

He nui nā ala e hiki aku ai: Factors Influencing Phonetic Variation In The Hawaiian Word *kēia*

Katie Drager, Bethany Kaleialohapau‘ole Chun Comstock, and
Hina Puamohala Kneubuhl

University of Hawai‘i at Mānoa

Apart from a handful of studies (e.g., Kinney 1956), linguists know little about what variation exists in Hawaiian and what factors constrain the variation. In this paper, we present an analysis of phonetic variation in the word *kēia*, meaning ‘this’, examining the social, linguistic, and probabilistic factors that constrain the variation. The word *kēia* can be pronounced with a constricted glottis (e.g., as creak or a glottal stop) or without one (Pukui & Elbert 1986: 142) and, like many words in Hawaiian, it can undergo phonetic reduction. The analysis was conducted on interviews with eight native-speaking *kūpuna* (elders) who were recorded in the 1970s. We find that the likelihood of the word being realized with a constricted glottis decreases if the word immediately following *kēia* begins with an oral stop or if the speaker is a man. Additionally, we observe a higher likelihood of phonetic reduction as word sequences (*kēia* + the following word(s)) are repeated during the interaction. The results contribute to current models of speech production and planning, and they inform work aimed at supporting the ongoing efforts to conserve and revitalize the Hawaiian language.

Ma waho a‘e o kekahi mau papahana noi‘i he ‘u‘uku wale (e la‘a ‘o Kinney 1956), ‘a‘ole nō nui ka ‘ike o ka po‘e kālai ‘ōlelo e pili ana i nā ‘ano like ‘ole o ka ho‘opuka ‘ia ‘ana o ka ‘ōlelo Hawai‘i a me kēlā mea kēia mea ho‘i e kaupalena ana i ia mau ‘ano. Ma ka papahana noi‘i nei, hō‘ike ‘ia aku ko mākou kālailai ‘ana i ka lolina kani leo o ka hua‘ōlelo ‘o *kēia*, ‘o ia ho‘i ‘o *this* ma ka ‘ōlelo haole, me ka nānā pū aku nō ho‘i i nā mea e kaupalena ana i ia lolina, ‘o ia ho‘i ka launa kanaka, nā loina ‘ōlelo, a me nā pili papaha. Ho‘opuka ‘ia ‘o *kēia* me ka ‘okina (ma ke ‘ano he ‘okina maoli a i ‘ole he ‘u‘ina pōkole) a me ka ‘ole o ka ‘okina (Pukui lāua ‘o Elbert 1986: 142). E like me nā hua‘ōlelo Hawai‘i he nui, ke ho‘opuka ‘ia mai ‘o *kēia*, lohe ‘ia ka mokuna kani leo i kekahi manawa. Ua kālailai ‘ia aku ka ‘ōlelo a ‘ewalu kūpuna mānaleo i ho‘opa‘a ‘ia ko lākou mau leo ma ka lola ma nā makahiki kanahiku. ‘O kekahi hua o *kēia* noi‘i, emi mai ka nui o ka lohe ‘ia o ka mokuna kani leo inā he hua ho‘okū ko mua o ka hua‘ōlelo e puka mai ana ma hope pono o *kēia*, a i ‘ole he kāne ka mea e ‘ōlelo ana. Eia hou, nui a‘e ka lohe ‘ia o ka mokuna kani leo inā kūpina‘i mai ka ho‘opuka ‘ia ‘ana o kekahi hopuna‘ōlelo (*kēia* + kekahi [mau] hua‘ōlelo) i loko o ke kama‘ilio ‘ana. He ho‘omaika‘i aku a ho‘ololi iki a‘e paha *kēia* mau hua noi‘i i nā kumu ho‘ohālike no ka ho‘opuka ‘ana i ka ‘ōlelo a me nā hana ho‘olālā ‘ōlelo, a he ho‘onui ‘ike nō ho‘i ia e kākō‘o ana i ka po‘e e alu nei i ka mālama a ho‘ōla hou ‘ia ‘ana o ka ‘ōlelo Hawai‘i.

1. INTRODUCTION. Healthy languages exhibit a great deal of systematic variation. Much of this variation is predictable based on language-internal factors (e.g., phonological environment), social factors (e.g., aspects of the speaker's identity), or an interaction between such factors. Speakers produce the variation in consistent and predictable ways (Labov 1966; Labov et al. 1972; Cedergren 1973; Milroy 1980), and the variation influences what social characteristics are attributed to the speaker (Campbell-Kibler 2007; Levon 2007) even when speakers and listeners are not aware that the variation exists.¹

When language loss occurs, overgeneralization of certain features can occur, and obligatory rules can become optional (Rankin 1978; Campbell & Muntzel 1989), ultimately resulting in a simpler linguistic system than what was present when the language was healthy. In such cases, linguistic variation is reduced, leaving gaps in the speakers' (socio)linguistic repertoires (Anderson 1982). Much of the work examining the simplification of variables in endangered languages has been concerned with addressing why the incoming variant takes the form it does, demonstrating, for example, that the outcome is influenced by markedness² (Campbell & Muntzel 1989) and that different speakers or groups of speakers can have different strategies in how they overgeneralize (Hill 1983: 267).

But a number of questions about variation in endangered languages remain, such as: how does an innovative form diffuse through a speech community?³ In what ways do linguistic and social factors interact in their effects on form variation and change? And, when a language is being revitalized, can and do learners acquire variation that is affected by internal and external factors and interactions between such factors? Variationist sociolinguistic methods are well-suited to answer such questions. They can be used to examine change as it occurs, exploring how the change begins and subsequently diffuses through a community while also considering the effects of (and interactions between) internal and external pressures on different types of linguistic variables. This is exemplified by numerous variationist studies that have examined, for example, change within the context of language and/or dialect contact (e.g., Britain 2002; Kerswill 2003; Cheshire et al. 2011).

Variationists do not assume that a linguistic form is categorical in a language or dialect, meaning that variation can be systematic even if the linguistic constraints on the form do not affect the output 100% of the time. For example, different speakers

¹ We would like to thank Clinton Kanahale for conducting the interviews, and the Joseph F. Smith Library at Brigham Young University – Hawai'i for granting us permission to use the recordings for research purposes. We would also like to thank the kūpuna for graciously agreeing to be recorded, and for sharing their stories and their language with us. Without the efforts and generosity of all of those above, this work would not exist. We would like to thank Kamuela Yim for his help with the abstract, and Keao NeSmith, 'ŌiwiHawaiiokalani Parker Jones, Kristine Hildebrandt and two anonymous reviewers for their helpful comments on earlier versions of this paper. We would also like to acknowledge Robert Fromont and Aitor Alvarez for their technical support. Of course, all errors remain our own.

² In order to make this paper more readable to non-linguists who are interested in Hawaiian language, we include a number of footnotes to explain linguistics jargon. Here, *markedness* refers to standing out because either the marked variant is different from the norm, is not found in the first language (L1), or is difficult to acquire. For a discussion (and critique) of the range of uses of the term, see Haspelmath (2006).

³ It is possible that the modes of diffusion in endangered language communities are very different from those observed in communities where the language is not threatened. See, for example, the manner in which lexical items in Rapa (Austral Islands, French Polynesia) are undergoing change (Walworth 2015).

favor some variants over others, and single individuals shift which variant they use depending on the social context (Podesva 2007) and their social goals during a single interaction (Bucholtz 2011). Knowing which variant to use when, with whom, and at what rate is part of a speaker's communicative competence (Hymes 1972). Taking a variationist sociolinguistic approach when documenting endangered and understudied languages has the potential to make substantial contributions to sociolinguistic theory since it can shed light on the wide range of culturally-determined social information that influences variation, and it can uncover linguistic variation not found in more commonly studied languages. Furthermore, variationist sociolinguistic work – as a form of language documentation – can inform revitalization efforts. Descriptions of variation should, therefore, be included in efforts to document and conserve a language.

The goals of this paper are twofold: (1) to present results from a variationist sociolinguistic analysis of phonetic variation in Hawaiian and (2) to present an argument for integrating variationist sociolinguistic methods into language documentation efforts more generally. We begin in Section 2 by arguing for the benefits of conducting variationist sociolinguistic work on lesser-studied languages and on languages that are threatened or endangered, in particular. In Sections 3 and 4, we provide an example, presenting results from an analysis of glottal stop variation in the Hawaiian word *kēia*, meaning 'this'. In Section 5, we discuss the theoretical implications of the findings from our study as well as a description of our efforts to disseminate the information regarding the variation. Finally, we make a call for variationist sociolinguists to work closely with language documentarians in order to describe the large amount of variation that often goes unstudied for endangered and threatened languages.

2. BACKGROUND. In this section, we discuss the benefits of conducting variationist sociolinguistic work as part of a language documentation project. To do this, we draw on previous work that has used this approach to study endangered and/or understudied languages as well as studies that investigate second language learners' acquisition of sociolinguistic variables.⁴ Then, in order to inform the analysis and discussion of variation in Hawaiian, we turn to a discussion of probabilistic factors influencing phonetic reduction, which is defined and discussed in Section 2.2. In Section 2.3, we present a discussion of Hawaiian and Hawai'i-based language contact.

2.1 Variationist sociolinguistics as language documentation.

2.1.1 Social factors and linguistic variation. The vast majority of work that uses a variationist sociolinguistic approach is based on major European languages, and English especially. As such, most of the work is conducted in Western societies and in major cities (though see contributions to Stanford & Preston 2009). This means that our understanding of sociolinguistic variation and change is based largely on a small percentage of the ways in which language is used in societies around the world. If we focus our efforts on a wider array of cultures, we can examine social factors that are different or absent from those found in more commonly studied societies. Likewise, social factors that have typically been the focus of variationist sociolinguistic work

⁴ A linguistic *variable* is a linguistic entity that has more than one form, such as (ING) in the word *fishing*, which can be realized with a velar nasal (*fishɪŋ*) or an alveolar nasal (*fishɪn*). A *variant* is one of the forms of a variable.

may be irrelevant, or may take an unexpected form. Work using a variationist sociolinguistic approach to uncover variation in lesser-studied languages demonstrates a wider range of social factors than what is generally described. For example, Stanford (2009) describes how clan serves as an influential social factor in Sui villages in China. As discussed in the other contributions in this volume, other work that has used a variationist sociolinguistic approach to study variation in minority (and often understudied) languages includes Eckert (1980), Maclagan et al. (2009), Babel (2008), Rau (2009), and Nagy (2011).

Additionally, studying underdescribed languages from a variationist perspective also means that we can look at less typologically-common linguistic variables. This is especially important for sociolinguistic theory given that different variables behave differently from one another and attract the speaker's attention to varying degrees.

2.1.2 Using variationist sociolinguistic work to inform language revitalization.

Many of the world's languages remain undescribed from a variationist standpoint, so why do we believe that the focus of our efforts should be on those languages that are endangered or threatened? First, there is of course the very real danger that the languages will only continue to exist for a short time. Second, language shift is often a result of the speakers' social and political situation, potentially giving socio-political and attitudinal factors a different (and potentially greater) role in their relationship to linguistic variation. Third, descriptions of sociolinguistic variation can be used to inform language revitalization, including the regeneration, renewal, and reclamation of the language.

A crucial part of speaking a language is knowing which linguistic variants to use across different contexts and for what social goals. Some variants are more frequent in formal contexts (Labov 1972), some vary as a function of characteristics attributed to an interlocutor (Bell 1984), and some contribute to a speaker's personal style (Eckert 2000; Mendoza-Denton 2008). Failing to produce the appropriate linguistic variants in a given situation can lead to confusion, misunderstanding, and negative judgments (Labov 1972).⁵ This aspect of linguistic knowledge falls under communicative competence⁶ (Hymes 1972), the absence of which "is often painfully salient to discriminating listeners" (L. Wong 1999: 95). It is something that speakers of healthy languages learn through exposure and without explicit instruction, but when exposure is impaired, so is a speaker's communicative competence. Therefore, some scholars have argued that "mastering [variable rules] should be looked upon as an important pedagogical goal in second language teaching" (Rehner & Mougeon 1999: 126). Variationist sociolinguistics provides a means of describing the social and linguistic factors related to the variation. Instructors and learners alike can then use the findings, teaching and learning communicative competence in the language as appropriate.

Variationist sociolinguistics is a descriptive (as opposed to prescriptive) enterprise: we describe the patterns we observe in speech and assume there is not a "right way" to say something. Non-categorical inter- and intra-speaker variation is widely

⁵ Of course – in the interest of social progress – sometimes it may be necessary and appropriate to challenge what is considered acceptable. However, not conforming can be a social or political stance, and is not the same as failing to conform due to lack of exposure.

⁶ *Communicative competence* is a term used to refer to a speaker's grammatical and social knowledge of language use, as exemplified by how they use it.

observed; speakers draw from different parts of their linguistic repertoires in sometimes nuanced ways in order to meet both their communicative needs and their social goals. While the nuanced aspects of variation may be difficult if not impossible to master by second language learners, the authors of this paper assume that many language learners wish to approximate native speakers' ways of talking as closely as possible, including the range of styles found within the speech of a single native-speaking individual.⁷ When focusing on threatened and endangered languages such as Hawaiian, it is important to remember that changes due to language loss are not the same as changes that occur in communities where the language is healthy "for the very fact that non-native speakers of Hawaiian disproportionately influence the direction of that change" (K.L. Wong 2011:152). Thus, describing the variation that is found in the speech of native speaking elders serves to shed light on linguistic variants that appear to have been widespread in conversational Hawaiian at one time but that are less common today following the break in intergenerational transmission. That said, we wish to be clear that we are not saying there is a "correct" way to use the variants; indeed, different individuals who are native speakers use them differently and to different degrees. Instead, our hope is that – through the description of the variation and the dissemination of the findings – learners are better able to comprehend the speech of native speakers and that, with awareness of the variables, comes the freedom to use each of the variants some of the time.

To what extent do language learners acquire sociolinguistic variables without explicit instruction? Research on the acquisition of sociolinguistic variables shows that, if exposed to the variation in natural speech, learners can acquire it, but they need a large amount of exposure and, even among immersion students, the pattern is weaker than for native speakers (Warner 1996; Mougeon et al. 2004). Additionally, learners acquire some variables more readily than others (Warner 1996; Mougeon et al. 2004). One way for learners to acquire a sociolinguistic variable is through explicit instruction (Lyster 1994): once their attention is drawn to the variable and the contexts in which each variant tends to be used, learners are more likely to use the variant in the appropriate contexts. As Tarone & Swain (1995) argue in their work with immersion students, learners who acquire linguistic variables associated with informal styles (in addition to the formal-associated variants used for teaching) may feel more inclined to use the language with one another. This could, in turn, motivate greater acquisition of the language in general. Learning about linguistic variation is also likely to aid comprehension of native speech, which – for many learners – is the most difficult part of second language acquisition; being made aware that the variation exists reduces the chances that learners will misunderstand non-standard variant(s). Finally, listeners attribute at least some social characteristics to stylistic variants used by L2 speakers (Hardeman Guthrie 2016), so if learners wish to construct their identities within the context of their L2, it is helpful to know something about what sociolinguistic variants exist. Taken together, this body of work suggests that there are benefits to exposing learners to the myriad ways of talking that native speakers use.

2.2 Phonetic reduction and predictability. One advantage of analyzing variation in an understudied language is the opportunity to test linguistic theories for which most supporting evidence comes from work on heavily studied European languages. For the current study we focus on one such factor: the effect of predictability on phonetic reduction. Phonetic reduction is a process known to occur in many languages, and it

⁷ This is certainly true for the authors, who are second language speakers of Hawaiian, each at a different stage of acquiring the language. However, see NeSmith (2012) for evidence that some learners do not.

can come in many forms. When reduction occurs, sounds can be shorter (so they have a measurably shorter duration), lenited (making them less consonant-like), or elided (omitted altogether).

There is a large body of work that demonstrates how highly predictable words are more likely to undergo reduction (Lieberman 1963; Hunnicutt 1985; Gregory et al. 1999; Aylett & Turk 2004; Baker & Bradlow 2009; Bell et al. 2009; Díaz-Campos & Gradoville 2011, Hay & Foulkes 2016), confirming earlier claims along these lines (Zipf 1929). In their Smooth Signal Reduction Hypothesis, Aylett & Turk (2004) propose that a token's probability (what they refer to as language redundancy) and the phonetic reduction it undergoes (which they refer to as acoustic redundancy) are balanced over the course of an utterance, so that syllables with higher probability are more reduced, and vice versa. While much of the work in this area has focused on words (or sequences of words), there is also evidence that sounds at word boundaries are more likely to be reduced if the word sequence is highly frequent (Bush 2001) and that high predictability also leads to greater reduction for bound morphemes (Davis 2003; Blevins 2005; Pluymaekers et al. 2005, Rose et al. 2015).

The predictability of a word depends on a variety of different factors, including its token frequency (i.e., how many times a word was produced) (Zipf 1929, Bybee 2001, Gahl 2008, Hay & Foulkes 2016), the likelihood of the syntactic structure it is part of (Gahl & Garnsey 2004; Tily et al. 2009), the contextual probability of the word given the words immediately surrounding it (Jurafsky et al. 2002), and the number of times the speaker has produced the word previously in the conversation (Fowler & Housum 1987, Fowler 1988). For the present study, work related to the last two of these (contextual probability and the number of repetitions) is especially relevant, so we focus on these two probabilistic factors for the remainder of this subsection.

Phonetic reduction appears to be related to the probability of a word that stems from its surrounding context; a word is more likely to be phonetically reduced if it is likely to occur given its overall frequency combined with the frequency in which it occurs alongside the words that immediately precede and/or follow it (Gregory et al. 1999; Jurafsky et al. 2002; Bell et al. 2009). There is some evidence that such effects are only found among high frequency items (Fosler-Lussier & Morgan 1999). However, Bell et al. (2009) observe an effect of conditional probability for both content words (an open class of words that carry the semantic load of a sentence e.g., nouns and main verbs) and function words (a closed class of words that are used to make sentences grammatical e.g., prepositions and articles). Words with a higher probability of occurrence given their context are more likely to be reduced.

Additionally, there is a well-established effect of repetition on phonetic reduction; as words are repeated during the course of an interaction, their realizations tend to be phonetically reduced (Fowler & Housum 1987; Fowler 1988; Gregory et al. 1999). The effect, however, seems to be limited to content words. In a study on probabilistic factors influencing phonetic reduction in conversational speech, Bell et al. (2009) observed an effect of repetition on content words but found no such effect for function words. When taken together with work that demonstrates that high frequency words or word sequences are both looked at for less time (suggesting faster retrieval) and comprehended more quickly than those with lower frequency (Rayner & Duffy 1986; Arnon & Snider 2010), the tendency for words to undergo reduction when highly frequent suggests that probabilistic information such as token frequency is stored in the mind, or that mental representations of words are stored or tracked in a

way that permits on-line computation of the probabilistic information. Hence, the findings that phonetic factors are linked with word-based probabilistic factors are relevant to cognitive models of speech production and perception, and to our understanding of language in general. Despite the wide-ranging implications of such findings, the vast majority of work along these lines has focused on English or closely related languages, like Dutch (Pluymaekers et al. 2005), with some exceptions (Van Son et al. 2004; Zhao & Jurafsky 2007; Wiener 2012), stemming at least partially from the availability of corpus data for the more heavily studied languages. In the current study, we test the assertion that predictability influences phonetic reduction by examining whether probabilistic factors influence reduction in Hawaiian.

Following Bell et al. (2009), we examined the probability of a word given the following word, and we also tested another factor influencing predictability: the number of repetitions of word sequences containing the function word. We did this because while the previous work by Bell and colleagues found no effect of repetition for function words, they examined repetition of the function word itself rather than repetition of word sequences. We hypothesized that due to the high frequency of function words and the tendency for some degree of lexicalization⁸ to occur between function words and the words immediately surrounding them, we would observe an effect of word sequence repetition on phonetic reduction in function words that is analogous to that found with the repetition of content words.

Probabilistic factors have not yet been investigated as predictive of phonetic reduction in any Polynesian language. In this paper, we examine the roles that contextual probability and word sequence repetition play in the reduction of a function word in the Polynesian language, Hawaiian. The history of Hawaiian and the socio-historical context in which the language is embedded are complex, but they are also critical to understanding the current state of the language and the motivation behind the current study. Thus, we now turn our discussion to a brief history of the language, including select historical events that resulted in the endangerment of the language and the adverse impact on the Hawaiian people.

2.3 The shift from Hawaiian to English. Hawaiian was once a language of daily use in all domains in Hawai‘i. Hawaiian was widely spoken throughout the 1800s, with almost universal literacy among its speakers (Warschauer et al. 1997). However, like many endangered languages, a number of factors and historical events, including occupation by the United States (Sai 2011: 114-121), have contributed to its drastic decline. Contact with English began with the arrival of Captain Cook in 1778. Westerners introduced novel diseases, which resulted in the death of over 90% of the Hawaiian population (Warner 2008[1982]: 134). Another key event was the arrival of Protestant missionaries from New England in 1820; they introduced Christianity and, together with the Hawaiian people, developed the first orthography and established schools in both Hawaiian and English.

The social and linguistic landscape was further affected when American sugar plantations began to dominate the economy in the mid 1800s. With the influx of immigrant workers from Asia and Europe, the immediate need to communicate gave rise to an English-lexified creole, known locally as Pidgin and referred to by linguists as Hawai‘i Creole or Hawai‘i Creole English. Meanwhile, American businessmen were strengthening their hold on political power in the kingdom. Land was divided

⁸ ‘Lexicalization’ refers here to the storage of word sequences as word-like entities in the mental lexicon.

and distributed, but only 1% was awarded to the Hawaiian people (Warner 1999: 70). In 1893, with the support of US forces, the American businessmen engineered the overthrow of Queen Lili‘uokalani and the sovereign government of Hawai‘i (Coffman 2009; Sai 2011). By that time, only about 5% of the schools were conducted in Hawaiian. The occupying government instituted an act stating that English was to be the only language of instruction unless otherwise authorized by the government (Act 57), and by 1902 there were no longer any Hawaiian language schools (Schütz 1994: 352 gleaned from Reinecke 1969: 70-72). Children were beaten or humiliated for speaking Hawaiian, and “the illusion of future prosperity resulting from the abandonment of Hawaiian in favor of English was inculcated into the Hawaiian people” (Warner 2008: 135). Pukui tells of being hit on the head as a young child in public school, and of humiliation and revocation of privileges as a teenager (Pukui et al. 2002: 61). In some instances, the use of Hawaiian was even considered a crime (Lucas 2000: 7-8). It was not uncommon for Hawaiians to be compelled—both willingly and unwillingly—to renounce their language, culture, and even their names in order to become *haole* (i.e., like affluent whites) and, therefore, better able to achieve “success” socially, economically, and intellectually.

The effect on the vitality of Hawaiian was devastating. By the early 1980s, the only sizeable population of speakers (>200) under the age of 18 were from Ni‘ihau (NeSmith, personal communication, June 2016); outside of Ni‘ihau, it is estimated that there were fewer than 50 individuals under the age of 18 who could speak the language (Kawai‘ae‘a et al. 2007: 183; Oliveira 2014: 82).

Concerned by the declining numbers of Hawaiian speakers and cultural practitioners, a grassroots renaissance movement began as early as the 1930s (Schütz 1994: 361) and by the 1970s, the movement was in full force. It was during this time that the data for the present study was collected, spurred by the concern for the loss of the language and the desire to preserve or revitalize it. Thus, the elderly speakers included in this study grew up during a period in which speaking Hawaiian was typically discouraged or prohibited.

Thanks to efforts to revitalize the language, today there is a growing number of Hawaiian language speakers (Wilson & Kamanā 2008[1982]). Among the revitalization efforts are the *Pūnana Leo* Hawaiian immersion preschools (modeled after the *Kōhanga Reo* or “language nests” in New Zealand) and the *Papahana Kaiapuni Hawai‘i* Hawaiian-medium schools, which can be found on every major island excluding Ni‘ihau and Lāna‘i. Nearly 2,000 students were enrolled in immersion schools in the 2010-2011 school year (NeSmith 2012: 27). The language is also being taught in a number of other contexts, including the University of Hawai‘i at Mānoa, where an extensive Hawaiian language program has blossomed since the introduction of Hawaiian language classes in the 1920s (Oliveria 2014), and the University of Hawai‘i at Hilo, where the first Hawaiian language college (Ka Haka ‘Ula O Ke‘elikōlani) was established and which currently offers a doctorate in Hawaiian and Indigenous Language and Culture Revitalization. Additionally, community efforts have established other venues, such as ‘Ōiwi TV, where the language is used and/or heard.

While the revitalization efforts certainly have successes, the number of speakers who have a native-like command of the language (complete with rich sociolinguistic variation) remains small. Estimates of all highly fluent speakers who use the language daily are as low as 1,000. Thus, L2 learners make up the majority of Hawaiian language speakers today. Many of these learners have little to no contact with

native speakers of the language, so they do not have a chance to acquire more nuanced linguistic variation (Kneubuhl 2014: 5). Even immersion school students mainly interact with L2 speakers (K.L. Wong 2011: 152). In Hawai‘i, defining who counts as a native speaker is not always straight-forward. Parker Jones (2017) argues that immersion students should be counted as native speakers because many of them speak it as an L1 (first language); others point out that the variety spoken by immersion students has been influenced by L2 speech (Warner 1996; NeSmith forthcoming). But, even if we assume that all speakers who learned Hawaiian as a first language are native speakers, native speakers are still outnumbered by the number of people who have learned Hawaiian as a second language.

One of the ways in which exposure to L2 speech has influenced Hawaiian involves a decrease in linguistic variation (NeSmith 2005). Efforts to standardize Hawaiian have also led to reduction in linguistic variation (Kinney 1956: 282). A goal of the current study is to identify variation found in the speech of native-speaking kūpuna and to disseminate the information so that learners can better comprehend the variation when they encounter it and so that they can adopt the variation into their own speech should they wish to do so.

2.4 Variation in Hawaiian. Relatively little is known about what linguistic variation exists in Hawaiian and what factors influence variation that is present. Very little of the linguistic variation that has been mentioned in the literature was investigated empirically or quantitatively using natural, conversational data; exceptions can be found in Kinney (1956), Warner (1996), and Kneubuhl (2014), but much more in-depth description of variation in the speech of native speakers is needed for all levels of the grammar (Warner 1996: 262). To date, there has been no quantitative variationist sociolinguistic work conducted on Hawaiian.

There were traditionally at least two regional dialects⁹ of Hawaiian. One of the dialects, referred to as the Ni‘ihau dialect, is spoken today by all of the residents of the island of Ni‘ihau as well as a small percentage of people on Kaua‘i. The Ni‘ihau dialect is the only variety of Hawaiian spoken today for which the majority of speakers learned the language through an unbroken chain of intergenerational transmission. Other possible dialect regions include Maui, Hawai‘i island, O‘ahu (Kimura 1983:174-175) and Moloka‘i-Lāna‘i (Elbert & Pukui 1979:25), though Wilson suggests the differences may only have been minor (Wilson 1998:126).

In addition to the Ni‘ihau dialect, the other dialect that is commonly spoken today is variously referred to in English as Neo Hawaiian (NeSmith 2005) or Standard Hawaiian (Parker Jones 2016). It is based on the variety traditionally found on O‘ahu (Kimura 1983:175) and Hawai‘i Island (Parker Jones, personal communication June 2016), and it is used largely by immersion students and second language learners, some of whom are highly proficient in the language and use it as their main mode of conversation. NeSmith (2005, forthcoming) argues that linguistic forms found in this variety have been heavily influenced by the speakers’ extensive exposure to L2 speech. This would be unsurprising given the prolonged gap in intergenerational transmission for most families who do not live on Ni‘ihau or Kaua‘i. Work by Warner (1996) on syntactic variation in the speech of Hawaiian immersion children lends some support to this claim (Warner 1996:128-132) though, for some variables, the

⁹ In linguistics, the term *dialect* is used to describe a variety that is distinct from another variety of the same language. Thus, everyone speaks a dialect. It does not have the negative connotations that are sometimes present for the word when used by non-linguists.

patterns exhibited by immersion students were similar to those used by *kūpuna*. The full extent to which influence from L2 has occurred remains an open question.

There are a number of differences between the Ni‘ihau dialect and varieties spoken on the other islands. For example, contrary to widely held beliefs that Hawaiian has no alveolar plosive, [t] is a common allophone¹⁰ of /k/ in the Ni‘ihau dialect (Newbrand 1951: 106-107), and there is some evidence that /t/ occurs phonemically in the Ni‘ihau dialect as well (A.K. Wong 2010: 161). The alveolar plosive [t] can also be found in other dialects allophonically, especially in ritualized contexts such as *oli* ‘chants’, but it is not as common as on Ni‘ihau. There is also a number of lexical differences between the Ni‘ihau dialect and other varieties (Elbert & Pukui 1979:23-27), as well as additional phonological differences, such as variation between [l] and [r] as well as between [l] and [n] (Tava & Keale 1989:15). The faster speech rate of Ni‘ihau dialect speakers has also been noted.

Kinney (1956) provides a quantitative analysis of phonetic variation in select lexical items. Her analysis was conducted on conversational data from native-speaking *kūpuna* who were born between 1865 and 1905. The results reveal a great deal of phonetic variation in their speech, especially in regard to raising of /ai/ to [ei] in select lexical items. Further analysis of this vowel in other lexical items is necessary to determine what factors influence the variation across different contexts and speakers, and within the speech of a single individual.

NeSmith (2005) describes a number of other linguistic variables that are found in the speech of native-speaking *kūpuna*. This work provides an excellent starting place for variationist sociolinguistic work because it points to linguistic variables that may co-vary with linguistic or social factors.¹¹ In this paper, we focus on phonetic variation in the word *kēia*, meaning ‘this’. The phonetic realization of *kēia* is not a variable that is listed in NeSmith (2005) but, like those variables, it is one that varies in the speech of native-speaking *kūpuna* who learned the language through intergenerational transmission but tends to be homogenous in the speech of many speakers who learned Hawaiian as a second language.

According to Pukui & Elbert, there are two variants of *kēia*: one with a glottal stop [keʔia]¹² and one with a long vowel [ke:ia] (Pukui & Elbert 1986: 142). The variant with the long vowel is consistent with the standard spelling of the word. While some textbooks explicitly state that there are multiple ways of pronouncing certain words (e.g., Hopkins, 1992: 2), the full range of phonetic variants that are found in conversational Hawaiian, including the various realizations of *kēia*, is not found in textbooks. The realization with the long vowel is generally the only one that is taught in the classroom, and it is the form that is predominantly used by most second language speakers. One reason for this is that most L2 speakers have only a limited amount of exposure to L1 speech, so they do not have the opportunity to achieve an L1-like command of the range of phonetic realizations of this word. A second reason is that many L2 learners’ use of Hawaiian has been influenced by spelling (K.L. Wong

¹⁰ An *allophone* is a variant of a sound that is actually produced. For example, the sound /p/ in English is pronounced differently depending on whether it is preceded by a /s/ or is word initial, and these different pronunciations are called *allophones*.

¹¹ Preliminary findings from an acoustic phonetic analysis of vowels from two speakers (one a native speaker from Ni‘ihau and the other a fluent L2 speaker from O‘ahu) provides some evidence of differences between their vowel realizations (Piccolo 2005).

¹² We are using the International Phonetic Alphabet (IPA) for the transcriptions. The symbol [ʔ] is used for the glottal stop which in Hawaiian is written using the ‘okina, and [:] is used to denote a long sound.

1999). A third reason is that the variable has never been studied, so instructors (many of whom are second language learners themselves) do not yet know how to explain when to use each variant. Yet linguistic variation is viewed as an important aspect of the Hawaiian language by many of its speakers: “those who insist upon standardizing the language may not realize the danger of rendering any Polynesian dialect, or any other language for that matter, nearly incapable of flexibility” (Kinney 1956: 282). Variation provides speakers with the opportunity to do more with language than simply convey the semantic content; through using certain variants in their speech, they can construct aspects of their identities (e.g., use a linguistic variant to signal that they are from a given island) and shift in subtle ways to imply nuanced meanings (see Wong 2011).

2.5 Stress and rhythm in Hawaiian. One of the variables we are investigating – degree of phonetic reduction – is linked to stress and rhythm. For words in isolation in Hawaiian, long vowels and diphthongs obligatorily carry stress, and stress can also occur on short vowels (Elbert & Pukui 1979:14). Phonetic reduction and changes in stress are common when speaking Hawaiian quickly (Kinney 1956; Elbert & Pukui 1979:22). The word *kēia* is normally unstressed unless it is at an intonational phrase boundary¹³ or carries focus. Our analysis focuses on those instances where it does not carry stress (i.e., where we expect some degree of reduction). Unstressed tokens make up the vast majority of tokens that are available in the dataset.

How Hawaiian should be classified in terms of its rhythm is less clear. Work using the Pairwise Variability Index (PVI) suggests that Hawaiian is stress-timed¹⁴ (Parker Jones 2006), whereas some scholars imply that it is syllable-timed (Cutler 1991: 159). Additionally, some scholars have suggested that a closely-related language, Māori, is mora-timed (Bauer 1981).¹⁵ Regardless, it is safe to say that a great deal of phonetic reduction can be observed in unstressed syllables in conversational Hawaiian produced by native-speaking elders.

3. METHODOLOGY.

3.1 Speakers. The data come from eight speakers, born between 1884 and 1894, who were interviewed by Clinton Kanahēle in 1970. Of the eight speakers, four are male and four are female. The home islands of the eight speakers are O‘ahu (the most heavily populated island) and Hawai‘i (the largest island and the one from which the archipelago gets its name). The distribution of speakers is shown in Table 1.

	Hawai‘i	O‘ahu
males	2	2

¹³ An intonational phrase roughly corresponds to a clause but is determined according to intonation (the rise and fall of pitch when speaking).

¹⁴ Some scholars argue that languages can be identified as stress-timed (stressed syllables are longer than unstressed syllables), syllable-timed (syllables are roughly equal length), and mora-timed (syllable weight determines length). As examples, English is said to be stress-timed, Spanish is said to be syllable-timed, and Japanese is said to be mora-timed.

¹⁵ While Māori is often assumed to be mora-timed, there is little evidence to support this claim. Additionally, the extent to which Hawaiian and Māori have the same rhythm is unclear.

females	2	2
---------	---	---

TABLE 1: The distribution of speakers, by island and gender.

The interviews were conducted for the purpose of preserving the Hawaiian language, and individuals with expertise in Hawaiian history, culture, and language were sought to take part. Because different individuals possessed different expertise, the specific topics across speakers differ, but the interviews all focus on aspects of Hawaiian culture and history, and on the Hawai‘i in which the speakers were raised.

3.2 Data preparation. The sound files had been previously digitized as mp3 files, and these were the sound files that were made available to us. Also made available to us were pdfs of typed transcriptions and translations of the interviews done by Clinton Kanahele.

To prepare the data for analysis, the mp3 files were converted to wave files and the transcripts were used as a guide to create new, time-aligned and searchable transcripts in ELAN. Once transcription was complete, the sound files and transcripts were uploaded to SOLIS, a multi-language corpus created by the first author that has been optimized for quantitative work in linguistics. The corpus was then mined for instances of *kēia* using LaBB-CAT (Fromont & Hay 2012). Despite the age of the recordings and the lossy format in which the files were originally digitized, the spectrograms and waveforms are clear, with clear formants and timing, as shown in Figure 1.

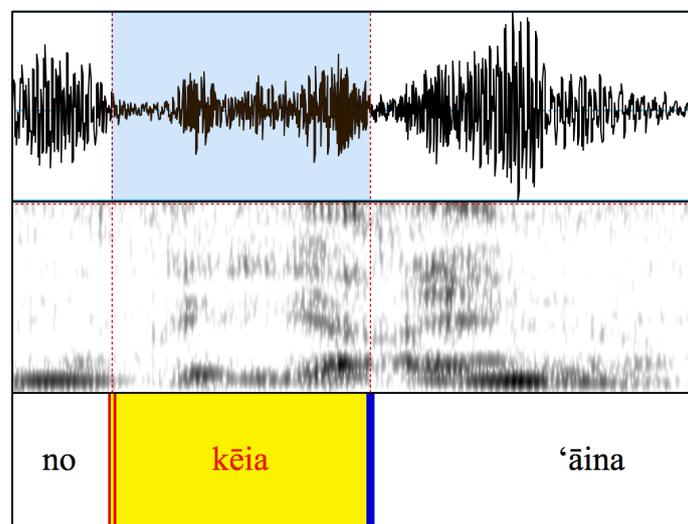


FIGURE 1: Example waveform and spectrogram from a male speaker

3.3 Analysis. Auditory analysis of the tokens was conducted by the first author. The full range of realizations observed in the data are shown in Appendix A. Using Praat, tokens were played over Bose noise-canceling headphones, and narrow phonetic transcription was conducted. The waveform and spectrogram of tokens identified as being produced with a constricted glottis¹⁶ were then inspected in order to investigate

¹⁶ Restricted glottis refers to a range of articulations at the glottal place of articulation, including creak and glottal stop.

the acoustic properties of the sound. In the data, there are very few tokens of *kēia* that contain a true glottal stop, where there is complete occlusion of the vocal folds ($n=4$). While not the focus of this paper, it is worthwhile to note that a phonetic glottal stop is also uncommon in words where the canonical form contains the glottal stop phonemically. The phoneme¹⁷ is much more commonly realized as creak, and the few tokens that contain a stop also contain creak. This is not particularly surprising given that glottal stops in other languages can be produced with a range of realizations (Nolan 1995; Docherty & Foulkes 1999), and it is consistent with other work arguing that the Hawaiian glottal stop can be realized as creak word-medially (Parker Jones, 2017). The tokens where the glottal stop is realized as creak often also have a drop in both f_0 and intensity for the duration of the sound; a drop in f_0 alone has been shown to result in the percept of glottalization (Pierrehumbert & Frisch 1997). Additional work is needed to explore the range of realizations further and to determine which factors influence the phonetic variation in the realizations of this sound. The analysis presented below treats the realization as binary (constricted glottis or not).

To determine what potential factors might influence realizations of *kēia*, the analysis was conducted on all tokens found in the middle of an intonational phrase ($n=391$), referred to henceforth as IP-medial. To analyze reduction, we binned tokens into one of four categories depending on the number of phones realized, including the glottal stop. Heavy syllable nuclei (diphthongs and long vowels) were counted as two phones. Thus, [keʔia] and [ke:ia] were coded as “1” (no reduction), [keʔi], [keʔe], and [keia] were coded as “2”, [ke:] and [kea] were coded as “3”, and [ke]¹⁸ and [ki] were coded as “4” (fully reduced). There were between 80 and 120 tokens in each bin, and all speakers produce realizations that fall into at least three of the four categories, with six speakers producing realizations in all four categories. The examples provided above for each category represent the most frequent realizations in each of the categories, though other realizations (e.g., [te:ia]) were also observed and were categorized using the same criteria.

4. RESULTS. Two separate logistic mixed effects models were fit to the data by hand using R, the first investigating factors influencing the presence of a constricted glottis and the second examining effects on phonetic reduction of the word. Only factors reaching significance of $p<.05$ or smaller were included in the models.

4.1 Factors influencing the presence of a constricted glottis. For the analysis presented here, a constricted glottis is treated as a binary dependent variable (present or absent). Due to the relatively small sample size ($n=391$), several other analyses were conducted: an analysis of the subset of the dataset that included only the first time a word sequence was produced ($n=142$) and an analysis of the subset that included only repeated word sequences. This was done to confirm the reported trends, ensuring that a single word sequence did not bias the results and also removing a potential confound of the hypothesized effect of word sequence repetition. The results reported herein were observed in all of the subsets of data tested. As an additional

¹⁷ *Phoneme* refers to the underlying representation of a sound, regardless of how it’s actually produced. For example, the /p/ in the words *pot* and *spot* are pronounced differently but native speakers of English think of it as a single sound, a phoneme.

¹⁸ Tokens that were reduced to [ke] are analyzed as reduced forms of *kēia* rather than definite article *ke* because (1) two different transcribers perceived and transcribed the word as *kēia* despite the reduced realization and (2) the realization is found with nouns that take *ka* (an alternate definite article) instead of *ke* in addition to those that take *ke*.

precaution, the extent to which individual speakers converge on the overall trends is also discussed.

Using R Version 3.2.4 for Mac (R Core Team, 2016) and lme4 (Bates et al. 2012), we fit a mixed effects model to the binary variable of constricted glottis presence using backwards selection, and models were compared using anova. Speaker and word were treated as random effects in the models. Tested as fixed effects were a number of factors concerned with the first sound of the word immediately following the token. These included whether or not the initial sound was phonemically a glottal stop (treated as binary), a plosive (i.e., /k/, /p/, /t/, or /ʔ/) or not (treated as binary), and the sonority level of the initial sound (treated as ranked), all of which were hypothesized to influence the realization of a constricted glottis due to either assimilation or dissimilation, in part or full. Also tested were factors involving the first syllable of the following word: whether the first syllable was stressed and whether it was heavy or light. Finally, the two social factors – the speaker’s gender and island of origin – were also tested. In cases when there is colinearity between factors, separate models were fit and the model with the weaker predictor was discarded. The results indicate a significant main effect of whether the speaker is male or female ($\beta = -1.59$, $t(389) = -3.29$, $p=.001$), with male speakers being significantly less likely to produce a constricted glottis. Also significant in the model is an effect of whether the following word starts with a plosive; a constricted glottis was significantly less likely to be present when the word immediately following *kēia* began with a stop consonant ($\beta = -1.38$, $t(389) = -3.87$, $p=.0001$). There was no effect observed for the speaker’s island of origin. A table with the full model is shown in Appendix C.

To examine these factors further, the percentage of tokens realized with a constricted glottis in each of these categories is shown in Table 2. Despite the small token numbers, the effect of the following word beginning with a plosive is robust. We observe a higher percentage of tokens realized with a constricted glottis when the following word does not begin with a plosive for both subsets of the data tested (i.e., the first time the word sequence was uttered in the interaction and all other tokens); both predicting factors reach significance in both models. Additionally, when investigated individually for each speaker, we observe the effect of a following plosive for all except two speakers.

following word starts with:	females		males	
	percent	total N	percent	total N
a plosive (/p, k, ʔ/)	42	69	18	102
other	69	96	35	124

TABLE 2: Percentage of tokens realized as with a constricted glottis for male and female speakers, depending on characteristics of the following word.

The tendency for the female speakers to produce a larger percentage of tokens with a constricted glottis compared to the male speakers is also robust: the rate of a constricted glottis when in the phonological contexts that *inhibits* a constricted glottis is 42% for the female speakers, which is even greater than the rate in the context which *favors* a constricted glottis (35%) for the male speakers.

4.2 Factors influencing phonetic reduction. For the second set of results, a mixed effects model was fit to the subset of the data without a constricted glottis.¹⁹ This was done in order to tease apart the effects of phonetic reduction and a constricted glottis. Thus, subsetting the data in this way means that we can test factors influencing reduction, and we can be sure that the effects are not a byproduct of the variable presence of a constricted glottis. Treated as the dependent variable was a binary distinction between “reduced” and “unreduced”. Tokens were categorized in this way using by-speaker means for reduction. To do this, the mean level of reduction (from 1-4) was calculated for each speaker. Then, each token was coded as being reduced if its realization corresponded to a value greater than the mean for that speaker and as unreduced if it was smaller than the mean. This way, we were able to categorize a token as reduced in a way that depended on the speaker’s overall tendency to reduce rather than use a researcher-imposed threshold of reduction.²⁰ The mean rather than the median was used to calculate these values since there are only four possible values and means did not correspond to any of them, making a binary split (over or below the mean) possible. Tested as predicting factors in the model were: stress of the syllable immediately following the token, the number of morae in the syllable immediately following the token, speaker-based and corpus-based measures of bigram²¹ probability, and word sequence repetition.²² For word sequence repetition, two related but different measures were tested: (1) the number of times *kēia* plus the following word had been previously produced and (2) the number of times *kēia* plus the noun phrase had been previously produced. The critical difference between these two measures stems from the frequent occurrence of the plural marker *mau* immediately following *kēia*. We hypothesized that we would observe a stronger effect from measure (1) if predictability stemmed from the immediately surrounding context, and from measure (2) if the effect was due to predictability as related to the lexicalization of the phrase. No difference was observed between the measures, so the model discussed here includes only the second measure, and the two measures will not be compared further. Future research can test the differences between these values using a larger dataset and a wider range of function words, and can also investigate reduction of the content words in the phrase.

The only factor that reached significance in the model was word sequence repetition. The greater number of times that *kēia* plus the following noun phrase was repeated by a speaker, the more likely the token of *kēia* was to be phonetically reduced ($\beta = 0.14$, $t(233) = 2.12$, $p=.034$). This effect is shown graphically in Figure 2 for tokens in the two categories of reduction. The graph is based on the raw data; it is not generated within the context of the model. Evident in Figure 2 is a greater tendency for reduction to occur as word sequences are repeated.

¹⁹ All trends reported here remain significant in a model fit to the full dataset.

²⁰ See Walker et al. (forthcoming) for a similar analysis of a scaled variable.

²¹ Here ‘bigram’ refers to sequences of two words.

²² While speech rate is also hypothesized to be related to reduction, it was not included in this analysis; ongoing work that analyzes reduction in a larger number of Hawaiian function words will include speech rate as a predictor. Additionally, it is possible that the reason we do not observe an effect of speaker-based bigram probability is due to the small number of speakers and tokens per speaker that were analyzed; future work analyzing a larger number of tokens from a larger number of speakers will revisit this factor.

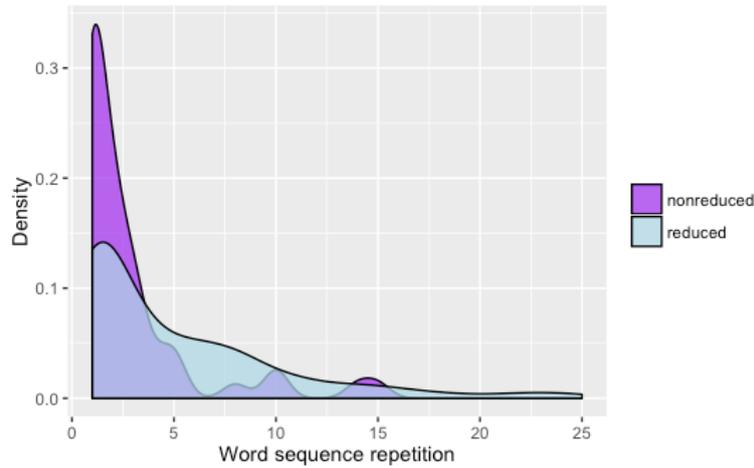


FIGURE 2: Density of tokens across word sequence repetitions, shown for non-reduced tokens (in purple) and reduced tokens (in light blue)

The results suggest that repetition of word sequences can affect phonetic reduction in the function word *kēia* in much the same way that content words undergo reduction with repetition. We return to a discussion of this result in Section 5.

5. DISCUSSION. The results provide evidence that phonetic realizations in Hawaiian are linked with social factors about the speaker (gender) as well as linguistic factors (following environment). Specifically, the results indicate that a constricted glottis is more likely in the word *kēia* when produced by women than by men, and it is less likely to be found if the word immediately following *kēia* begins with a stop. Furthermore, the results provide evidence that *kēia* is more likely to be phonetically reduced as the sequence of words is repeated during the interaction.

The results demonstrate that there is indeed systematic (socio)phonetic variation to be found in Hawaiian. Furthermore, we have identified a great deal of phonetic reduction in spoken Hawaiian. While the presence of phonetic reduction in rapid, conversational speech is well-known to Hawaiian language scholars (Kinney 1956: 285; Elbert & Pukui 1979: 22-23), there seems to be a belief among at least some non-Hawaiian-speaking linguists that Hawaiian exhibits little to no phonetic reduction; we hope the analysis presented herein goes some way toward dispelling that myth.

In the remainder of this section, we discuss the implications of the findings for linguistic theory (Section 5.1) and present our efforts to disseminate the findings to the Hawaiian-speaking community in order to inform the conservation and revitalization of the language (Section 5.2).

5.1 Implications for linguists.

5.1.1 Factors influencing a constricted glottis. The results for constricted glottis demonstrate the importance of treating linguistic variables stochastically rather than labeling non-categorical variation as free variation; social and linguistic factors are systematically linked with the variable, just not categorically so. While this is well-established for widely-studied languages like English, it is often missing from descriptions of lesser-studied endangered languages. Additionally, the patterns

observed for constricted glottis demonstrate the advantages of examining phonetic and phonological processes in connected speech since both the extent of the variation and the main linguistic factor constraining the variation can only be observed in connected speech.

The effect of following stops on glottal stop presence obeys the Obligatory Contour Principle (OCP), in which consonants that share characteristics are less likely to be realized when they are found in close proximity to one another (McCarthy 1986; Yip 1988 and many others). Gradient OCP effects are found in other Austronesian languages, such as Tagalog (Shih & Zuraw, under review) and Māori (Rácz et al. forthcoming). Rácz et al. (forthcoming) report that Māori stops that share their place of articulation are less likely to be realized when found in CVCV sequences in words without reduplicated sequences. A similar phenomenon (OCP for manner of articulation) may be responsible for variation in the presence of a constricted glottis in Hawaiian *kēia*. An extensive analysis of consonant variation in a wide range of words in Hawaiian is needed to explore this further. An additional future direction for research may be to investigate glottal stop variation in Tahitian *teie* ‘this’ and Rarotongan Māori *tēia* ‘this’, where the variable presence of the glottal stop is also found (Walworth, personal communication Dec 2015 for Tahitian; NeSmith, personal communication June 2016 for Rarotongan Māori).

In addition to environmental factors, the speaker’s gender had an effect on whether or not the word was realized with a constricted glottis. Female speakers were more likely to produce *kēia* with a constricted glottis. Further work is needed to determine the extent to which this pattern can be found beyond the word *kēia* and the extent to which the variation is socially meaningful.

5.1.2 Repetition of word sequences affecting phonetic reduction. In addition to the results for constricted glottis, we observed an effect of word sequence repetition on phonetic reduction in *kēia*, lending support to the Probabilistic Reduction Hypothesis (Jurafsky et al. 2001) and the Smooth Signal Reduction Hypothesis (Aylett & Turk 2004). This result is surprising for two reasons. The first is that probabilistic effects such as this have largely remained undescribed (and probably largely untested) in Austronesian languages (though see Davis 2003; Blevins 2005). Even more surprising is the fact that we have observed an effect of repetition on a function word at all; work on English has observed repetition effects for content words (Fowler & Housum 1987; Fowler 1988) but not function words (Bell et al. 2009: 104). One reason we may have observed an effect is that we counted repetitions of word sequences rather than counting the function words in isolation, as others have done. While function words in isolation may be so frequent that repetition has little to no effect at the word level, it is possible that lexicalized phrases undergo reduction in much the same way as content words in isolation do. If this is the case, other words within the phrase may also be realized as more reduced as the phrase is repeated, a hypothesis that will be tested in future research. Future research will also focus on a wider range of function words.

The first set of results (i.e., the variability of the presence of a constricted glottis) is in accordance with models of speech production that can account for systematic variation that is non-categorical, such as variable rules (Labov 1969; Cedergren & Sankoff 1974; Guy 1991), exemplar-based models (Johnson 1997; Pierrehumbert 2001, 2006), and Bayesian models (Norris & McQueen 2008). The second set of results, however, present a challenge to models that do not track or store

repetitions of words and sequences of words. To account for this set of results, an exemplar-based or Bayesian-based account is most promising.

There are, of course, limitations to the work presented herein. We acknowledge that our sample size of eight speakers is smaller than what has become the norm for many variationists. A small sample size is one of the drawbacks of conducting quantitative work on an endangered language. But while it may limit the types of variables one can investigate, it is still a worthwhile enterprise, as we have shown. Given the robustness of our results – particularly the results for phonological environment and speaker gender – we anticipate that our ongoing research in this area, incorporating a larger number of tokens, words, and predicting variables as well as drawing on data from a larger number of speakers, will confirm the findings presented here.

5.2 Dissemination to speakers. There are a number of ways that sociolinguistic analyses of variation in endangered languages can assist language learners and instructors. However, it is unlikely to have much of an impact if the report exists solely in an academic book or journal because only a handful of learners with no linguistics training will choose to wade through the technical terms and abstract theoretical concepts to learn about the variable they are interested in. Therefore, we have developed an online resource to aid in the teaching and acquisition of linguistic variation. The goal of the website is to make information about variation in Hawaiian more accessible for learners and instructors alike. In this section we begin by discussing select online resources for learners of Hawaiian. We then present a description of the website that we developed to assist in the teaching and acquisition of sociolinguistic variation in Hawaiian. Finally, we provide a brief discussion about our personal connection to the project, describing what we feel is our responsibility as linguists and Hawaiian language learners.

The internet provides a means of reaching a wide variety of learners, and it can also provide a united platform for a speech community that is dispersed geographically. Online tools provide a means for all learners – not only those who are associated with an institution or who gain special permission – to have access to recordings and tools that can assist in their progression toward fluency. Therefore, websites are an excellent medium for Hawaiian language learners, most of whom can gain access through a public library if not from the comfort of their home. Learners of Hawaiian, especially those who have already achieved some level of proficiency, tend to be highly motivated to improve their language skills. Thus, it is not particularly surprising that online resources made available for such learners are heavily used. For example, in 2011, ‘Ōiwi TV (found online at www.oiwi.tv and on Oceanic Digital Channel 326) reported as many as 225,000 views per month to their television channel and website after only two years of existence (Kamehameha Schools 2011). At a smaller scale, the Kealopiko²³ blog, with weekly posts on the various meanings and uses of select lexical items, receives hundreds of visits to each of its posts, with hundreds of people also engaging in other online formats (e.g., Instagram). Thus the internet is proving to be a viable way of both disseminating information about and increasing exposure to the Hawaiian language.

There are two main modes of acquiring sociolinguistic variables: explicit instruction and ample exposure to the pattern. We first discuss the benefits of exposure

²³ Kealopiko is a clothing company geared toward raising awareness and prestige of Hawaiian language, culture, and biodiversity.

in regard to acquiring variation. Contact with native speakers results in increased exposure to patterns of sociolinguistic variables, which can, in turn, aid both comprehension and production. However, when there is little opportunity to interact with native speakers, other methods of exposure become necessary. Luckily, there are several existing archives of conversational speech produced by native speakers of Hawaiian who learned the language through an unbroken chain of intergenerational transmission.²⁴ These recordings are an invaluable resource for Hawaiian language learners and scholars. Thanks to the initiative of a small number of forward-thinking individuals, including the kūpuna who agreed to be interviewed, there are several archives that contain such speech. One of these is the Kanahēle Collection, from which the data analyzed for this project were gleaned. The Kanahēle Collection is accessible online, with mp3 files and corresponding pdfs of transcriptions and translations available for download. Another archive is made up of two sets of interviews from *Ka Leo Hawai'i*, a radio show begun by Larry Kimura. The first set was recorded between 1972-1989 (Kimura 2015) and, through Kimura's efforts, is in the process of being made available. Under the direction of Puakea Nogelmeier, the second set of *Ka Leo Hawai'i* recordings was made between 1991-2000. These recordings are available online for download as mp3 files (<https://evols.library.manoa.hawaii.edu/handle/10524/47857>). The largest archive of Hawaiian language recordings is held at the Bishop Museum and was collected in the 1950s and 1960s by Mary Kawena Pukui. Pukui, a native speaker of Hawaiian, interviewed people from six different islands. While of course not as beneficial as interacting with native speakers, listening to these collections is an excellent way for learners who have little contact with native speakers to be exposed to the linguistic patterns that are found in the recordings in addition to the worldviews expressed therein (Kneubuhl 2014: 72). Note that ownership of the recordings ultimately rests with those who produced and maintain them. Thus, while the recordings are freely available to the public and learners are encouraged to listen to the recordings, researchers who were not directly involved with data collection should seek permission from the owners of the archives before treating them as data or disseminating them through other venues.

The increased accessibility of recordings discussed above can help learners achieve a higher level of fluency in Hawaiian, including the incorporation of non-categorical linguistic variation and an increased understanding of Hawaiian concepts and culture. Explicit instruction, however, can assist with the comprehension and acquisition of the variables, especially if received alongside regular exposure to the variable in speech through, for example, frequently listening to the archived recordings. Explicit instruction draws learners' attention to the pattern so it can be noticed and then learned (Schmidt 1990). Instructors can incorporate sociolinguistic variants when using a Communicative Language Teaching approach (Richards 2006), treating L1 speech as a model and giving students practice using the variants. Additionally, explicit instruction that includes exposure to short clips of speech from the archives may help lessen the degree to which the recordings are perceived as daunting to some beginning learners. As an instructor of Hawaiian language at the University of Hawai'i at Mānoa, the third author found that, while some learners readily embrace listening to the recordings even if they don't understand most of what is being said, many others find the task intimidating. By sharing smaller clips along with the associ-

²⁴ To our knowledge, there is not yet an archive of speech produced by L1-speaking children of L2-speaking parents though this would be an excellent resource.

ated transcriptions on the website, we hope to make the task of listening a little less intimidating.

With an aim of reaching a wide range of learners (i.e., not only those who are taking a Hawaiian language class), we have created a website (<http://nuinaala.wixsite.com/variationinhawaiian>) that explains the variation that we report in this paper, with plans to add new pages describing other variables as descriptions emerge. The linguistic variable and select factors constraining the variable are presented in language intended to engage learners who have little to no training in linguistics. The site makes available short clips of speech that contain different phonetic realizations of *kēia*, alongside transcription of the clips. Additionally, examples of lesson plans and in-class exercises are provided to assist any instructors who wish to incorporate a discussion of linguistic variation in the classroom but do not already have the resources to do so. The goals of the website are to pique users' interest in variation, teach them something about what we found and, for beginning learners especially, make the native speaker resources less intimidating through presenting short clips alongside the transcription. The website draws the viewer's attention to the variation without prescribing a "right" way to use the variation. This is especially important given that speaker identity is complex. There are likely complex interactions between predicting variables and subtle ways the variation is used for social purposes that is completely missing from our description. Also, the website is not intended as a standalone site for acquiring the language; it is assumed that learners are acquiring Hawaiian through other methods and that they will use the site merely as a supplementary tool to learn something about the range of different pronunciations found in the speech of native speakers.

This aspect of our work – the creation of resources to assist language learners – immediately becomes part of our *kuleana* when examining variation in the language at all. The Hawaiian concept of *kuleana* refers to one's right, responsibility, and province, which, for linguists, includes helping to responsibly steward aspects of language we study to speakers of the language. Assisting learners in their acquisition of Hawaiian has been central to the motivations behind this project. We believe that an examination of sociolinguistic variation in Hawaiian helps to maintain a link between the language as it is used today and how the ancestors spoke, and that the information gleaned from sociolinguistic analyses is an important step in the maintenance of that link. Thus, we view the creation of this website as a critical component of our research investigating variation in Hawaiian. In fact, for us, dissemination of the results to interested community members is just as important as disseminating the information among linguists. Whether learners choose to use the site and whether they then include variation in their realizations of *kēia* remains to be seen. It is not our *kuleana* to say they should.

6. FINAL THOUGHTS. Taken together, the results reveal a great deal of phonetic variation, even though we focus our analysis on only one word: *kēia*. The results provide a description of sociolinguistic variation in Hawaiian, and they inform current models of linguistic theory. Additionally, we are using the results to construct resources for language learners who wish to adopt the patterns of variation we have described.

The work demonstrates the advantages of using a variationist approach to analyze data from a lesser-studied language; we believe using variationist methodologies on a wider range of languages will yield an unforeseen number of insights. We en-

courage variationists to describe variation in a wider range of languages and we urge language documentarians to view non-categorical linguistic variation as an integral part of language. We recognize, of course, that limited resources and the overwhelming number of endangered languages in the world makes our recommendation difficult to fulfill. Additionally, we face something of a conundrum: variationist analysis on endangered or threatened languages is especially feasible for languages that already have some documentation so that researchers can focus their attention on those variables that are most likely to vary in interesting ways. However, populations who speak endangered languages are aging and, as language shift occurs, sociolinguistic variation is lost. It is desirable (if not essential) that appropriate recordings be made prior to the language reaching the status of a severely endangered language. Word lists and other formal methods of eliciting data, while useful, normally elicit only the most formal styles of speech, meaning that they do not provide the wide range of styles that speakers use. In variationist work, an analysis of such data is often compared with an analysis of conversations between speakers in order to more fully describe the range of variation. Ideally, conversations between speakers who grew up speaking the language are recorded. Then, transcription and translation of the materials – with help from highly fluent speakers – is conducted. Of course, in an endangered language context, such a scenario is not always possible. A variationist analysis can still be conducted (see e.g., Babel 2008), but – just as in any linguistic study – the range, richness, and robustness of the variation decreases as language loss increases. Thus, in order to counteract the conundrum, we recommend recording interactions between speakers of endangered languages early in the documentation process, even if one must wait to analyze the sociolinguistic variation until after documentation of more categorical trends has been completed. In order to accomplish this goal, we see the most fruitful way forward being one where – when specialization in both subfields by a single individual is not (immediately) possible – close collaborations between variationist sociolinguists and language documentarians are established. Likewise, graduate students can benefit from insights of a diverse faculty, gaining the expertise they need to conduct a sociolinguistic variationist analysis as part of their language documentation project.

For Hawaiian, there is still much work to be done, both in terms of describing variation and disseminating information about the variation to language learners. Therefore, this paper marks what is merely the beginning of a research program aimed at describing variation found in Hawaiian, conducted with the joint purpose of advancing linguistic theory and informing ongoing Hawaiian language revitalization efforts.

APPENDICES

Appendix A

Realizations of *kēia* observed in the full dataset (i.e., all phrasal positions), shown in decreasing order of occurrence. Number of tokens in the restricted dataset that was used to analyze the constricted glottis variation shown in parentheses. Tokens with no number correspond to attested realizations that were not included in the analysis presented herein due to noise, occurring at an intonational phrase boundary, or being followed by a pause.

ke (80)	ke‘i (3)
ke: (77)	keo (3)
ke‘i (56)	ke:i (2)
ke‘ia (32)	keu (2)
ke‘e (30)	ēia (1)
ke:ia (22)	g ^h e (1)
ke:‘ia (15)	gu (1)
kea (14)	ka (1)
ka‘i (10)	ke:a (1)
kei (7)	ke‘a (1)
keia (7)	ke‘i: (1)
ke‘ea (6)	kuja (1)
ka‘ia (4)	te (1)
ki (4)	ke:‘i
kia (4)	te:ia
kai (3)	

Appendix B

Model of constricted glottis

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.9898	0.3886	2.547	0.0109
gender = M	-1.5904	0.4833	-3.291	0.0010
following = plosive	-1.3787	0.3562	-3.870	0.0001

Model of reduction

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.8785	0.6014	3.124	0.0018
word sequence repetition	0.1447	0.0683	2.119	0.0341

Appendix C

By-speaker means and medians for reduction.

speaker	mean	median
Hilda	3.34	3
James	3.64	4
JohnB	4.26	5
JohnC	4.23	4
Luka	3.21	3
Lydia	3.33	3
Minnie	2.89	3
Peter	3.42	4

REFERENCES

- Act 57, sec 30 of 1896 Laws of the Republic of Hawai‘i passed by the Legislature – 1896, p. 189, Sanford B. Dole.
- Anderson, Roger W. 1982. Determining the linguistic attributes of language attrition. In Richard D. Lambert & Barbara F. Freed (eds.) *The Loss of Language Skills*, Rowley, MA: Newbury House, 83-118.
- Aylett, Matthew & Alice Turk. 2004. The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47(1). 31-56. doi: 10.1177/00238309040470010201
- Babel, Molly. 2008. The phonetic and phonological effects of moribundity. *University of Pennsylvania Working Papers in Linguistics* (142). 25-34.
- Baker, Rachel E. & Ann R. Bradlow. 2009. Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech* (524). 391-413.
- Bauer, Matt. 2005. Lenition of the flap in American English. *Penn Working Papers in Linguistics*, 102.
- Bell, Alan. 1984. Language style as audience design. *Language in Society* 13(2). 145-204.
- Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand & Dan Jurafsky. 2009. Predictability effects on duration of content and function words in conversational English, *Journal of Memory and Language* 60. 92-111.
- Blevins, Juliette. 2005. The role of phonological predictability in sound change: Privileged reduction in Oceanic reduplicated substrings, *Oceanic Linguistics* 44(2). 517-526.
- Bradlow, Ann R., Gina M. Torretta, & David B. Pisoni. 1996. Intelligibility of normal speech 1: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20(3-4). 255–272.
- Bucholtz, Mary. 2011. *White Kids: Language, race, and style of youth identity*. Cambridge University Press.
- Bush, Nathan. 2001). Frequency effects and word-boundary palatalization in English, in Joan Bybee & Paul Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*, Benjamins, Amsterdam. 255–280.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Byrd, Dani. 1994. Relations of sex and dialect to reduction, *Speech Communication*, 15(1-2). 39-54.
- Campbell, Lyle & Martha C. Muntzel. 1989. The structural consequences of language death, In Nancy C. Dorian (ed.) *Investigating Obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press. 181-196.
- Cedergren, Henrietta J. 1973. *The Interplay of Social and Linguistic Factors in Panama*. Cornell University Ph.D. dissertation.
- Cedergren, Henrietta J. & David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2). 333-355.
- Cheshire, Jenny, Paul Kerswill, Susan Fox & Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15(2). 151-196.

- Coffman, Tom. 2009. *Nation Within: The history of the American occupation of Hawai'i*. Kihei: Koa Books. Revised Edition.
- Cutler, Anne. 1991. Linguistic rhythm and speech segmentation. In J Sundberg, L. Nord & R. Carlson (eds.) *Music, language, speech and brain*. London: Macmillan, 157-166.
- Davis, Karen. 2003. *A Grammar of the Hoava Language, Western Solomons*. Canberra : Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Díaz-Campos, Manuel & Michael Gradoville. 2011. An analysis of frequency as a factor contributing to the diffusion of variable phenomena: Evidence from Spanish data. In Luis A. Ortiz-López (ed.) *Selected Proceedings of the 13th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project. 224-238.
- Eckert, Penelope. 1980. The structure of a long-term phonological process: The back vowel chain shift in Soulatan Gascon. In William Labov (ed.) *Locating Language in Time and Space*. 179-219.
- Eckert, Penelope. 2000. *Language Variation as Social practice: The Linguistic Construction of Identity in Belten High*. Wiley-Blackwell.
- Elbert, Samuel H. & Mary Kawena Pukui. 1979. *Hawaiian Grammar*. Honolulu: University of Hawai'i Press.
- Fosler-Lussier, Eric & Nelson Morgan. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communication* 29. 137-158.
- Fowler, Carol. 1988. Differential shortening of repeated content words produced in various communicative contexts, *Language and Speech* 31(4). 307-319.
- Fowler, Carol & Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26(5). 489-504
- Fromont, Robert & Jennifer Hay. 2012. LaBB-CAT: An annotation store. In Paul Cook & Scott Nowson (eds.) *Proceedings of the Australasian Language Technology Association Workshop*. Australasian Language Technology Association 113-117.
- Gahl, Susanne. 2008. 'Time' and 'thyme' are not homophones: The effect of lemma frequency on word durations in a corpus of spontaneous speech. *Language* 84(3). 474-496.
- Gahl, Susanne & Susan M. Garnsey. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80. 748-775.
- Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, & Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. In Sabrina J. Billings, John P. Boyle & Aaron M. Griffith (eds.) *Chicago Linguistics Society: The Panels* 35(2).151-166.
- Guy, Gregory R. 1991. Contextual conditioning in variable lexical phonology, *Language Variation and Change* 3. 223-239.
- Hardeman Guthrie, Kate. 2016. Gender and second language style: American learner perceptions of Mandarin *sajiao*. *Asia-Pacific Language Variation* 2(2). 157-187.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25-70.

- Hay, Jennifer & Paul Foulkes. 2016. The evolution of medial /t/ over real and remembered time. *Language* 92(2). 298-330.
- Heffernan, Kevin. 2010. Mumbling is macho: Phonetic distinctiveness in the speech of American radio DJs. *American Speech*, 85(1). 67–90.
- Hopkins, Alberta P. 1992. *Ka lei ha'aeo: Beginning Hawaiian*. Honolulu: University of Hawai'i Press.
- Hunnicut, Sharon. 1985. Intelligibility versus redundancy – conditions of dependency. *Language and Speech* 28(1). 47–56.
- Hymes, Dell. 1972. On communicative competence. In John B. Pride and Janet Holmes (eds.) *Sociolinguistics: Selected readings*, Harmondsworth: Penguin. 269-293.
- Jurafsky, Daniel, Alan Bell & Cynthia Girand. 2002. The role of the lemma in form variation, in Carlos Gussenhoven & Natasha Warner (eds.) *Papers in Laboratory Phonology VII*. New York: Mouton de Gruyter. 1-34.
- Kamehameha Schools Community Relations and Communications Group. 2011, Summer. A station for the nation. *Imua*, 2011(2). 1, 14-15.
- Kawai'ae'a, Keiki K.C., Alohalani Kaluhiokalani (Kaina) Housman, Makalapua (Ka'awa) Alencastre. 2007. Pū'ā i ka 'ōlelo, ola ka 'ohana: Three generations of Hawaiian language revitalization. *Hūlili: Multidisciplinary research on Hawaiian Well-being* 4(1). 183-237.
- Kerswill, Paul. 2003. Dialect levelling and eographical diffusion in British English. In Dave Britain and Jenny Cheshire (eds.) *Social Dialectology. In honor of Peter Trudgill*. Amsterdam: Benjamins. 223-243.
- Kimura, Larry L. 1983. Native Hawaiian culture. *Native Hawaiians study commission: Report on the culture, needs, and concerns of native Hawaiians* 1. 173-203, Washington, D.C.: U.S. Department of the Interior.
- Kimura, Larry. 2015. Ka Leo Hawai 'i tapes and the genesis of Hawaiian. Paper presented at the 4th International Conference on Language Documentation and Conservation (ICLDC). Talk available for download at: <http://hdl.handle.net/10125/25329>
- Kinney, Ruby Kawena. 1956. A non-purist view of morphomorphemic variations in Hawaiian speech. *The Journal of the Polynesian Society* 65(3). 282-286.
- Kirtley, M. Joelle. 2015. *Identity, Language and Non-binary Gender in Hawai'i*. University of Hawai'i at Mānoa PhD dissertation.
- Kneubuhl, Hina Puamohala. 2014. *Ke mau ne nō nā leo kūpuna: He papahana noi 'i no ke ki'a hopuna 'ōlelo e like me kona ho'opuka 'ia e nā kūpuna*. University of Hawai'i at Mānoa MA thesis.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington D.C.: Center for Applied Linguistics.
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715-762.
- Labov, William. 1972. *Sociolinguistic patterns*. University of Pennsylvania Press.
- Lieberman, Philip. 1963. Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech* 6(3). 172-187.
- Lucas, Paul F. N. 2002. E ola mau kakou i ka 'olelo makuahine: Hawaiian language policy and the courts. *Hawaiian Journal of History*, 34. 1-28.
- Lyster, Roy. 1994. The effect of functional-analytical teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics* 15. 263-287.

- Maclagan, Margaret, Catherine I. Watson, Ray Harlow, Jeannette King & Peter Keegan. 2009. /u/ fronting and /t/ aspiration in Māori and New Zealand English. *Language Variation and Change*, 21(2). 175-192.
- Milroy, Leslie. 1980. *Language and Social Networks*. Oxford: Basil Blackwell.
- Mougeon, Raymond, Katherine Rehner & Terry Nadasdi. 2004. The learning of spoken French variation by immersion students from Toronto, Canada. *Journal of Sociolinguistics* 8(3). 408-432.
- Nagy, Naomi. 2011. Lexical change and language contact: Faetar in Italy and Canada, *Journal of Sociolinguistics* 15(3). 366-382.
- NeSmith, R. Keao. 2005. Tūtū's Hawaiian and the emergence of a Neo Hawaiian language, *Ōiwi: A Native Hawaiian Journal* 3. 135-152.
- NeSmith, R. Keao. 2012. *The Teaching and Learning of Hawaiian in Mainstream Educational Contexts in Hawai'i: Time for Change?* University of Waikato Ph.D. dissertation.
- NeSmith, R. Keao. Forthcoming. Let native speakers be native speakers and let Hawaiian language learners be like them too. To appear in *Hūlili: Multidisciplinary research on Hawaiian Well-being*.
- Newbrand, Helene Luise. 1951. *A Phonemic Analysis of Hawaiian*. University of Hawai'i at Mānoa M.A. thesis.
- Nolan, Francis. 1995. The role of the jaw – active or passive? In Bruce Connell and Amalia Arvaniti (eds.) *Phonology and Phonetic Evidence: Papers in Laboratory Phonology*. Cambridge University Press. 343-360.
- Oliveira, Katrina-Ann R. Kapā'anaokalākeola Nākoa. 2014. E Ola Mau ka 'Ōlelo Hawai'i: The Hawaiian language revitalization movement. In Noelani Good-year-Ka'opua, Ikaika Hussey & Erin Kahuawaika'ala Wright (eds.) *A Nation Rising: Hawaiian movements for life, land, and sovereignty*. Duke University Press, 78-85.
- Parker Jones, 'ŌiwiHawai'iokalani. 2006. Durational variability and stress-timing in Hawaiian. In Paul Warren & Catherine I. Watson (eds.) *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*. Auckland: Australasian Speech Science and Technology Association, 417-420.
- Parker Jones, 'ŌiwiHawai'iokalani. 2017. Hawaiian. *Journal of the International Phonetic Association*, first view 1-13. DOI: <https://doi-org.eres.library.manoa.hawaii.edu/10.1017/S0025100316000438>
- Piccolo, Fabiana. 2005. Where is the Hawaiian language headed? A phonetic study. *University of Hawai'i at Mānoa Working Papers in Linguistics* 36(1). 1-24.
- Pierrehumbert, Janet B. & Stefan Frisch. 1997. Synthesizing allophonic glottalization. In Jan P.H. van Santen, Joseph P. Olive, Richard W. Sproat & Julia Hirschberg (eds.) *Progress in Speech Synthesis*. New York:Springer. 9-26.
- Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118(4). 2561-2569.
- Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11(4). 478-504.
- Pukui, Mary Kawena & Samuel H. Elbert. 1986. *Hawaiian Dictionary*. Honolulu: University of Hawai'i Press.
- Pukui, Mary Kawena, Haertig, E.W. & Catherine A. Lee. 2002. *Nānā i ke kumu, Volume II*. Honolulu: Hui Hānai.
- R Core Team. 2016. R: A language and environment for statistical computing. Vien-

- na: R Foundation for Statistical Computing. Downloaded from <https://www.R-project.org>
- Rácz, Peter, Jennifer Hay, Jeremy Needle, Jeanette King & Janet B. Pierrehumbert. Forthcoming. Gradient Māori phonotactics. To appear in *Te Reo*.
- Rankin, Robert L. 1978. The unmarking of Quapaw phonology: A study of language death. *Kansas Working Papers in Linguistics* 3. University of Kansas Department of Linguistics, 45-52.
- Raynor, Keith & Susan A. Duffy. 1986. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition* 14(3). 191-201.
- Rehner, Katherine & Raymond Mougeon. 1999. Variation in the spoken French of immersion students: To *ne* or not to *ne*, that is the sociolinguistic question. *The Canadian Modern Language Review* 56(1). 124-154.
- Richards, Jack C. 2006. *Communicative Language Teaching Today*. Cambridge University Press.
- Sai, David Keanu. 2011. *Ua mau ke ea, Sovereignty endures: An overview of the political and legal history of the Hawaiian Islands*. Honolulu: Pū‘ā Foundation.
- Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11(2). 129-158.
- Schütz, Albert J. 1994. *The Voices of Eden: A History of Hawaiian Language Studies*. Honolulu: University of Hawai‘i Press.
- Shih, Stephanie S. & Kie Zuraw. Under review. Phonological conditions on variable adjective-noun word order in Tagalog.
- Stanford, James N. 2009. Clan as a sociolinguistic variable: Three approaches to Sui clans. In James N. Stanford & Dennis R. Preston (eds.) *Variation in Indigenous Minority Languages*. Amsterdam/Philadelphia: John Benjamins. 263-484.
- Tarone, Elaine & Merrill Swain. 1995. A sociolinguistic perspective on second language use in immersion classrooms. *The Modern Language Journal* 79. 166-178.
- Tava, Rerioterai & Moses K. Keala. 1989. *Niihau: The Traditions of an Hawaiian Island*. Honolulu: Mutual.
- Tily, Harry, Susanne Gahl, Inbal Arnon, Anubha Kothari, Neal Snider & Joan Bresnan. 2009. Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition*, 1. 147-165.
- Van Son, R.J.J.H., Olga Bolotova, Mietta Lennes & Louis C.W. Pols. 2004. Frequency effects on vowel reduction in three typologically different languages (Dutch, Finnish, Russian), *Proceedings of Interspeech 2004*. 1277-1280.
- Walker, Abby, Jennifer Hay, Katie Drager & Kauyumari Sanchez. Forthcoming. Divergence in speech perception. To appear in *Linguistics*.
- Walworth, Mary. 2015. *The Language of Rapa Iti: Description of a Language in Change*. University of Hawai‘i at Mānoa PhD dissertation.
- Warner, Sam L. No‘eau. 1996. *I Ola ka ‘Ōlelo i nā Keiki: Ka ‘apo ‘ia ‘ana o ka ‘ōlelo Hawai‘i e nā keiki ma ke kula kaiapuni*. University of Hawai‘i at Mānoa. Ph.D. dissertation.
- Warner, Sam L. No‘eau. 1999. Kuleana: The right, responsibility, and authority of

- indigenous peoples to speak and make decisions for themselves in language and cultural revitalization. *Anthropology and Education Quarterly* 30(1). 68-93.
- Warner, Sam L. No‘eau. 2008. The movement to revitalize Hawaiian language and culture. In Leanne Hinton and Ken Hale (eds.) *The Green Book of Language Revitalization in Practice* (3rd edition; 1st edition published in 1982. Bingley: Emerald, 133-144.
- Warschauer, Mark, Keola Donaghy & Hale Kuamo‘o. 1997. Leoki: A powerful voice of Hawaiian language. *Computer Assisted Language Learning*, 10(4). 349- 61.
- Wiener, Seth, Shari R. Speer & Claire Shank. 2012. Effects of frequency, repetition and prosodic location on ambiguous Mandarin word production. In *Proceedings of the 6th International Conference on Speech Prosody*. 528-531.
- Wilson, William H. 1998. I ka ‘ōlelo Hawai‘i ke ola, Life is found in the Hawaiian language. *International Journal of the Sociology of Language* 132. 123-137.
- Wilson, William H. & Kauanoē Kamanā. 2008. “Mai loko mai o ka ‘i‘ini: Proceeding from a dream” The ‘Aha Pūnana Leo Connection in Hawaiian Language Revitalization. In Leanne Hinton and Ken Hale (eds.) *The Green Book of Language Revitalization in Practice* (3rd edition; 1st edition published in 1982. Bingley: Emerald. 146-176.
- Wong, Annette Kuuipolani. 2010. *Pukaiki kula maniania no Niihau, na ka Niihau*. University of Hawai‘i at Mānoa MA thesis.
- Wong, K. Laiana. 1999. Authenticity and the revitalization of Hawaiian. *Anthropology & Education Quarterly* 30(1). 94-115.
- Wong, K. Laiana. 2011. Hawaiian methodologies of indirection: Point-less vs. point-less. In Kathryn A. Davis (ed.) *Critical Qualitative Research in Second Language Studies: Agency and advocacy*, Charlotte: Information Age. 151-170.
- Zhao, Yuan & Daniel Jurafsky. 2007. The effect of lexical frequency on tone production. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the International Congress of Phonetic Sciences XVI*, International Phonetic Association. 477-480.
- Zipf, George K. 1929. Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology* 15. 1-95.

Corresponding author: Katie Drager
kdrager@hawaii.edu