where changes in its normalized text frequency can be tracked relatively straightforwardly.” (2017, p. 165). This dissertation will employ both frequentist methods (section 4.4) and variationist methods (section 4.5) in order to present a broad yet nuanced picture of the ways discourse-pragmatic *like* has been used in Hawai’i, and what changes are taking place in both Pidgin and Hawai’i English.

⁴ While I recognize multiple genders, none of the participants in this study reported identifying as trans, non-binary, or any other gender. Thus, this dissertation will use the terms “women” and “men” when discussing participant gender, as that is how participants chose to self-identify.
4.3. Speakers selected for analysis

The data used for the analysis of discourse-pragmatic functions in Pidgin and Hawai‘i English comes from the same two corpora analyzed by Drager et. al. (in prep). The data for Pidgin comes from the Influences and Variation in Hawai‘i Creole English project spearheaded by Jeff Siegel and Kent Sakoda. The analysis is focused on data from 16 sociolinguistic interviews conducted in Pidgin, by interviewers who were familiar with the interviewees. For the purpose of this dissertation, names have been changed to pseudonyms to protect the identity and privacy of the speakers. The speakers selected for analysis are balanced for age cohort (young and middle) as well as gender. A total of 54802 words from this corpus were transcribed for analysis in this dissertation. Speaker gender and birth year are shown in Table 4 below.

<table>
<thead>
<tr>
<th></th>
<th>women</th>
<th>men</th>
</tr>
</thead>
<tbody>
<tr>
<td>younger (1983-1988)</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>middle (1949-1967)</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

The data for Hawai‘i English comes from sociolinguistic interviews conducted for the Language in Hawai‘i Project, headed by Katie Drager. The analysis will focus on data from 42 selected participants, balanced for age cohort (young, middle, and older) as well as gender. Interviews were conducted in locations familiar to the interviewee, such as their home or the home of a friend. Participants were recruited through word of mouth, and in most cases, participants were not acquainted with the researcher prior to the interview. All participants in the data set are from one of three areas of O‘ahu: Kalihi, Kaimuki, or Kāne‘ohe. Some of the speakers in this corpus chose to pick out pseudonyms at the time of recording, while others preferred that their real names be used in publication. A total of 1602569 words from this corpus were used for analysis in this dissertation. Speaker demographics are shown in Table 5.

<table>
<thead>
<tr>
<th></th>
<th>women</th>
<th>men</th>
</tr>
</thead>
<tbody>
<tr>
<td>younger (1982-1993)</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>middle (1950-1970)</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>older (1921-1947)</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

Some differences between the two corpuses should be noted. While the Hawai‘i English corpus contains data from speakers born prior to 1949, and is constrained by neighborhood, the Pidgin corpus does not contain data from an older cohort of speakers, and is not balanced by speaker neighborhood. Notably, the Pidgin

---

5 Interviews were conducted by several researchers from the University of Hawai‘i at Mānoa, including the current author.
interviews were conducted by someone with whom the interviewee was familiar; this is not the case with the Hawai’i English data set. Discourse markers are associated with informal speech, and therefore we might expect to see lower rates of discourse markers with unfamiliar interlocutors. However, to foreshadow the results, that is not what we see with the Hawai’i English data set (section 4.4.2.2), which may provide some evidence that the speakers were using a relatively informal style during the interviews.

Despite these differences, the two corpuses are the only available and relatively comparable corpuses of Pidgin and Hawai’i English currently available. While having perfectly parallel corpores would be ideal, this high standard is often unattainable for marginalized and minority languages such as Pidgin. In the case of understudied languages, we mustn’t let the perfect be the enemy of the good.

4.4. Coding of the corpus
For the purposes of coding, an utterance was defined as containing exactly one matrix clause (plus any subordinate clauses). This is meant as an operational definition for the purposes of this data set and research question, and is not intended to take any theoretical stance.

4.4.1. Frequentist analysis of like
4.4.1.1. Frequentist analysis methods
For the frequentist analysis all tokens of like were tagged using the 12 types described in D’Arcy (2017) and presented in chapter 2. Following D’Arcy (2017), utterance position was used to distinguish which type of discourse-pragmatic token was present. Tokens in initial position were coded as discourse markers. Tokens in final position were coded as sentence adverbs. Tokens in medial position were further disambiguated based on what type of information they were immediately followed by; tokens immediately followed by reported speech or thought were coded as quotatives, tokens immediately followed by a numeral quantifier were coded as approximate adverbs, and all other tokens were coded as discourse particles (which can occur in all other utterance medial positions (see D’Arcy, 2017, p. 15 for an overview). An example of the coding scheme used is provided in Example 11 below, and example sentences from the English data set are provided in Example 12.

<table>
<thead>
<tr>
<th>discourse-pragmatic token</th>
<th>criteria</th>
<th>code</th>
</tr>
</thead>
<tbody>
<tr>
<td>discourse marker</td>
<td>utterance initial</td>
<td>dm</td>
</tr>
<tr>
<td>quotative</td>
<td>utterance medial preceding quoted material</td>
<td>q</td>
</tr>
<tr>
<td>approximate adverb</td>
<td>utterance medial preceding a numeral quantifier</td>
<td>aadv</td>
</tr>
<tr>
<td>discourse particle</td>
<td>utterance medial</td>
<td>dp</td>
</tr>
<tr>
<td>sentence adverb</td>
<td>utterance final</td>
<td>sadv</td>
</tr>
</tbody>
</table>
**Example 12:** Sample sentences showing discourse-pragmatic *like* in the Hawai’i English data set

- **dm** “We have taken a little bit from everywhere and just made our own language. *Like* there’s Hawaiian right, there’s the actual language...this our own variation, Pidgin.”
- **q** “They said something I’m *like*. ‘Oh god you’re. from Maui aren’t you.’”
- **aadv** “That’s why we went through *like* three weed whackers already.”
- **dp** “Had one pool hall right around the corner it had all these *like* killer looking Filipinos used to drink beer in and play pool there.”
- **sadv** “You know by the headrest *like*.”

Frequency of types was calculated via the number of tokens of a particular type of *like* out of every 100 words in the corpus or relevant subset (i.e. young women) (Biber, Conrad, & Reppen, 1998; Drager, 2011; Drager, 2016).

### 4.4.1.2. Pidgin

This section presents results for the frequentist analysis of *like* in both Pidgin and Hawai’i English. Figure 6 shows frequency by age group for all types of *like* analyzed. In keeping with variationist work, use of *like* is plotted as a line graph when represented visually, in order to show change over time. This is the standard way of visually representing apparent time data in variationist work, making the data presented here interpretable alongside the variationist analysis presented in section 4.4.2. Plots shown in this section are calculated using the frequentist approach, via normalization per 100 words.

In addition to the twelve types of *like* analyzed by D’Arcy (2017), Pidgin also has a second verbal form of *like*. There is the form familiar to English speakers, as in “I *like* to go to Vegas” with a meaning similar to “enjoy”, and a second form which will be familiar to Pidgin speaking readers; “I *like* go Vegas” where the meaning of *like* “indicates volition or wanting to do something” (Sakoda & Siegel, 2003, p. 41). These are coded in the data as v and v1 respectively, and not analyzed further in this dissertation.

Figure 6 indicates that use of the discourse marker (dark green) and discourse particle (pink) have experienced a marked increase in apparent time. The quotative (light orange) and approximate adverb (light blue) also show increases in apparent time. These findings are in line with D’Arcy’s findings for global Englishes (2017). Also evident in the graph is the fact that the preposition (red), conjunction (light green), comparative complementizer (dark blue), and verb (light purple) have experienced a slight decline in use.
Taking a closer look at the data, Figure 7 examines distribution of the various tokens of *like* for both women and men. While both women and men have increased their use of the discourse particle (pink) and approximate adverb (light blue), use of the other discourse-pragmatic functions are markedly different. Women show more of an increase in use of the quotative (light orange) than men. Also notable is the fact that men have strongly increased their use of the discourse marker (dark green), while women’s use of this function increases only slightly across apparent time.
4.4.1.3. Hawai‘i English

Figure 7: Proportion of *like* in Pidgin shown for women (left panel) and men (right panel)

![Graph showing proportion of like in Pidgin by gender](image)

Figure 8 displays change over time in the system of Hawai‘i English. As evident in the graph, speakers of Hawai‘i English increased their use of the discourse particle across apparent time. This increase appears to be peaking in use among the youngest cohort of speakers recorded. Also evident in the graph is the fact use of the discourse marker and quotative have increased over time. These finding are consistent with the
Looking closer at the data reveals that trends are similar for women and men speaking Hawai‘i English. While D’Arcy (2007) finds that women use the quotative and the marker more frequently than men, this appears not to be the case with Hawai‘i English. Here we see that rates of use are quite similar. This non-effect of gender is consistent with our variationist analysis of the quotative (Drager et. al., in prep), which potentially points to a larger trend in Hawai‘i.
Comparing Pidgin and Hawai‘i English

Comparing the results for Pidgin and Hawai‘i English reveals both similarities as well as differences in use of *like* over time. In both Pidgin and English, use of the discourse particle is increasing in apparent time for both women and men. Men use the discourse particle at higher rates than women in both languages, a finding which follows D’Arcy’s finding for global Englishes. Also in line with other studied varieties is the fact that use of the quotative and discourse marker have increased in both languages.
Trends diverge with respect to the approximate adverb. In Hawai‘i English, use of the approximate adverb appears relatively stable across apparent time, while in Pidgin use has increased only slightly. These findings contrast with D’Arcy’s (2006) findings on the approximate adverb, which has undergone a recent rapid expansion across global Englishes. At this point, it is unclear why Pidgin and English diverge from global trends regarding the use of the approximate adverb respectively.

Interestingly, Pidgin and English diverge with respect to usage of the discourse marker. In Hawai‘i English, women and men produce discourse marker *like* at similar rates to one another. In Pidgin however, young men use discourse marker *like* at higher rates than women (see Figure 7). This is in contrast with D’Arcy’s findings for global Englishes, where women use the discourse marker at higher rates than men (D’Arcy, 2007, p. 396). This surprising find merits further analysis.

The following section seeks to answer the question: how does *like* operate within the ecosystem of discourse markers in each language? This question is explored using variationist methodology, in section 4.4.2.

### 4.4.2. Variationist analysis of discourse markers

The variationist analysis presented in this section plots each discourse marker as a percentage of all discourse markers analyzed, as discussed by D’Arcy (2017). Using this method enables examination of change in apparent time within the system of discourse markers.

In addition to investigating the five discourse-pragmatic functions of *like* exemplified in chapter 2, this dissertation also seeks to examine discourse marker variants which alternate with *like*. A variationist analysis was conducted to explore the alternation between discourse markers in both language varieties. This analysis employed the proportional method discussed by D’Arcy (2017, p. 163), by analyzing each token out of all discourse marker tokens in the respective corpus in order to investigate the distribution and variation of discourse markers alternating with *like* in the systems of Pidgin and Hawai‘i English. While previous work has used the variationist method to analyze discourse particle *like* (in addition to other functions) (D’Arcy, 2017), this type of analysis is outside the scope of this dissertation, as it requires analyzing all utterance-medial positions where a discourse particle could potentially occur. Thus, this dissertation is focused on a variationist analysis of discourse markers, leaving the particle for future work.

In addition to *like*, 12 other discourse markers were selected for analysis. The majority of the discourse markers chosen for analysis were selected based on previous work in other studied varieties (*but, just, so, oh, well, you know, I mean, like, yeah*) (Jucker & Ziv, 1998; Tagliamonte S. A., 2005; Kastronic, 2011; Fraser, 1988; Koops & Lohmann, 2015). In addition to these previously studied discourse markers, this dissertation also
analyzes discourse markers used in a Local context (brah, eh, and ho), in order to
examine variation in the systems of both Pidgin and Hawai‘i English (Sakoda & Siegel, 
2003). A by-speaker analysis was also conducted in order to investigate individual 
variation. Discourse markers within reported speech were not analyzed. There are no 
meaningful changes to the results when brah is not included in the analysis.

4.4.2.1. Pidgin

Figure 10 shows change over time in the discourse marker system of Pidgin. Compare 
this with Figure 6, which shows a frequentist analysis of all types of like in Pidgin, 
showing an increase in overall use of discourse marker like from middle-aged to 
younger speakers of Pidgin. Comparatively, Figure 10 provides a variationist analysis 
of these discourse markers, examining the proportional frequency of each token.

Evident in the graph is the fact that discourse marker you know is experiencing a sharp 
decline in apparent time, dropping from over 25% use by middle-aged speakers to 5% 
among younger speakers. Discourse markers well and but have also undergone a 
decline in apparent time, though not nearly as steep as you know. The variants oh, 
like, yeah, and ho appear to be increasing in use, while just, eh, and brah show no 
noticeable change in use.

Figure 10: Pidgin discourse markers; change in apparent time (variationist analysis)

![Figure 10: Pidgin discourse markers; change in apparent time (variationist analysis)](image)

Figure 11 plots change over time in the system of discourse markers between women 
and men speaking Pidgin. Evident in the graph is the fact that both women and men 
have drastically reduced their use of discourse marker you know. Also evident in the 
graph is the fact that women have increased their use of discourse marker oh much 
more than men. Also notable is the fact that middle-aged men use discourse marker 
well at much higher rates than middle-aged women (4.497% and 3.035%, respectively). 
Men use discourse marker like at much higher rates than women speaking Pidgin, a
result that contrasts with D'Arcy’s findings (2017). The difference between women’s and men’s production rates is especially noticeable for the younger cohort of speakers, with *like* making up less than five percent (2.88%) of discourse markers produced by young women, and comprising 8.87% of the discourse markers produced by young men, respectively.

![Figure 11: Discourse markers in Pidgin](image)

For statistical analysis, mixed effects logistic regression models were fit to the binary dependent variable of whether or not the produced variant was *like*. Included in the model as fixed effects were speaker age and gender, interaction of speaker age and gender, as well as random intercepts of speaker. The reference level for speaker gender was set to men, and the reference level for speaker age was set to younger speakers, as these are the groups predicted to have higher levels of discourse marker...
like.

The statistical analysis of discourse marker *like* production in Table 6 shows a significant effect of gender (p<0.05), with men being significantly more likely to produce discourse marker *like* than other discourse markers, compared to women. Quite surprisingly, age is not a predictive factor in the production of discourse marker *like*, and the interaction between age and gender failed to reach significance.

**Table 6: Model fit to discourse marker *like* production in Pidgin**

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-2.599</td>
<td>0.23</td>
<td>-11.299</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>age group = y</td>
<td>0.3532</td>
<td>0.2269</td>
<td>1.557</td>
<td>0.1195</td>
</tr>
<tr>
<td>gender = m</td>
<td>0.5064</td>
<td>0.2279</td>
<td>2.222</td>
<td>p&lt;0.05</td>
</tr>
<tr>
<td>age group = y : gender = m</td>
<td>0.1118</td>
<td>0.2269</td>
<td>0.493</td>
<td>0.6223</td>
</tr>
</tbody>
</table>

4.4.2.1.1. **Individual variation in Pidgin**

While the information presented thus far in this section has given an overview of how speakers of Pidgin and Hawai’i English are using discourse-pragmatic *like*, this information has been presented in terms of broad demographic categories. In order to gain a better understanding of the variation described in the use of discourse markers, an analysis of individual speaker variation has been conducted to determine if the trends discussed are consistent across individuals, or if there is a high degree of variation within the studied population. Figure 12 shows individual variation in the proportional use of discourse marker *like* by Pidgin speaking women. Evident in the graph is a great degree of individual variation. Carla⁶ could be considered a leader of change, having a production rate of discourse marker *like* three times that of the other middle-aged women. Notable as well are the differences between the younger women, whose production rates range from 12% (Lena) to 0% (Starla). The high degree of individual variation shown here helps to explain why age group does not reach significance in the statistical model.

**Figure 12: Individual variation in proportional use of discourse marker *like* by women**

---

⁶ All speaker names used in the Pidgin corpus are pseudonyms, in contrast with the Hawai’i English corpus where some speakers elected to use their full names.
Figure 13 shows individual variation in the proportional use of discourse marker *like* by Pidgin speaking men. The amount of individual variation displayed in the plot explains the discrepancy between the trend seen in Figure 11 and the fact that age is not a predictive in the use of discourse marker *like*, in the model shown in Table 6. The young speakers with higher rates of discourse marker *like*, Lena and Mina, are driving the trend seen in Figure 11, but this trend fails to reach significance due to the high degree of variation. Starla does not use discourse marker *like* at all, and Carla’s rates are much more similar to the young women than the other middle-aged women. Carla’s high rates are potentially explained by her profession as a sex worker, where she frequently interacts with diverse groups of speakers, potentially accommodating to aspects of their speech. Future work involving a larger sample size which takes these types of social factors into account would help to determine if the trend seen in Figure 11 reaches significance in a statistical model.

*Figure 13:* Individual variation in proportional use of discourse marker *like* among Pidgin speaking men
4.4.2.2. Hawai’i English

Figure 14 shows change over time within the system of Hawai’i English discourse markers. Evident in the graph is a steep decline in the use of discourse markers so (10%) and you know (8%) in apparent time. Discourse marker but (5.5%) also appears to be decreasing in apparent time. However, when looking at raw numbers, it is not clear that but is decreasing, but proportionally it appears that way because several new discourse markers have entered the system. Discourse markers yeah, like, and just appear to be increasing in apparent time. The Pidgin discourse markers brah, ho, and eh appear relatively stable across apparent time.

Figure 14: Hawai’i English discourse markers; change in apparent time (proportional analysis)
Figure 15 shows change over time in the system of discourse markers between women and men in Hawai‘i English. While trends for some discourse markers appear similar between women and men, women appear to be abandoning discourse marker *you know* at a much faster rate than men. This is similar to what is observed in Pidgin, where women are also abandoning use of discourse marker *you know*. Women also appear to have increased their use of discourse marker *so*, while the opposite is true of Hawai‘i English speaking men. Interestingly, while Pidgin speaking men are leading the use of discourse marker *like*, this is not the case in Hawai‘i English, where women and men’s rates are 3.95% and 4.66% for younger speaking women and men respectively.
Multilevel logistic regression models were fit to the subset of the data containing discourse markers. The dependent measure was a binary factor of whether or not the produced variant was *like*. Included in the model as fixed effects were speaker age, speaker gender, as well as interaction effects of age and gender, as well as a random intercepts for each speaker. The reference level for speaker gender was set to men, and the reference level for speaker age was set to younger speakers, as these are the groups predicted to have higher levels of discourse marker *like*.

Results of the statistical analysis of discourse marker *like* production in Table 7 shows a significant effect of age (p<0.05), with younger speakers being significantly more likely to produce discourse marker *like* than other discourse markers. Contrary to the popular ideologies discussed in chapter 2, gender does not significantly predict a difference in the frequency of production of discourse marker *like* in Hawai‘i English. This is particularly interesting, given that gender reached significance as a fixed effect for the Pidgin data set (see Table 6). No interactions reached significance for the Hawai‘i English model (Table 7). These results are consistent with the findings of the frequentist analysis (section 1.4.3) which show that younger speakers have the highest rates of discourse marker *like*, and trends do not appear different for women and men.
Table 7: Model fit to discourse marker *like* production in Hawai’i English

dm.model<- glmer(dms ~ age.group * subj.gender + (1|speaker), data = dms, family = binomial)

|                     | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------------|----------|------------|---------|---------|
| (Intercept)         | -3.914   | 0.2932     | -13.35  | p < 0.001 |
| age group = y       | 1.665    | 0.378      | 4.405   | p < 0.001 |
| age group = m       | -0.2008  | 0.3978     | -0.505  | 0.614   |
| gender = m          | -0.1178  | 0.2824     | -0.417  | 0.677   |
| age group = y : gender = m | 0.3731 | 0.372      | 1.003   | 0.316   |
| age group = m : gender = m | -0.3371 | 0.3989     | -0.845  | 0.398   |

4.4.2.2.1. Individual variation in Hawai’i English

Figure 16 shows individual variation in the proportional use of discourse marker *like* by Hawai’i English speaking women. Immediately evident in the graph is the high degree of individual variation among the younger cohort of women; ranging from the lowest use by Lei (2.198%) to 17.582% (Keiko). This high degree of individual variation is possibly indicative of a change still in progress, or perhaps this is due to stylistic choice on the part of the individual speakers.

Figure 16: Individual variation in proportional use of discourse marker *like* by women

Figure 17 shows individual variation in the proportional use of discourse marker *like* by Hawai’i English speaking men. Immediately evident in the graph is the fact that the younger speaker with the lowest production rate of *like* (Justin Taomae) is still producing discourse marker *like* at a higher rate than the highest producers from the older and middle-aged groups of speakers. Also evident is a high degree of individual variation within the younger speakers, ranging from Jeremy’s 27.272% to Justin Taomae’s 4.839% production rate. There is less individual variation in the middle-aged and older groups of men; none of the middle-aged men produce discourse marker *like*
at rates higher than 4%, and none of the older men produce discourse marker *like* more than 2% of the time. Three of the older men and two of the middle-aged men produce no tokens of discourse marker *like* at all.

**Figure 17:** Individual variation in proportional use of discourse marker *like* by men

Discourse pragmatic *like* is on the rise in Hawai‘i; in the frequentist analysis we see that four types of discourse pragmatic *like* (discourse marker, discourse particle, quotative, and approximate adverb) are increasingly used by younger speakers of both Pidgin and Hawai‘i English. This finding fits in with trends seen in global Englishes (see both Tagliamonte et. al. 2016 as well as D’Arcy 2017 for an overview) as well as with analysis of the quotative in Pidgin and Hawai‘i English conducted by Drager and colleagues (in prep a, b). When we examine the discourse marker in particular, we encounter some interesting deviations from global trends. Globally, D’Arcy finds that discourse marker *like* is used more by women than men (who she finds use the particle more (2007; 2017)). However, in Hawai‘i English, we see that rates are similar between women and men. An unexpected finding is the fact that in Pidgin, men use the discourse marker *like* at higher rates than women, and this is supported by the statistical analysis presented in section 1.5.1. When we examine individual speakers’ production of discourse marker *like*, we find a high degree of individual variation, particularly in the smaller Pidgin data set. In Section 1.6, we turn our attention to how speakers are using discourse marker *like* in Pidgin and Hawai‘i English.
4.4.3. Pragmatic usage of discourse marker *like*

4.4.3.1. Methods

To further examine the use of discourse marker *like*, this dissertation analyzes the functions that discourse marker *like* serves within the discourse context. While previous work has looked at the pragmatic functions that discourse markers serve within the discourse context, as of yet, no research has been conducted to examine the distribution of pragmatic uses of discourse marker *like* across age and gender. This analysis focuses on how discourse marker *like* is used by speakers of each age and gender group in both Pidgin and Hawai‘i English. These results can then be compared within the context of Hawai‘i, as well as the context of language change taking place in global Englishes. While much of the work focused on *like* has been quantitative, very little of this work has focused on which groups are using *like* in what ways. This analysis examines how this incoming form is being used, and tracks changes in how it is used by different demographic groups. This largely unexplored area has the potential to yield interesting results related to variation and change with the use of discourse-pragmatic *like*.

Each token of discourse marker *like* was examined within the context of the surrounding discourse, and categorized according to the criteria in Table 8 below. The tokens were analyzed using the variationist method, plotting each pragmatic type of discourse marker *like* out of all tokens of discourse marker *like* in the corpus or relevant subset.

<table>
<thead>
<tr>
<th>Pragmatic Use</th>
<th>Coding Criteria</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>clarification</td>
<td>Clearing up existing or potential confusion</td>
<td>Eric younger man</td>
</tr>
<tr>
<td></td>
<td>“Like my teacher always used to make fun of him aeswai” explaining how he knows of a person he’s never met</td>
<td></td>
</tr>
<tr>
<td>elaboration</td>
<td>Holding the floor</td>
<td>Sarah younger woman</td>
</tr>
<tr>
<td></td>
<td>“Like living with one married lady oh boy” holding the floor when the interviewer begins to speak</td>
<td></td>
</tr>
<tr>
<td>focus</td>
<td>Highlighting new information</td>
<td>Kaleo younger man</td>
</tr>
<tr>
<td></td>
<td>“Like me I’m from up country so going to Hana is …” providing interlocutor with new information about where he is from</td>
<td></td>
</tr>
<tr>
<td>illustration</td>
<td>Providing an illustrative example</td>
<td>Kaleo younger man</td>
</tr>
<tr>
<td></td>
<td>“Like they always say ‘oh cab’” – Kaleo, younger man providing an example of regional variation in Kaua‘i Pidgin</td>
<td></td>
</tr>
<tr>
<td>stance</td>
<td>The speaker’s epistemic stance</td>
<td>Alika younger man</td>
</tr>
<tr>
<td></td>
<td>“Like to me the kine halau that come down from Japan (…)” discussing which type of visitors to Hawai‘i he believes display pono behavior</td>
<td></td>
</tr>
<tr>
<td>topic</td>
<td>Relating the following utterance to prior discourse</td>
<td></td>
</tr>
</tbody>
</table>
From the Pidgin data set, two tokens were coded as ambiguous and removed from analysis, constituting 1.4% of the data. From the English data set, a total of six tokens were coded as ambiguous and removed from analysis, constituting 0.2% of the data.

4.4.3.2. Pragmatic usage of discourse marker *like* in Pidgin

Figure 18 displays the distribution of the different pragmatic uses of discourse marker *like* in Pidgin across middle-aged and younger speakers. Evident in the graph is a sharp decline in the use of the illustrative function, coupled with a rise in the clarification use, and a slight (5%) rise of the elaborative use. Other variants (focus, stance, and topic) remain relatively stable across apparent time.

![Figure 18: Use of discourse marker *like* in Pidgin, shown over time](image)

Figure 19 shows variation in the pragmatic use of discourse marker *like* by gender. For both women and men, the illustrative use of discourse marker *like* has decreased with apparent time, while use of discourse marker *like* to provide clarification has increased. Women have increased their use of discourse marker *like* to signal elaboration, while men have not increased how often they use it for this purpose. With only two examples in the corpus of discourse marker *like* being used to signal stance, it is impossible to determine any sort of trend.
Figure 19: Pragmatic use of discourse marker *like* in Pidgin by women and men (percent of all dm *like*)

**Pragmatic use of discourse marker *like* in Pidgin by gender**

Pragmatic usage of discourse marker *like* in Hawai‘i English

Figure 20 displays pragmatic use of discourse marker *like* in Hawai‘i English across apparent time. Evident in the graph is a decline in the use of the illustrative and focus usages. Hawai‘i English speakers appear to have increased their use of discourse marker *like* to signal elaboration, clarification, and stance work. Textual use of discourse marker *like* is used by the middle-aged cohort of speakers, but with only two tokens in the corpus, it is difficult to say if this is an outlier or indicative of a trend.

Figure 20: Pragmatic use of discourse marker *like* in Hawai‘i English
Figure 21 displays proportional frequency of the pragmatic uses of discourse marker *like* in Hawai‘i English by gender. While women initially used discourse marker *like* for illustrative purposes, this use has decreased with middle-aged and younger women. Among middle-aged and younger women, the uses of discourse marker *like* appear to be in a state of flux; clarification and elaboration may have plateaued slightly, focus and topic marking are trending downwards, and using discourse marker *like* to mark stance appears to be increasing slightly in apparent time, though no one use of discourse marker *like* is a clear favorite at this point in apparent time.

When looking at Hawai‘i English speaking men, we see a few patterns emerge. Middle aged men appear to have abandoned the use of discourse marker *like* to mark focus, though this does increase slightly with younger men. Younger men most often use discourse marker *like* to elaborate, or “hold the floor”, and have also slightly increased using discourse marker *like* for clarification purposes.

From this analysis, we can see that while both women and men are increasing their use of discourse marker *like*, they are using it to do different things within the context of discourse. This is especially clear with the older and middle-aged speakers in this data set. The proportion of pragmatic uses appears to be stabilizing for the younger group of speakers.

**Figure 21:** Proportional frequency of pragmatic uses of discourse marker *like* in Hawai‘i English by women and men

<table>
<thead>
<tr>
<th>Pragmatic use of discourse marker like in Hawai‘i English by gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>women</td>
</tr>
<tr>
<td>men</td>
</tr>
</tbody>
</table>

**4.4.4. Comparing Pidgin and English**

The frequentist analysis suggests that the frequency of discourse pragmatic functions of *like* (excluding the sentence adverb) is increasing across age cohorts in both Pidgin and English. This finding is verified for the discourse marker through using the
variationist analysis, which also provides evidence that discourse marker *like* is increasing in apparent time. This effect appears to be driven by men in the Pidgin corpus, whereas both young women and men in the Hawai`i English corpus have the highest rates of discourse marker *like* use. However, age fails to reach significance in the statistical models. The high degree of individual variation may help to explain why age does not reach significance, and helps to highlight the need for more robust corpora of minority and marginalized language varieties.

Examining the pragmatic usage of discourse marker *like* within the discourse context reveals that the elaborative function, which seems to “hold the floor”, is increasing in use across apparent time in both Pidgin and English. Additionally, Pidgin speakers appear to be employing discourse marker *like* for the purpose of clarification, but in Hawai`i English younger speakers are not using this function more than their middle aged counterparts.

These results provide a detailed look how discourse pragmatic *like* is used in the multilingual context of Hawai`i, and highlights the strength of global trends which persist in this situation of close language contact.
5. Discussion

5.1. Research questions, revisited
This dissertation sets out to examine a global phenomenon within a local context. Since Butters’ seminal study in 1982, sociolinguists have been captivated by the rise of discourse-pragmatic *like* in English varieties worldwide, as well as their neighboring contact languages. This body of research has carefully documented the worldwide increase in the discourse-pragmatic functions of *like*; with particular attention to the quotative, which entered English varieties rapidly and uniformly, patterning similarly in studied varieties across the globe. In the context of this rapid worldwide language change, this dissertation examines discourse-pragmatic *like* in detail in both Pidgin and Hawai’i English in order to determine if either or both language varieties pattern similarly to these worldwide trends, and to examine how discourse-pragmatic *like* is being used and socially interpreted within the context of Hawai’i.

5.2. Brief overview of perception results
The results from the perception experiment demonstrate that the most perceptually attended to aspect of the stimuli was Pidgin. The analysis reveals a significant main effect of language condition, suggesting that the presence of Pidgin is highly perceptually noticeable to Hawai’i Locals. In fact, it is likely that the high noticeability of Pidgin overrode other aspects of the stimuli for participants; participants may have rated quotative *be like* and quotative *tell* as equally local when paired with Pidgin due to the presence of Pidgin in this experimental context. In other words, the use of Pidgin may identify someone as Local regardless of what quotative forms are present, as long as those forms are observed in spontaneously produced Pidgin. This is explored further in section four.

Participant gender failed to reach significance in any of the statistical models, suggesting that for young Hawai’i locals, women and men have similar perceptions of quotatives *be like* and *tell*. Future work could examine interaction between participant gender and talker gender.

In terms of perception of speaker gender, participants were more likely to assign a higher femininity rating when Pidgin was paired with quotative *tell*, than when it was paired with quotative *like*. This is potentially related to ideologies around Pidgin and gender (Meyerhoff, 2004), and is explored further in section 5. This was the only case where gender reached significance in the perception study.

5.3. Brief overview of corpus analysis results
The data from the corpus study shows that discourse-pragmatic *like* in Hawai’i patterns similarly with other studied varieties worldwide. The frequentist analysis provided in chapter 4 shows that in both Pidgin and Hawai’i English, the discourse-pragmatic functions of *like* have increased in use among the younger cohort of speakers. Results also show that among Pidgin speakers, young men are using
discourse marker *like* at higher rates than young women. In Hawai‘i English, patterns of use are more consistent between women and men.

The variationist analysis provided in chapter 4 provides a slightly different perspective on variation in the discourse marker systems of both Pidgin and Hawai‘i English. The results of the variationist analysis provide evidence that young Pidgin speaking men in Hawai‘i are using discourse marker *like* at higher rates than young women. In Hawai‘i English, however, women and men pattern more similarly. This finding deviates from patterns observed in other studied varieties, where young men are using discourse marker *like* at lower rates than young women (D'Arcy, 2007). This surprising finding naturally leads to the question of why young Hawai‘i Locals are behaving differently with respect to discourse marker *like*.

By expanding the variationist approach to observe how discourse marker *like* is being used in discourse by speakers of different ages, we can see that it is being used differently by its early adopters. This novel approach, developed and implemented in this dissertation, allows us to investigate not just the frequency and proportional production rates of discourse marker *like*, but gives us an idea of what may be driving speakers to use certain forms within the context of discourse. Understanding the pragmatic uses that speakers employ specific forms for may help us to understand the changes that occur within the system.

In Pidgin, discourse marker *like* was most frequently used to provide illustrative commentary among middle-aged speakers, particularly women, but this function has fallen out of favor among younger speakers, and has largely been eclipsed by its use for elaboration and clarification. This finding raises the question of whether this gap has been filled by some other discourse marker, or whether speakers no longer feel the need for the illustrative functions, perhaps due to more narrative strategies. The rise of quotative *like* has been linked to changing narrative structure (Tagliamonte & D'Arcy, 2007; D'Arcy, 2012). Future work investigating the connection between discourse marker *like*, quotative *like*, and changing discursive practice could provide much needed insight into the factors surrounding the rise of discourse-pragmatic *like*. While there are differences in gender among the older and middle-aged cohorts of speakers, discourse marker *like* appears to be stabilizing in its pragmatic use among younger speakers.

Previous work finds that while discourse-pragmatic *like* is socially evaluated as being more feminine, empirical work shows that this picture is far more complex (Daily-O'Cain, 2000; D'Arcy, 2007). D'Arcy (2017, p. 120) finds that young women use discourse marker *like* significantly more than young men; the effect peaks with speakers born between 1984 and 1988. Given this finding, we would expect to find young women using discourse marker *like* more frequently than young men in Hawai‘i English, and possibly in Pidgin as well. However, the results presented in this dissertation do not line up with this prediction. In Pidgin, young men have higher rates of discourse marker *like* than young women, and in Hawai‘i English, young women and men have
similar rates of production for discourse marker *like*.

In Hawai‘i English, both the focus marking and illustrative uses of discourse-marker *like* were used more by older and middle-aged speakers. Middle-aged women often use discourse marker *like* to signal focus, while younger speakers of Hawai‘i English are using discourse marker *like* to signal elaboration, or to continue “holding the floor”.

The results of the variationist analysis and frequentist analysis of discourse markers are consistent, and when compared together with the results of recent work on quotatives in Hawai‘i (Drager et. al. in prep; Schutz et. al. in prep), provide robust evidence that discourse-pragmatic *like* is on the rise in Hawai‘i. The frequentist analysis of *like* provided in chapter 4 demonstrates that quotative *be like* is more frequently used by the youngest cohorts of speakers in both languages, and this result concurs with the results of the variationist analyses provided by the Hawai‘i research team who find that quotative *be like* is on the rise in both Pidgin and Hawai‘i English. Furthermore, they find no gender-based differences in quotative use in both Pidgin and Hawai‘i English, consistent with the production results presented herein. Previous work demonstrates that English speakers conceptualize all tokens of *like* as being the same (Daily-O’Cain, 2000; D’Arcy, 2007), indicating that the studies of quotative *be like* and discourse marker *like* are indeed comparable. Taken together, these results show that younger listeners in Hawai‘i have prior experience with discourse-pragmatic *like* in both Pidgin and Hawai‘i English input.

Looking at the results for Pidgin and Hawaii English together, we are able to observe not only that discourse-pragmatic *like* is on the rise in Hawaii, but that its use among speakers appears to be stabilizing and converging with the youngest observed group of speakers.

5.4. **Localness and *like*: comparing production and perception**

Listeners have been shown to demonstrate perceptual sensitivity to associations between linguistic forms and social attributes across a wide variety of contexts. Native English speakers show sensitivity to the alternation between the -ing and -in variants of the English (ING) variable, attaching social meaning to the different variants in a way that depends on what other social information is attributed to the talker (Campbell-Kibler, 2011). Variants are ideologically linked to multiple social meanings, and which of those meanings becomes relevant depends on the context of the social interaction (Podesva, 2007). This social meaning is driven not just by stylistic choices of the speaker, but is also dependent on listener uptake (Eckert, 2011). Associations between linguistic features and social meanings have been found to be enabled by stereotypical attitudes and ideologies about speaker categories, demonstrating that a combination of cognitive and ideological factors guides perception of social meaning (Levon, 2014).

In addition to this perceptual sensitivity, discourse-pragmatic *like* is metalinguistically
commented upon (Daily-O'Cain, 2000; D'Arcy, 2007), so we would expect that the
listeners in the perception study would demonstrate sensitivity to its presence.
However, they do not.

Given this effect of prior experience, the surprising results of the perception study can
be explained by the results of the production study. The results presented in chapter
4 demonstrate that young listeners in Hawai'i have prior experience hearing discourse-
pragmatic like in Pidgin and Hawai'i English. It is important to note that all of the
participants in the perception experiment were young, in the age band of speakers
with the highest production rates of quotative be like (Drager et. al., in prep; Schutz et.
al., in prep), and the highest production rates of all forms of discourse-pragmatic like,
as presented in chapter 4. This high degree of contextual frequency indicates that
discourse-pragmatic like is not highly salient for the population of younger Locals
tested in the perception experiment, where salience is defined as “the increased
attention that results from a lack of experience with a variable” in line with Hay,
Drager, & Gibson, 2018 (p. 365). In other words, because discourse-pragmatic like is
common in the input, it is not very noticeable to this population. Since like is not
noticeable, listeners may be less likely to assign social meaning to it. These results
explain why this group of listeners does not demonstrate sensitivity to the presence of
discourse-pragmatic like in the perception experiment.

Another possibility is that the strength of association between Pidgin and Localness is
so strong that it overrides any “pop out effect” that like has. The perception study
clearly demonstrates that listeners are sensitive to who is an authentic native speaker
of Pidgin; they can tell that the speaker producing the stimuli is a native speaker
rather than a second language speaker. With the strong perceptual link between
Pidgin and Localness, listeners may hear most anything produced by an authentic
Pidgin speaker as being acceptable in Pidgin. In other words, listeners do not pay any
heed to which quotative the speaker is producing, because they can tell that she is an
authentic Pidgin speaker.

5.5. like and gender
Previous work finds strong language ideologies linked with discourse-pragmatic like. Much attention has been paid to the relationship between discourse-pragmatic like
and gender, both in popular ideologies and media, as well as in the scientific
literature. Discourse-pragmatic like is often popularly associated with the speech of
young women (Daily-O'Cain, 2000; Buchstaller, 2013; D'Arcy, 2007; D'Arcy, 2017).
Previous work has found that there is an association between Pidgin and masculinity
(Meyerhoff, 2004). Additionally, quotative tell is found at higher rates in Pidgin than
every English variety described to date (Drager et. al. in prep). Given this context, we
might expect that in Hawai'i, discourse-pragmatic like would be rated as more
feminine than masculine in the perception experiment (chapter 3) since tell is
frequent in Pidgin, and Pidgin is perceived as masculine. However, the results
indicate young Hawai'i locals, regardless of their gender, do not rate quotative be like
higher than quotative *tell* in terms of femininity.

Perhaps the ideological link between *like* and women is more contextually relevant in English, and the presence of Pidgin overrides those social associations that are linked with English (as suggested by Levon (2014)). To put this another way, it is possible that Pidgin is more perceptually prominent than English, and as a result, listeners attend selectively to attitudes and ideologies associated with Pidgin, while giving less attention to those associated with English. While this could be considered a task effect of the perception experiment (close proximity between Pidgin and Hawai‘i English stimuli) it is not uncommon to have both languages being used within the same discourse act in Hawai‘i. Many Hawai‘i Locals are fluent in both Pidgin and Hawai‘i English, so speaker-internal language contact is not uncommon either. Given this, it is possible that what is driving this result is the strong ideological link between Pidgin and Localness; when a Local speaker is producing forms found in both the local variety of English as well as the local creole, those forms may not be perceived as less Local.

Again, when we consider this result in a larger context, it becomes less surprising. In their variationist study of quotatives, Drager (et. al. in prep) found no gender difference in the production of quotative *be like* in either Pidgin or Hawai‘i English. The results of the corpus analyses presented in chapter four demonstrate that by and large, there are few gendered differences in the production rates of discourse-pragmatic *like* among younger Hawai‘i Locals. While we do see that young Pidgin speaking men are producing discourse marker *like* at higher rates than women, it is important to keep in mind that speakers perceive all types of *like* as monolithic (Daily-O’Cain, 2000; D’Arcy, 2007). It is clear then, that differences in women’s and men’s rates of discourse-pragmatic *like* production are not prominent in the linguistic landscape of Hawai‘i. Knowing this, it is not so unexpected that I find only minor differences in perception of gender. Given all of this information as a whole, it appears that younger speakers in Hawai‘i are converging with respect to their production of discourse-pragmatic *like*, and as a result they may not perceive use of discourse-pragmatic *like* to be strongly linked with gender.

Perhaps it is some combination of all of these aforementioned possibilities that is driving the results of the perception experiment. Future research focused on teasing apart these ideologies and associations would be a fruitful area to explore further.

5.6. The value of multiple approaches
This dissertation examines discourse-pragmatic *like* within the multilingual context of Hawai‘i using multiple methodological approaches. For production data, both frequentist and variationist approaches were utilized to examine the data set. A perception experiment examined Local listeners’ sensitivity to a particular type of discourse-pragmatic *like*. By examining the phenomenon of discourse-pragmatic *like* using multiple approaches, we gain a better understanding of how *like* is entering and
being used in these language varieties. Examining discourse-pragmatic *like* across both production and perception data allows us to see not just what speakers are doing with this feature, but allows us to also examine how they interpret its use within their own Local context.

Another advantage of using multiple methodological approaches is the ability to engage a diverse community of researchers. Using multiple approaches can tie disparate literatures together in such a way that researchers from different sub-disciplines can engage in meaningful discussion at these intersections. The work presented herein draws from corpus linguistics, variationist (both first wave production studies as well as third wave perception research) linguistics, as well as from the literature on perception and salience, and contextualizes these approaches within the multilingual setting of Hawai’i.

The value of using multiple approaches lies not just in the ability to speak to a wide array of disciplines, but also in the ability to make recommendations for future work on the basis of well-rounded analyses. By utilizing multiple methods and approaches to examine linguistic phenomena, we are able to “zoom out” and get a better picture of how things are related. Using more than one approach allows us to see a slightly different, more encompassing picture of the phenomena we are examining.

Because this dissertation incorporates both perception as well as production data, the results of the corpus study are contextualized and tempered by the insights gleaned from the perception experiment. The production data shows that discourse-pragmatic *like* is more frequent in Hawai’i English than in Pidgin, but the perception experiment enables us to see that while this distinction may exist numerically, listeners are not necessarily perceptually sensitive to it. Drawing from work on perception and salience allows us to consider possible explanations for this result (Campbell-Kibler, 2011; Eckert, 2011; Levon, 2014; Hay, Drager, & Gibson, 2018).

In addition, by conducting multiple analyses of the corpus data, this dissertation is able to study this ongoing language change in depth; examining not just which groups are using discourse-pragmatic *like*, but which variants are being used, and how speakers are using them. By taking this approach, we are able to see that initially, as discourse marker *like* entered the linguistic systems of Pidgin and Hawai’i English, it was used for different pragmatic functions by women and men who were early adopters of the form. Over time, these uses have stabilized and are converging with respect to how younger speakers are using discourse marker *like*. This variationist analysis of the pragmatic use of discourse marker *like* helps us to speak to the results of the perception experiment. Once we are able to see that the use of discourse-pragmatic *like* is stabilizing, and women and men are converging, it becomes far less surprising that listeners do not perceive discourse-pragmatic *like* to be more feminine. Future work should seek to take this analysis one step further, and survey participants on their explicit language ideologies as Daily-O’Cain did (2000).
The results of these multiple analyses together allow us to observe both broadly and in detail the changes that are taking place within the discourse-pragmatic systems of Pidgin and Hawai‘i English, and understand how Locals are perceiving those changes. Each of the analyses conducted in this dissertation adds a unique angle, coming together to form a more complete picture of discourse-pragmatic *like* in Hawai‘i.
6. Conclusion
This dissertation presents a detailed description of discourse-pragmatic *like* in Pidgin and Hawai‘i English. By combining multiple methodological approaches, this dissertation is able to provide a uniquely detailed lens into the changes that are occurring within the discourse-pragmatic systems of Pidgin and Hawai‘i English, as well as provide a window into how Locals perceive these changes, and how they are using the discourse marker *like* within the context of the discourse event. These results can then be compared with global patterns, highlighting ways in which discourse-pragmatic *like* in Hawai‘i patterns with globally studied Englishes, and patterns of use that are unique to Hawai‘i.

The results of the perception experiment indicate that for young Hawai‘i Locals, the presence of Pidgin is highly noticeable, and likely overrides other socially noticeable aspects of the stimuli. Results also demonstrate that for young Hawai‘i locals, women and men have similar perceptions of quotatives *be like* and *tell*, suggesting that quotative *be like* is fully accepted into the grammar for this group of speakers. Future work should seek to examine if these perceptions are held by middle-aged and older speakers of Pidgin and Hawai‘i English, or if these groups perceive discourse-pragmatic *like* differently.

While perception work conducted on other varieties shows that speakers mainly associate discourse-pragmatic *like* with women, the matched-guise study demonstrates that young Hawai‘i Locals do not perceive discourse-pragmatic *like* to be particularly feminine. This research underscores the need for locally contextualized perception research, particularly in the case of global language change, where it may be tempting to assume that language ideologies pattern similarly across varieties.

The major findings of the corpus study showed that in Hawai‘i, discourse-pragmatic *like* is patterning similarly with other varieties studied worldwide. While a change in apparent time is evident in both the frequentist and variationist analyses there are important differences to note. The frequentist analysis shows how discourse-pragmatic *like* patterns relative the grammatical functions of *like* (e.g. the verb). The variationist analysis examines how discourse marker *like* is increasing in apparent time relative to other “competing” discourse markers (e.g. *well*). Taken together, these results indicate that not only is discourse marker *like* increasing in fundamental frequency across the corpus, but that over the course of apparent time, speakers are increasingly employing it rather than using other discourse markers. The methodological implications are immediately evident. By employing both the frequentist and variationist methods, I am able to paint a more complete picture of the complexity of discourse marker *like* in Hawai‘i. Using either of these methods alone would not provide the whole story. If this study had employed only the frequentist method, we might conclude that discourse marker *like* is increasing in apparent time because speakers are using more discourse markers in general. This analysis would miss the important detail that discourse marker *like* appears to be
replacing some other variants (e.g. you know). Similarly, employing only the variationist method might lead us to the conclusion that discourse marker like is on the rise simply because all variables of like are on the rise. By employing the frequentist analysis, we are able to see that in actuality, several of the grammatical variables have remained stable over the course of apparent time.

Interestingly, both the frequentist and variationist analyses show young Pidgin speaking men use discourse marker like at higher rates than women. This finding is different from patterns observed in other studied varieties, where young men are using discourse marker like at lower rates than young women (D'Arcy, 2007). If Pidgin speaking men continue to use discourse marker like at higher rates than women, it is possible that this feature will begin to be associated with masculinity in Hawai‘i. This trend does not hold for Hawai‘i English, presenting the interesting possibility that in the future, discourse marker like may come to be associated with Pidgin, and potentially localness as well. Future perception work should be conducted with younger groups of speakers to examine these possibilities and look for emerging stereotypes.

A novel approach developed in this dissertation examines how speakers are using discourse marker like within the context of the surrounding discourse – for example, to elaborate, clarify, or provide illustrative commentary – and then quantifying the patterns of use. By providing a method for understanding the pragmatic uses that speakers employ specific forms for, this dissertation has taken an innovative step towards understanding what social factors may be driving the changes that are occurring within the system. This dissertation examines not just which demographic groups are using discourse-pragmatic like, but which functions of discourse marker like are being used, and what social work they are performing for the speakers who use them. By taking this unique approach, this dissertation is able to show that initially, as discourse marker like entered the linguistic systems of Pidgin and Hawai‘i English, it was used for different pragmatic functions by the women and men who were early adopters of the form. Over time, these uses have stabilized and are may now be converging with respect to how younger speakers use the discourse marker like.

By combining perception, production, and taking the novel approach of examining how speakers are using discourse-marker like, this dissertation sheds new light on the phenomenon of language change within a local context. In taking this innovative approach, this dissertation lays the groundwork for future study of discourse pragmatic like in Hawai‘i, and provides methodological recommendations for future work conducted on discourse-pragmatic elements.
7. Appendix

Table 9 displays a frequentist analysis of *like* in Pidgin with normalized frequency calculated per 100 words of the corpus.

<table>
<thead>
<tr>
<th>Token</th>
<th>Middle Aged Speakers</th>
<th>Younger Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Normalized Frequency</td>
</tr>
<tr>
<td>aadv</td>
<td>18</td>
<td>0.063%</td>
</tr>
<tr>
<td>comp</td>
<td>13</td>
<td>0.0454%</td>
</tr>
<tr>
<td>conj</td>
<td>13</td>
<td>0.0454%</td>
</tr>
<tr>
<td>dm</td>
<td>42</td>
<td>0.147%</td>
</tr>
<tr>
<td>dp</td>
<td>39</td>
<td>0.136%</td>
</tr>
<tr>
<td>prep</td>
<td>51</td>
<td>0.178%</td>
</tr>
<tr>
<td>q</td>
<td>23</td>
<td>0.080%</td>
</tr>
<tr>
<td>sadv</td>
<td>1</td>
<td>0.003%</td>
</tr>
<tr>
<td>v</td>
<td>26</td>
<td>0.091%</td>
</tr>
<tr>
<td>v1</td>
<td>50</td>
<td>0.175%</td>
</tr>
</tbody>
</table>

Table 10 displays a frequentist analysis of *like* in Pidgin as produced by women and men.

<table>
<thead>
<tr>
<th>Like Type</th>
<th>Women</th>
<th>Normalized Frequency</th>
<th>Men</th>
<th>Normalized Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aadv</td>
<td>21</td>
<td>0.038%</td>
<td>41</td>
<td>0.075%</td>
</tr>
<tr>
<td>comp</td>
<td>6</td>
<td>0.011%</td>
<td>12</td>
<td>0.022%</td>
</tr>
<tr>
<td>conj</td>
<td>4</td>
<td>0.007%</td>
<td>11</td>
<td>0.020%</td>
</tr>
<tr>
<td>dm</td>
<td>34</td>
<td>0.062%</td>
<td>104</td>
<td>0.190%</td>
</tr>
<tr>
<td>dp</td>
<td>58</td>
<td>0.106%</td>
<td>84</td>
<td>0.153%</td>
</tr>
<tr>
<td>prep</td>
<td>29</td>
<td>0.053%</td>
<td>71</td>
<td>0.130%</td>
</tr>
<tr>
<td>q</td>
<td>34</td>
<td>0.062%</td>
<td>41</td>
<td>0.075%</td>
</tr>
<tr>
<td>sadv</td>
<td>0</td>
<td>0.000%</td>
<td>1</td>
<td>0.002%</td>
</tr>
<tr>
<td>v</td>
<td>22</td>
<td>0.040%</td>
<td>26</td>
<td>0.047%</td>
</tr>
<tr>
<td>v1</td>
<td>62</td>
<td>0.113%</td>
<td>43</td>
<td>0.078%</td>
</tr>
</tbody>
</table>
Table 11 displays the raw numbers and normalized frequency (per 100 words) for each type of *like* displayed in Figure 8 above.

**Table 11**: normalized frequency of *like* in Hawai’i English

<table>
<thead>
<tr>
<th>Older speakers</th>
<th>Middle aged speakers</th>
<th>Younger speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>like</em> type</td>
<td>n</td>
<td>normalized frequency</td>
</tr>
<tr>
<td>aadv</td>
<td>29</td>
<td>0.058%</td>
</tr>
<tr>
<td>comp</td>
<td>13</td>
<td>0.026%</td>
</tr>
<tr>
<td>conj</td>
<td>17</td>
<td>0.034%</td>
</tr>
<tr>
<td>dm</td>
<td>12</td>
<td>0.024%</td>
</tr>
<tr>
<td>dp</td>
<td>25</td>
<td>0.050%</td>
</tr>
<tr>
<td>prep</td>
<td>82</td>
<td>0.165%</td>
</tr>
<tr>
<td>q</td>
<td>6</td>
<td>0.012%</td>
</tr>
<tr>
<td>sadv</td>
<td>0</td>
<td>0.000%</td>
</tr>
<tr>
<td>suffix</td>
<td>1</td>
<td>0.002%</td>
</tr>
<tr>
<td>v</td>
<td>22</td>
<td>0.044%</td>
</tr>
</tbody>
</table>

**Table 12**: Normalized frequency of *like* in Hawai’i English as produced by women and men

<table>
<thead>
<tr>
<th>Women</th>
<th>Middle aged speakers</th>
<th>Younger speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>like</em> type</td>
<td>n</td>
<td>normalized frequency</td>
</tr>
<tr>
<td>aadv</td>
<td>79</td>
<td>0.105%</td>
</tr>
<tr>
<td>comp</td>
<td>30</td>
<td>0.040%</td>
</tr>
<tr>
<td>conj</td>
<td>36</td>
<td>0.048%</td>
</tr>
</tbody>
</table>
Table 13 shows raw numbers and proportional frequency of discourse markers in Pidgin for middle aged and younger speakers. As with Figure 10 above, proportion here is calculated via a particular discourse marker out of all discourse markers analyzed.

**Table 13: Proportional frequency of discourse markers in Pidgin by age**

<table>
<thead>
<tr>
<th>token</th>
<th>n</th>
<th>frequency</th>
<th>token</th>
<th>n</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>brah</td>
<td>1</td>
<td>0.131%</td>
<td>brah</td>
<td>3</td>
<td>0.404%</td>
</tr>
<tr>
<td>but</td>
<td>149</td>
<td>19.554%</td>
<td>but</td>
<td>117</td>
<td>15.768%</td>
</tr>
<tr>
<td>eh</td>
<td>11</td>
<td>1.444%</td>
<td>eh</td>
<td>9</td>
<td>1.213%</td>
</tr>
<tr>
<td>ho</td>
<td>11</td>
<td>1.444%</td>
<td>ho</td>
<td>38</td>
<td>5.121%</td>
</tr>
<tr>
<td>I mean</td>
<td>26</td>
<td>3.412%</td>
<td>I mean</td>
<td>37</td>
<td>4.987%</td>
</tr>
<tr>
<td>just</td>
<td>15</td>
<td>1.969%</td>
<td>just</td>
<td>19</td>
<td>2.561%</td>
</tr>
<tr>
<td>like</td>
<td>42</td>
<td>5.512%</td>
<td>like</td>
<td>94</td>
<td>12.669%</td>
</tr>
<tr>
<td>oh</td>
<td>82</td>
<td>10.761%</td>
<td>oh</td>
<td>147</td>
<td>19.811%</td>
</tr>
<tr>
<td>so</td>
<td>114</td>
<td>14.961%</td>
<td>so</td>
<td>101</td>
<td>13.612%</td>
</tr>
<tr>
<td>well</td>
<td>58</td>
<td>7.612%</td>
<td>well</td>
<td>28</td>
<td>3.774%</td>
</tr>
<tr>
<td>yeah</td>
<td>56</td>
<td>7.349%</td>
<td>yeah</td>
<td>110</td>
<td>14.825%</td>
</tr>
<tr>
<td>you know</td>
<td>197</td>
<td>25.853%</td>
<td>you know</td>
<td>39</td>
<td>5.256%</td>
</tr>
</tbody>
</table>

Table 13 shows raw numbers and proportional frequency of discourse markers in Pidgin for women and men.

**Table 14: Proportional frequency of discourse markers in Pidgin by gender**

<table>
<thead>
<tr>
<th>women middle-aged speakers</th>
<th>younger speakers</th>
<th>men middle-aged speakers</th>
<th>younger speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>token</td>
<td>n</td>
<td>proportional frequency</td>
<td>token</td>
</tr>
<tr>
<td>brah</td>
<td>1</td>
<td>0.152%</td>
<td>1</td>
</tr>
<tr>
<td>but</td>
<td>81</td>
<td>12.291%</td>
<td>59</td>
</tr>
<tr>
<td>eh</td>
<td>8</td>
<td>1.214%</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 15 shows the raw numbers and proportional frequency for each discourse marker represented in Figure 14 above.

Table 15: Proportional frequency of discourse markers in Hawai'i English

<table>
<thead>
<tr>
<th>token</th>
<th>older speakers</th>
<th>proportional frequency</th>
<th>middle-aged speakers</th>
<th>proportional frequency</th>
<th>younger speakers</th>
<th>proportional frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ho</td>
<td>1</td>
<td>0.152%</td>
<td>ho</td>
<td>10</td>
<td>1.183%</td>
<td>24</td>
</tr>
<tr>
<td>I mean</td>
<td>16</td>
<td>2.428%</td>
<td>I mean</td>
<td>10</td>
<td>1.183%</td>
<td>30</td>
</tr>
<tr>
<td>just</td>
<td>4</td>
<td>0.607%</td>
<td>just</td>
<td>11</td>
<td>1.302%</td>
<td>12</td>
</tr>
<tr>
<td>like</td>
<td>14</td>
<td>2.124%</td>
<td>like</td>
<td>28</td>
<td>3.314%</td>
<td>75</td>
</tr>
<tr>
<td>oh</td>
<td>44</td>
<td>6.677%</td>
<td>oh</td>
<td>38</td>
<td>4.497%</td>
<td>52</td>
</tr>
<tr>
<td>so</td>
<td>53</td>
<td>8.042%</td>
<td>so</td>
<td>61</td>
<td>7.219%</td>
<td>66</td>
</tr>
<tr>
<td>well</td>
<td>20</td>
<td>3.035%</td>
<td>well</td>
<td>38</td>
<td>4.497%</td>
<td>17</td>
</tr>
<tr>
<td>yeah</td>
<td>22</td>
<td>3.338%</td>
<td>yeah</td>
<td>34</td>
<td>4.024%</td>
<td>57</td>
</tr>
<tr>
<td>you know</td>
<td>87</td>
<td>13.202%</td>
<td>you know</td>
<td>110</td>
<td>13.018%</td>
<td>34</td>
</tr>
</tbody>
</table>

Table 16: Proportional frequency of discourse markers in Pidgin as produced by women

<table>
<thead>
<tr>
<th>token</th>
<th>older speakers</th>
<th>proportional frequency</th>
<th>middle-aged speakers</th>
<th>proportional frequency</th>
<th>younger speakers</th>
<th>proportional frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>brah</td>
<td>0</td>
<td>0.000%</td>
<td>brah</td>
<td>0</td>
<td>0.000%</td>
<td>brah</td>
</tr>
<tr>
<td>but</td>
<td>97</td>
<td>4.340%</td>
<td>but</td>
<td>160</td>
<td>6.987%</td>
<td>but</td>
</tr>
<tr>
<td>eh</td>
<td>1</td>
<td>0.045%</td>
<td>eh</td>
<td>0</td>
<td>0.000%</td>
<td>eh</td>
</tr>
<tr>
<td>ho</td>
<td>1</td>
<td>0.045%</td>
<td>ho</td>
<td>0</td>
<td>0.000%</td>
<td>ho</td>
</tr>
<tr>
<td>I mean</td>
<td>8</td>
<td>0.358%</td>
<td>I mean</td>
<td>45</td>
<td>1.965%</td>
<td>I mean</td>
</tr>
<tr>
<td>just</td>
<td>7</td>
<td>0.313%</td>
<td>just</td>
<td>14</td>
<td>0.611%</td>
<td>just</td>
</tr>
<tr>
<td>like</td>
<td>6</td>
<td>0.268%</td>
<td>like</td>
<td>25</td>
<td>1.092%</td>
<td>like</td>
</tr>
<tr>
<td>Token</td>
<td>Older Speakers</td>
<td>Proportional Frequency</td>
<td>Middle-Aged Speakers</td>
<td>Proportional Frequency</td>
<td>Younger Speakers</td>
<td>Proportional Frequency</td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>------------------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>brah</td>
<td>0</td>
<td>0.000%</td>
<td>brah</td>
<td>1</td>
<td>0.038%</td>
<td>brah</td>
</tr>
<tr>
<td>but</td>
<td>172</td>
<td>6.464%</td>
<td>but</td>
<td>151</td>
<td>5.675%</td>
<td>but</td>
</tr>
<tr>
<td>eh</td>
<td>17</td>
<td>0.639%</td>
<td>eh</td>
<td>3</td>
<td>0.113%</td>
<td>eh</td>
</tr>
<tr>
<td>ho</td>
<td>1</td>
<td>0.038%</td>
<td>ho</td>
<td>7</td>
<td>0.263%</td>
<td>ho</td>
</tr>
<tr>
<td>I mean</td>
<td>37</td>
<td>1.390%</td>
<td>I mean</td>
<td>71</td>
<td>2.668%</td>
<td>I mean</td>
</tr>
<tr>
<td>just</td>
<td>6</td>
<td>0.225%</td>
<td>just</td>
<td>20</td>
<td>0.752%</td>
<td>just</td>
</tr>
<tr>
<td>like</td>
<td>6</td>
<td>0.225%</td>
<td>like</td>
<td>17</td>
<td>0.639%</td>
<td>like</td>
</tr>
<tr>
<td>oh</td>
<td>87</td>
<td>3.269%</td>
<td>oh</td>
<td>84</td>
<td>3.157%</td>
<td>oh</td>
</tr>
<tr>
<td>so</td>
<td>222</td>
<td>8.343%</td>
<td>so</td>
<td>231</td>
<td>8.681%</td>
<td>so</td>
</tr>
<tr>
<td>well</td>
<td>48</td>
<td>1.804%</td>
<td>well</td>
<td>46</td>
<td>1.729%</td>
<td>well</td>
</tr>
<tr>
<td>yeah</td>
<td>69</td>
<td>2.593%</td>
<td>yeah</td>
<td>135</td>
<td>5.073%</td>
<td>yeah</td>
</tr>
<tr>
<td>you know</td>
<td>161</td>
<td>6.050%</td>
<td>you know</td>
<td>182</td>
<td>6.840%</td>
<td>you know</td>
</tr>
</tbody>
</table>
8. References


Butters, R. (1982). Editor's note [on be like 'think']. American Speech, 149.


Höhn, N. (2012). "And they were all like 'what's going on?'": New quotatives in Jamaican and Irish English. In M. Hundt, & U. Gut (Eds.), *Mapping unity and diversity world-wide: Corpus-based studies of...*


