

Exploring the connections between Pidgin and ‘ōlelo Hawai‘i

by

Michelle Kamigaki-Baron

B.A., The University of Southern California, 2014

M.A., The University of Hawai‘i at Mānoa, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Linguistics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2026

© Michelle Kamigaki-Baron 2026

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Exploring the connections between Pidgin and ‘ōlelo Hawai‘i

submitted by Michelle Kamigaki-Baron in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Linguistics

Examining Committee:

Molly Babel, Professor, Linguistics, UBC

Supervisor

Carla Hudson-Kam, Professor, Linguistics, UBC

Supervisory Committee Member

Scott Saft, Professor, Linguistics, University of Hawai‘i at Hilo

Supervisory Committee Member

Henry Davis, Professor, Linguistics, UBC

University Examiner

Alexis Black, Assistant Professor, Audiology and Speech Sciences, UBC

University Examiner

Jeanette King, Professor, Aotahi School of Maori and Indigenous Studies,

University of Canterbury

External Examiner

Abstract

This dissertation examines the potential psycholinguistic connections between Pidgin (Hawai‘i Creole) and ‘ōlelo Hawai‘i (Hawaiian), in contrast to English, addressing whether Pidgin offers an advantage in processing ‘ōlelo Hawai‘i. It considers community-driven questions about whether links between Pidgin and ‘ōlelo Hawai‘i support its use as a medium of instruction and whether Pidgin better facilitates ambient learning of ‘ōlelo Hawai‘i compared to English. The thesis begins with reflections on conducting this research in community contexts by use of mo‘olelo (storytelling).

Experiments 1 and 2 examine trilingual speakers of ‘ōlelo Hawai‘i, Pidgin, and English, assessing co-associations between Pidgin-‘ōlelo Hawai‘i and English-‘ōlelo Hawai‘i pairs. Prior work suggests that socially and structurally related codes can facilitate cross-language processing (e.g., Fricke et al. [2016]; Szakay et al. [2016]). Experiment 1 tests whether language switching context modulates lexical activation in ‘ōlelo Hawai‘i. Results suggest lower processing costs for Pidgin-to-‘ōlelo Hawai‘i switches. Experiment 2 uses auditory translation priming to test asymmetric cross-language associations; no overall priming effects were found.

Experiment 3 examines the proto-lexical knowledge of ‘ōlelo Hawai‘i across three non-‘ōlelo Hawai‘i speaking groups compared to ‘ōlelo Hawai‘i speakers: Pidgin speakers, English speakers from Hawai‘i, and English speakers from Canada. Prior work shows that ambient exposure to Māori supports the development of proto-lexical knowledge (e.g., Johnson [2016]) in adults without formal instruction (e.g., Oh et al. [2020], Panther et al. [2023b]). Pidgin speakers show more developed knowledge of phonotactics and word form than English speakers from Hawai‘i, and both groups outperform English speakers from Canada.

Overall, evidence for Pidgin-related advantages in the scope of ‘ōlelo Hawai‘i processing are mixed: results suggest that switching costs may be reduced, but priming effects are absent. However, when examining non-‘ōlelo Hawai‘i speakers’ relative knowledge of ‘ōlelo Hawai‘i, Pidgin speakers show evidence of stronger proto-lexical knowledge than English speakers from Hawai‘i, suggesting that gradient differences in exposure and/or structural

Abstract

overlap between Pidgin and ‘ōlelo Hawai‘i contribute to processing differences. Throughout this work, we conclude that Pidgin incurs no measurable disadvantage in any condition and yields measurable benefits in ‘ōlelo Hawai‘i proto-lexical knowledge relative to English speakers.

Lay Summary

This dissertation examines how people mentally connect Pidgin (the first language of most residents in Hawai‘i) and ‘ōlelo Hawai‘i (the Indigenous language of Hawai‘i), and whether Pidgin speakers have more ‘ōlelo Hawai‘i knowledge than English speakers. I start by describing my experiences understanding community perspectives by use of mo‘olelo (story-telling). Next, I tested whether trilinguals mentally link Pidgin and ‘ōlelo Hawai‘i more than English and ‘ōlelo Hawai‘i in Experiments 1 and 2. Then, I investigated how people who do not speak ‘ōlelo Hawai‘i differ in their ‘ōlelo Hawai‘i knowledge in Experiment 3. Overall, I found that while Pidgin and English may provide similar connections to ‘ōlelo Hawai‘i among trilinguals, Pidgin speakers show greater familiarity with ‘ōlelo Hawai‘i sound patterns compared to English speakers, which may provide them an advantage in future ‘ōlelo Hawai‘i learning.

Preface

This dissertation is an original intellectual product of the author Michelle Kamigaki-Baron. All written material was produced independently by the author. With the support of the author’s supervisor, Dr. Molly Babel, Michelle Kamigaki-Baron designed the research program, including the empirical methodologies and analyses presented in Chapters 3–5. The author independently selected the topic of exploring the psycholinguistic relationship between Pidgin and ‘ōlelo Hawai‘i and formulated the research questions. The author retains primary responsibility for the conception, design, data collection, and analysis of all included studies, unless otherwise indicated in individual chapters. Portions of this dissertation are intended for future publication.

All projects and associated methods were approved by the Behavioural Research Ethics Board of the University of British Columbia (certificate #H23-02161) and were reviewed and approved in compliance with the Institutional Review Board of the University of Hawai‘i at Mānoa.

Generative AI tools (e.g., ChatGPT and Grammarly, as well as AI features integrated into LaTeX editors) were occasionally used for editorial purposes only. In practice, text or code was input into these tools to obtain suggestions for debugging of LaTeX, Python, and R code, grammatical corrections, and assisting with the organization and formatting of bibliographic entries. Any suggestions provided by Generative AI were carefully reviewed and independently verified for accuracy. Generative AI was not used in any other capacity.

Table of Contents

Abstract	iii
Lay Summary	v
Preface	vi
Table of Contents	vii
List of Tables	xi
List of Figures	xvi
Glossary	xix
Acknowledgments	xxiii
Dedication	xxxii
1 Introduction	1
1.1 Research Focus	1
1.2 The Linguistic Landscape of Hawai‘i	5
1.3 Linguistic Description of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English	9
1.3.1 Research Context	14
1.4 Outline of Dissertation	16
2 Mo‘olelo, “talk story”, from work done in-community	17
2.1 Introduction	17
2.2 ‘Ohana and Hoa	18
2.2.1 Mo‘olelo I: Bridging the sacred space between ‘ike ku‘una and the present	18
2.2.2 Mo‘olelo II: Kapu and Kuleana	20
2.2.3 Mo‘olelo III: Pōhaku at the Pu‘uhonua O Hōnaunau	24

Table of Contents

2.2.4	Mo'olelo IV: Ka'aloa's Super J's Laulau	28
2.2.5	Mo'olelo V: Akamai Kid	30
2.3	Kumu and Haumāna	31
2.3.1	Mo'olelo I: Language Loss, Shame, and Collective Responsibility	31
2.3.2	Mo'olelo II: If you can, please kōkua me	35
2.3.3	Mo'olelo III: Tūtū lauhala lei hulu	36
2.3.4	Mo'olelo IV: Pidgin as Resistance in the Hawaiian Language Classroom	36
2.3.5	Mo'olelo V: Pidgin as Access, Care, and Relational Pedagogy	38
2.4	Lāhui, in Hawai'i and abroad	40
2.4.1	Mo'olelo I: Pidgin Stigma	40
2.4.2	Mo'olelo II: Hālau-Style Approaches to Revitalization	44
2.4.3	Mo'olelo III: On languages as living beings	46
2.4.4	Mo'olelo IV: Kānaka Maoli in the Diaspora: What happens when there are no speakers around you?	48
2.4.5	Mo'olelo V: Inter-community Work	49
2.5	Conclusion: Weaving the Lei	51
2.5.1	Kulāiwi	51
3	Experiment 1: A Psycholinguistic Exploration of Code-switching to 'ōlelo Hawai'i from Pidgin versus Hawai'i English	54
3.1	Introduction	54
3.2	Participants	63
3.2.1	Pidgin, English, 'ōlelo Hawai'i trilinguals	63
3.2.2	Inclusion and exclusion criteria	65
3.3	Materials	66
3.3.1	Phrases and Wordlist	66
3.3.2	Auditory stimuli	68
3.3.3	Visual stimuli	69
3.3.4	Language Background Questionnaire	69
3.4	Procedure	71
3.5	Analysis and Results	73
3.6	Discussion	91
3.7	Conclusion	95
3.8	Mo'olelo	96

Table of Contents

4	Experiment 2: Investigating Cross-Lingual Priming from Pidgin and English Primes to ‘ōlelo Hawai‘i Targets	97
4.1	Introduction	97
4.2	Participants	101
4.2.1	Inclusion and exclusion criteria	105
4.3	Materials	106
4.3.1	Wordlist	106
4.3.2	Creation of nonwords and selection of real words . . .	109
4.3.3	Creation of auditory stimuli	113
4.3.4	Language Background Questionnaire	114
4.4	Experimental Procedure	114
4.5	Analysis and Results	117
4.6	Discussion	123
4.7	Conclusion	131
4.8	Mo‘olelo	132
5	Experiment 3: Implicit knowledge of ‘ōlelo Hawai‘i by Pidgin speakers	134
5.1	Introduction	134
5.2	Participants	144
5.3	Experimental Procedure	146
5.3.1	Experimental Tasks	146
5.3.2	Well formedness Rating Task	147
5.3.3	Word Identification Task	148
5.3.4	Language background questionnaire and language attitudes survey	149
5.4	Stimuli and Materials	149
5.4.1	Overall stimuli	149
5.4.2	Corpus	150
5.4.3	Stimuli Limitations	151
5.4.4	Calculating phonotactic probabilities	152
5.4.5	Construction and Selection of nonword Stimuli	155
5.5	Data inclusion criteria and participant exclusion	157
5.6	Well Formedness Task: Statistical Analyses and Results . . .	157
5.7	Well formedness Task: Discussion	165
5.8	Word Identification Task: Statistical Analyses and Results .	170
5.8.1	Bayesian Model and Model Interpretation	170
5.8.2	Estimated Marginal Means and Trend Analyses	173
5.8.3	Results: Posterior Estimates from the Bayesian Mixed-Effects Model	176

Table of Contents

5.8.4 Results: Estimated Marginal Means and Trend Anal- yses	180
5.9 Word Identification Task: Discussion	187
5.10 Conclusion	192
5.11 Mo'olelo	195
6 Conclusion	199
Bibliography	206
 Appendices	
A Language Background Questionnaire	247
B Experimental Stimuli: Experiment 1	256
C Experimental Stimuli: Experiment 2	257
D Experimental Stimuli: Experiment 3 Task 1	293
E Experimental Stimuli: Experiment 3 Task 2	310

List of Tables

3.1	Code-Switched Stimuli List 1	67
3.2	Code-Switched Stimuli List 2	67
3.3	Mean accuracy and standard deviation by frame language and code-switching pattern. Accuracy reflects proportion of correct responses across trials.	74
3.4	Mean Accuracy for Participants 9796624 to 10163058 by Code-switch Condition and Frame Language, including number of trials completed after removal of trials with response times outside 1400–5000 ms.	77
3.5	Mean Accuracy for Participants 10199820 to 11155512 by Code-switch Condition and Frame Language, including number of trials completed after removal of trials with response times outside 1400–5000 ms.	78
3.6	Posterior estimates from the Bayesian multilevel logistic regression model predicting accuracy. The table reports fixed effects for code-switching pattern, frame language, and their interaction, as well as variance components for participant and item level random effects. Estimates are presented as posterior means with standard errors and 95% credible intervals.	79
3.7	Population-level (fixed-effect) estimates from the Bayesian mixed-effects model estimating log reaction times (log RTs). 95% credible intervals (CI) are shown alongside posterior mean estimates.	88

List of Tables

4.1	Cross-lingual stimulus types used in Experiment 2. Within each list, ‘ōlelo Hawai‘i target items were evenly divided between translation-equivalent trials and non-translation control trials, such that each target appeared either with a translation-equivalent prime or with an unrelated control prime. Prime languages were counterbalanced across targets, with half of translation and control trials preceded by Hawai‘i English primes and half by Pidgin primes. Filler trials consisted of unique prime–target pairings.	106
4.2	Cross-language translation equivalents	108
4.3	Overall accuracy across trial type after exclusion of participants performing under 80%.	118
4.4	Overall accuracy across prime language and target type after exclusion of participants performing under 80%.	118
4.5	Population-level effects from the Bayesian mixed-effects model estimating log reaction times (log RTs), including fixed effects for Prime Language (Pidgin vs. Hawai‘i English), Trial Type (translation equivalent vs. control), and their interaction. . .	122
4.6	Population-level effects from the Bayesian lognormal model for language consultant participants.	128
4.7	Population-level effects from the Bayesian binomial regression model of accuracy with participants that scores about 85% or more.	130
5.1	Fixed effects from a Bayesian-mixed effects cumulative probit model predicting participants’ ordinal well formedness ratings of Hawaiian-like nonwords. The model estimates how phonotactic probability and language background influence these judgments, as well as their interactions, with Hawaiian speakers serve as the reference group for all comparisons. Estimates with lower absolute magnitudes reflect response behavior that is less differentiated from the Hawaiian speaker baseline, indicating greater similarity in well formedness judgments. Positive estimates indicate higher well formedness ratings, while negative estimates indicate lower ratings relative to the Hawaiian speaker baseline. 95% CIs excluding zero are interpreted as meaningfully different from baseline, while intervals including zero suggest uncertainty or limited evidence of meaningful difference.	161

5.2	Population-level effects from the Bayesian cumulative probit model predicting participants' ordinal well formedness ratings of Hawaiian-like nonwords excluding responses from stimuli with uniquely Hawaiian orthographic features. Estimates show the effect of phonotactic probability and language group, as well as their interactions. Positive estimates indicate higher ratings relative to the reference group (Hawaiian speakers), while negative estimates indicate lower ratings. 95% CIs that exclude zero indicate a reliable difference from baseline. . . .	165
5.3	Population-level effects from the Bayesian model predicting confidence ratings. The model includes fixed effects for phonotactic probability, language background, and item type, as well as all interactions among these predictors. Estimates represent posterior means with 95% credible intervals (CIs). Effects whose CIs exclude zero are interpreted as credible differences. Positive coefficients indicate higher confidence ratings (i.e., stronger confidence that an item is a real word) as a function of the corresponding predictor. Interaction terms capture how these relationships vary across language backgrounds and item types.	177
5.4	Posterior contrasts of estimated marginal means (emmeans) analysis for confidence ratings presented on the probit scale. Each row shows the median posterior contrast comparing the first group listed to the second, along with the 95% highest posterior density interval. These contrasts summarize within-group differences between real and nonword item confidence ratings for Hawaiian speakers, Pidgin speakers, English speakers from Hawai'i, and English speakers from Canada.	180
5.5	Posterior contrasts of estimated marginal means (emmeans) for confidence ratings on the probit scale. Rows are grouped by word type (real vs nonword). Each contrast reports the median posterior estimate comparing the first group to the second, along with the 95% highest posterior density interval.	181

List of Tables

5.6	Posterior contrasts from the estimated trends (emmtrends) analysis, showing differences in phonotactic probability slopes across all combinations of language group and item type. Each contrast reports the median posterior difference in slope between the first group listed and the second, along with the associated 95% highest posterior density interval. These results summarize pairwise differences in how confidence ratings change as a function of phonotactic probability for Hawaiian, Pidgin, English from Hawai'i, and English from Canada speakers for both real and nonword items.	184
5.7	Posterior contrasts from the estimated trends (emmtrends) analysis, showing differences in phonotactic probability slopes across all combinations of language group and item type. Each contrast reports the median posterior difference in slope between the first group listed and the second, along with the associated 95% highest posterior density interval. These results summarize pairwise differences in how confidence ratings change as a function of phonotactic probability within each language population for real and nonword stimuli items. . . .	186
B.1	Unilingual Stimuli List (Control Items, Identical Across Participants)	256
C.1	List 1 cross-language priming task stimuli	257
C.2	List 2 cross-language priming task stimuli	266
C.3	List 3 cross-language priming task stimuli	275
C.4	List 4 cross-language priming task stimuli	283
D.1	Participants saw only the nonword stimuli. Phonotactic probabilities for the nonwords and their matched real words are included for comparison; matched real words were selected to minimize phonotactic probability differences. Note that phonotactic probability values are reported on a log scale; negative values reflect the logarithmic transformation of probabilities less than 1.	293

E.1	Full possible stimuli list for the Word Identification Task, including the complete pool of real and nonreal word stimuli and their calculated phonotactic probabilities. Phonotactic probability values are reported on a log scale, where negative values reflect the logarithmic transformation of probabilities less than 1. For clarity, the stimuli presented in the experiment were selected as subsets from this full stimulus pool, with each experimental list containing a different subset of items. Items were semi-randomly assigned to lists to ensure comparable phonotactic probability means and distributions across lists, as well as an equal number of real and nonreal stimuli and other experimental considerations. Participants were randomly assigned to one of the experimental lists. See Chapter 5, Methods for further information.	310
-----	--	-----

List of Figures

2.1	The Pōhaku card from the Mana Card Deck (Becker [1998]). Pōhaku, meaning stone, evokes our connection to the source, to others, and to our communities, offering strength, inspiration, and purpose. Photo taken by Katrina Karl.	27
2.2	Akamai Kid, published in 1974 at Konawaena High School, written by Karen Kamigaki	31
3.1	Cartoon data illustrating the interaction between frame language and trial type such that code-switching (CS) costs are lower in Hawai'i English on the left; and code-switching costs (CS) are lower in Pidgin on the right.	59
3.2	Distribution of age of acquisition (AoA) ratings for Pidgin, English, and 'ōlelo Hawai'i across participants. Values are plotted separately by language, allowing comparison of AoA distributions across the three language groups.	64
3.3	Distribution of self-reported proficiency ratings for Pidgin, English, and 'ōlelo Hawai'i across participants. Values are plotted separately by language, allowing comparison of proficiency distributions across the three language groups.	65
3.4	Mean proportion of correct responses (Accuracy) (0 = incorrect, 1 = correct) by code-switch condition, separated by frame language for participants among qualifying participants. Points represent condition means, lines connect means across conditions within each panel, and error bars indicate ± 1 standard error. Numeric labels show mean with standard deviation in parentheses.	75
3.5	Individual variation in accuracy by frame language and trial type without removal of participants. Each point represents a participant's mean accuracy for a given combination of frame language, Hawai'i English or Pidgin (on left, and right, respectively) and trial type (unilingual or code-switched to Hawaiian).	76

List of Figures

3.6	Mean response times (in milliseconds) by trial type (unilingual or code-switched speech) and frame language (Hawai'i English or Pidgin) from empirical results. Each point reflects the grand meanRT across participants for a given condition, with error bars representing the standard error of the subject-level means.	83
3.7	Individual Variation in Mean Reaction Times. Mean reaction times (RTs) by trial type and frame language for each participant. Error bars represent the standard error of the mean, capturing trial-level variability in response times within each condition. While group-level trends are visible, substantial individual variability is observed, particularly across language conditions. Participants who did not follow group-level patterns are marked with colored shapes next to their participant ID: red triangles indicate participants faster in code-switched than unilingual trials overall, blue circles indicate participants faster in Pidgin unilingual trials compared to Hawai'i English unilingual trials, and purple squares highlight participants who were faster in Hawai'i English unilingual trials and also faster in Hawai'i English code-switched trials.	85
3.8	Individual code-switching costs. Each point represents the mean response time difference between code-switched and unilingual trials for each participant. Positive values indicate slower reaction times for code-switched trials relative to unilingual trials, while negative values indicate the opposite.	90
4.1	Distribution of age of acquisition (AoA) ratings for Pidgin, English, and 'ōlelo Hawai'i across participants. Values are plotted separately by language, allowing comparison of AoA distributions across the three language groups.	104
4.2	Distribution of self-reported proficiency ratings for Pidgin, English, and 'ōlelo Hawai'i across participants. Values are plotted separately by language, allowing comparison of proficiency distributions across the three language groups.	105
4.3	Histograms with RT distributions for participants with greater than 80% accuracy.	119
4.4	RT differences by trial type and prime language for correct responses only.	120

List of Figures

- 5.1 Well formedness ratings as a function of phonotactic probability and language background. The figure displays empirical participant responses, with phonotactic probability on the x-axis and well formedness ratings on the y-axis, shown for each stimulus and stratified by language group (green = English participants from Hawai‘i, orange = English speakers from Canada, purple = Hawaiian speakers, pink = Pidgin speakers from Hawai‘i). Lines show group-specific linear fits to the observed responses, with shaded regions indicating standard error. Higher phonotactic probability values correspond to items that are more phonotactically attested in the corpus, and higher response values reflect greater perceived well formedness. Phonotactic scores appear as negative values because they are expressed on a log probability scale: more negative values indicate lower phonotactic probabilities, whereas values closer to zero indicate higher phonotactic probabilities. 160
- 5.2 Well formedness ratings for each stimulus as a function of phonotactic probability and language background, shown for items without uniquely Hawaiian orthography (kahakō or ‘okina). The figure displays empirical participant responses, with phonotactic probability on the x-axis and well formedness ratings on the y-axis, shown for each stimulus and stratified by language group (green = English participants from Hawai‘i, orange = English speakers from Canada, purple = Hawaiian speakers, pink = Pidgin speakers from Hawai‘i). 164
- 5.3 Empirical relationship between phonotactic probability and mean confidence rating responses in the word identification task across four language groups. Each panel displays one language group: Hawaiian speakers, Pidgin speakers, English speakers from Hawai‘i, and English speakers from outside Hawai‘i. The x-axis represents the phonotactic probability of each item (the likelihood of its sound sequence occurring in Hawaiian), while the y-axis shows the mean response across participants for that item. Data points represent individual items, with separate colors for real Hawaiian words (green) and nonword items (orange). 176

Glossary

‘ōlelo Hawai‘i Glossary

The following glossary provides English glosses for ‘ōlelo Hawai‘i terms used throughout this dissertation. In the main text, many ‘ōlelo Hawai‘i items appear within full utterances, quotations, or experimental stimuli; in these cases, glosses are provided in-text at first occurrence or where necessary for clarity, rather than exhaustively for every token. Glossing is therefore selective and context-dependent, with the aim of supporting interpretation without overburdening the reader. The following translation equivalents are largely drawn from Mary Kawena Pukui and Samuel Elbert’s *Hawaiian Dictionary* (Pukui and Elbert [1986]), found on the Ulukau website (Kuamo‘o [2009]). All mistakes are my own.

‘ōlelo Hawai‘i	English translation equivalent, gloss
‘ahi	tuna
‘āina	land, earth
‘āina momona	land of plenty
akamai	smart
ali‘i	royalty
aloha	love, affection, compassion, kindness, mercy, and charity, in addition to being a term used to greet, welcome, or express feelings of regard
aloha ‘āina	aloha for land, lit: aloha for that which nourishes; patriotism
ao ‘aumakua	the land, earth, realm of the ancestors
‘aumakua	family or personal gods, deified ancestors
ea	sovereignty, life, breath
haole	White person, foreigner
haumāna	students
hālau	long house, as for canoes or hula instruction
hoa	friend, friendship

‘ōlelo Hawai‘i	English translation equivalent, gloss
hoa‘āina	friends of ‘āina
hū	huh, oh, grunt, hum (interj.)
huaka‘i	journey
hula	dance
hulu	feather
imu	underground oven
‘ike	knowledge, understanding, to recognize, to perceive, to be aware, to understand
‘ike ‘āina	knowledge from or about the land
‘ike ku‘una	ancestral knowledge
ka/ke	definite singular article
kalo	taro
kama‘āina	native-born, one born in a place
Kānaka Maoli	native Hawaiian person
kahuna	priest
kapu	taboo, law, sacredness, prohibited, forbidden, sacred
keiki	child, children
kōkua	help
kuleana	a profound sense of responsibility, commitment, duty, right, claim
kumu	teacher, teachers
kūpuna	elder, elders
lāhui	nation, race, community
lauhala	the leaves of the hala tree (<i>Pandanus tectorius</i>)
lei	garland
lo‘i	pond
maika‘i	good, great
mālama	to take care of, tend, preserve, protect
maka‘āinana	common class
mana‘o	wisdom, thought, idea, belief, opinion
manini	convict tang (<i>Acanthurus triostegus</i>)
mo‘okū‘auhau	genealogy, of ties to place and people
mo‘olelo	story

‘ōlelo Hawai‘i Glossary

‘ōlelo Hawai‘i	English translation equivalent, gloss
Nī‘ihau	the seventh largest island in Hawai‘i, where ‘ōlelo Hawai‘i is spoken as a primary language and remains the only place where the Indigenous language has not been displaced by English
noni	Indian mulberry (<i>Morinda citrifolia</i>)
‘ohana	family
‘ōiwi	of the bone, native
‘ōlelo	language (n), to speak (v)
‘ōlelo Hawai‘i	the Hawaiian language (n), to speak the Hawaiian language (v)
‘ōlelo no‘eau	proverb
‘ōlelo Pa‘i‘ai	lit: pounded-taro language, ‘ōlelo Hawai‘i word for the Pidgin language
oli	chant
‘opihi	limpet
pilina	mutually sustaining relationships
pōhaku	rock, stone
pono	righteous, goodness, morality, correct, proper
pu‘uhonua	place of refuge
tūtū	grandmother / grandfather
wai	water
wana	sea urchin
wela	hot

Pidgin Glossary

The following glossary provides English glosses for Pidgin terms used throughout this dissertation. As with ‘ōlelo Hawai‘i terminology, in the main text, many Pidgin items appear within full utterances, quotations, or experimental stimuli; in these cases, glosses are provided in-text at first occurrence or where necessary for clarity, rather than exhaustively for every token. Pidgin lexical items included in this glossary are drawn from Sakoda and Siegel’s book *Pidgin Grammar* (Sakoda and Siegel [2003]) and the author’s own intuitions as a Pidgin speaker.

Pidgin	English translation equivalent, gloss
aunteh	middle aged woman
baddah	bro
braddah	dude
bumbai	later, next time, or else
kānaka	Hawaiian person
liddis	like this
lolo	idiot (n), stupid (adj)
tsubu	sea snail
unko	middle aged man

Acknowledgments

First and foremost, thank you to my supervisor, Molly Babel.

Mahalo, Molly, for your mentorship, generosity with your time, and unwavering support. Thank you for giving me the freedom to pursue the questions that mattered most to me while always encouraging me to challenge myself and achieve more than I ever thought possible. Thank you also for the financial support throughout the program and for the opportunities you provided that made this work possible and shaped my development as a scholar.

You have helped me grow in ways that will stay with me for the rest of my life. Thank you for taking a chance on me and for seeing potential in me even when I struggled to see it in myself. I feel incredibly lucky that you took me on as your student, and I will always be grateful I got to learn from you.

Thank you to the rest of my committee members, Carla Hudson Kam and Scott Saft.

Thank you, Carla, for your insight, your careful attention to detail, and the thoughtful questions you brought to this work. Thank you for making me feel that this research question was important and meaningful from the perspective of language acquisition, and for helping me see the broader contributions this work could make. Mahalo e Scott for your support, your enthusiasm, and for connecting me with folks at Ka Haka ‘Ula O Ke‘elikōlani. I learned so much from you and your ‘ike, particularly through your work bridging linguistics, community, and activism in Hawai‘i. Your commitment to using linguistics in service of language communities has been deeply inspiring to me.

Thank you to Henry Davis for your constant encouragement, from the moment you sent the acceptance letter to the PhD program, to the very end on my defense day. Thank you for being such a steadfast advocate for me and for being so loud about it; your confidence in me gave me the

Acknowledgments

courage to keep pushing forward. Thank you for knowing when to challenge me to push myself further and when to remind me not to take myself too seriously. Thank you for checking-in on me throughout this process, through field methods classes, QP2, Secwepemetsín research, and job applications, and for always reminding me to stay connected to the reasons that inspired me to do this work. It is incredible to see how many lives you have touched through your work. Seeing the relationships you have built and the impact you have had has been a powerful model for the kind of community-centered language revitalization work I hope to be able to do.

Thank you to Amanda Cardoso for teaching me sociophonetics in my first class ever as a PhD student, for advising my QP2, and for taking me on “walk meetings” with Masa when I first moved to Vancouver. I am grateful not only for your intellectual mentorship and guidance, but also for the kindness and friendship you showed me, especially during that transition.

Thank you to Márton Sóskuthy for teaching me statistics, corpus linguistics, and quantitative methodologies. Your classes were foundational to my training as a researcher and were instrumental in shaping the analyses presented in this dissertation. Thank you also for serving on my QP1 committee. I remember thinking that there were never enough statistics classes I could take to finally become someone who felt confident working in R, basic Python, or Bayesian analyses. Thanks to your patience and willingness to work through things with me, I now feel like I have reached a point where I am genuinely comfortable with these tools (although I remain acutely aware of how much there is left to learn). Thank you for your random “pep talks” whenever I thought I was only making a self-deprecating joke, and for letting me borrow a lab laptop, which I am currently writing this dissertation on. It has kept me going since my computer exploded five years ago, and I promise I will give it back to you soon!

Thank you to Chris Hammerly for your guidance during the early stages of Experiment 1 through your eye-tracking methodologies class, and for helping me see things differently when I couldn’t make sense of the results initially. I also really appreciated sitting in on your field methods class and having the opportunity to see how you engage with language consultants as a community member yourself. Your example has had a lasting impact on me, and I have always appreciated having someone I could trust and look up to as I navigated the many challenges and uncertainties of balancing academic expectations with community goals. Go Oilers!

Acknowledgments

Thank you to everyone else at the University of British Columbia who has ever taught me, mentored me, or supported me throughout my PhD: Kathleen Currie Hall, Bryan Gick, Jahurul Islam, Hotze Rullmann, Anne-Michelle Tessier, Strang Burton, Gunnar Hansson, Lisa Matthewson, Eva Csipak, Garrett Nicolai, and Miikka Silfverberg.

Mahalo to everyone at the University of Hawai'i at Mānoa for your support throughout my Master's program.

Thank you to Katie Drager for introducing me to research on Pidgin and sociolinguistics, for being my advisor when I came into the Masters program knowing absolutely nothing about linguistics, and for showing me how linguistic knowledge can be applied to the world around us. Thank you for believing that I had a place in graduate school and for making me feel like I had a chance. Mahalo nui for encouraging me to pursue work on Pidgin, for the opportunities you have opened for me, and for helping shape me into the linguist I am today.

Mahalo to Rory Turnbull for teaching me Phonetics for the first time. This class had a profound influence on my research and later inspired my own teaching style when I had the opportunity to teach Phonetics as an instructor. Your mentorship, especially during the time when I was deciding whether a PhD was the right choice for me, gave me confidence in my decision to take this next step in my academic career. I am incredibly grateful for your guidance, encouragement, and the role you played in shaping me as a scholar and educator.

Mahalo to Ryan Henke for teaching me introductory linguistics and morphology, for showing me that linguistics is ultimately about the communities we can serve, and for sitting down with me to help me weigh the pros and cons of pursuing my PhD at UBC when I was feeling a bit uncertain about my decision. I frequently returned to those conversations when I encountered obstacles throughout the program, and they helped ease many of the anxieties I faced. Thank you for being such a kind, genuine, and refreshingly human presence in academia. Thank you for reminding me that success in academia is not just about being the person who knows the most in the room, but also about respecting people and communities, having clear goals, and bringing our own life experiences into the work we do.

Mahalo to Amy Schafer for introducing me to psycholinguistics and for

Acknowledgments

opening my eyes to the questions that would eventually shape this work. I read several instrumental papers because of your recommendations, many of which became foundational influences on my research found here in this dissertation. Your support and confidence in me had a lasting impact, and I am deeply grateful for your mentorship.

Mahalo to everyone else at the University of Hawai'i who has supported me throughout my academic journey. Mahalo to Andrew Cheng and Theres Grüter for helping me recruit participants at the University of Hawai'i at Mānoa and for the many helpful conversations about the technical and community aspects of my research. Mahalo to Andrea Berez-Kroeker for accepting me into the Master's program, even when I approached you after the application deadline and had not yet taken any linguistics classes. Thank you for giving me a chance when all I had was a research question. Mahalo to William O'Grady for the reminder to always remember to "show you the evidence."

Kheyli mamnoon to Ladan Hamedani for your letter of recommendation into the Linguistics Master's program, for being the first person to introduce me to linguistics as a field, and for teaching me Farsi. This experience played a huge role in shaping my early research questions about the acquisition of sound patterns from the perspective of a new language learner. Man az shoma besiyar sepaşgozaram.

Mahalo e Kumu Larry Kimura and Kent Sakoda for being so kind to me. Meeting each of you for the first time was intimidating because you are both giants in the field and in our community, but how lovely it was to become acquainted with you both and to receive your mentorship, friendship, and care. How lucky Hawai'i is to have you both, who have dedicated so much to the community with such grace and humility behind your profound expertise. I am grateful to have found your support as I navigated conducting research in community myself. I hope to honor the guidance you have shared with me by continuing to approach research with aloha and a profound sense of kuleana, carrying out work that serves our families, friends, and the people of Hawai'i. I look forward to continuing to learn from you both.

Thank you to the Public Scholars Initiative for contributing to the funding of this research. Thank you to Serbulent Turan for your mentorship throughout the Public Scholars Program. It was an honor to be a Public Scholar, and the experience was invaluable for both my academic and

Acknowledgments

personal growth.

Thank you to those at Simon Fraser University who mentored me during my first year as a sessional instructor, especially Panayiotis Pappas, Eva Tchizmarova, Sylvia Cho, and Claudia Hein. I have learned so much.

Mahalo nui for your kōkua to everyone who I have consulted for your expertise on ‘ōlelo Hawai‘i linguistics, Hawaiian language education, and the Hawaiian activism that informs this work: Ha‘alilo Solomon, Kahiau Snyder, Bruce Torres Fischer, Keao NeSmith, Kalikoaloha Martin, Ali Rozet, Oiwi Parker Jones, Kalāmanamana Harman, Kekoa Harman, Toma Thomas Kettig, Kapeliela Gabriel Gilbert, Noah Dolim, Pili Kamakea-Young, BJ Akui, Welaahilani Wāhilani, Iāsona Ellinwood, and Hoaloha Westcott.

Mahalo to everyone who has participated in one of my studies and to the many people I have met along the way who have shared their time, stories, and mana‘o with me about language and different ways of expressing aloha ‘āina. To the kumu, haumāna, the kūpuna and keiki, to strangers and friends: this work could not have been done without you and your kōkua.

Kukwstsétsemc to Garlene Dodson, Louella Jules, Leona Calhoun, (the late) Daniel Calhoun, and Ronald Ignace for your patience and kindness in helping me learn about Secwepemctsín, as well as your culture and community. Miigwech to Koko Jones for telling me your fun stories about amik in Anishinaabemowin.

Thank you to the Speech in Context Lab, the OoPs Lab, and the Secwepemctsín Working Group for your feedback on my work at various stages. Thank you to Mandeep Khokhar for assistance with coding and cleaning the multilingual corpus, and to Jessica Lee and Emilie Russo for help with annotations.

Thank you to Roger Lo, my friend, classmate, and mentor. Thank you for always being willing to share your expertise and provide guidance on statistical and computational methods throughout the course of my PhD. I still remember the days in my first year when I was struggling to get even a single ggplot to run. From those early days of troubleshooting my first R scripts to helping me set up language models and measure phonotactic probabilities for Experiment 3 of this thesis, your guidance helped me develop the skills needed to approach increasingly complex computational analyses.

Acknowledgments

Mahalo e V. Pauahi Souza for collaborating with me on “Access to ‘ōlelo Hawai‘i Education for Kānaka Maoli Living in the Diaspora” and for inviting me to participate in the Moana Nui Podcast episode you led. Thank you for sharing your lived experiences as a diasporic Kanaka Maoli; your insight has deeply influenced how I approach this work and think about language, identity, and community. Mahalo for being my *hoa aloha* in British Columbia as we continue doing work for our communities in Hawai‘i from Canada.

Thank you to my dear colleagues and friends from my academic *huaka‘i* for always being there to laugh (and cry) with me, and for supporting me both intellectually and emotionally: Aani Perkins, Rachel Soo, Ana Arrieta Zamudio (thanks for hosting movie night and being such a wonderful person), Jessica Charest (would have croaked without you during the MA), Kate Zhou, Sander Nederveen (for co-authoring the “shoe theory” and for my beloved Salomons), Suyuan Liu, Ivan Bondoc, Micah Mizukami, Krystal Naumu, Chenxi Xu, Kaili Vesik, Sijia Zhang, Alex Ayala, Gazelle Khalaji Pirbaluti, Oksana Tkachman, Katherine Arnold, Weirui Chen, Gloria Mellesmoen, Abiodun Ibikunle, Noah Luntzlara, Bruce Oliver, Michael Schwan, Changbing Yang, Niko Andrés Schwarz-Acosta, Arian Shamei, Laura Griffin, Pejman Ghaemi, Christina Cook, Ruby Ching, and many others. Thank you Janet Onopa for hosting me for Sunday dinners during the first year of the PhD program.

Thank you, especially, to the following people for all the love and care a girl could ever dream of. Thank you for being my family, for celebrating every win, for mourning every loss with me, and for making even the most ordinary moments so joyful:

Daniel Cairns (for so many reasons, I could not have done this without you. words cannot describe what you mean to me. thank you for being my person and for our adventures. thank you for finding me).

Katrina Karl (+Thomas Goedecke + Emma gurl, I know ‘Iolani people don’t have emotions, but our friendship makes me feel that I can get through anything).

Riku Mizuta (thank you for being my bestie, for your endless support and love, and for literally crying with me and laughing with me through every high and low life brings).

Acknowledgments

Yoonseok Choi (tanks, ajushi, for letting me unleash my very “demure” auntie self every day).

Zachary Nanji (thank you for being the gentlest of raccoon friends, with a good amount of canadian goose energy).

Lucia Amore (thank God you became a real doctor and not me. thank you for the times spent together in Vancouver and Seattle, making time for me despite the stresses of medical residency, and for being my sista since the days we wore our lab goggles outside of our regular glasses like the super nerds we were/are. it’s been so beautiful growing up with you).

Marta Helgadóttir (the universe really thought it could get away with lazily making two copies of the same person and plopping us on opposite sides of the world, but of course we made a way to find each other. thank you for being my twin, for making me a godmother, and for giving me +3 baby friends. and thank you, +Mark Villegas, for jumping in to help with the Pidgin podcast, for somehow also fixing my mom’s broken car, and for taking care of my sister that time my sister came to your tattoo parlor).

Janet Tran (thank you for a friendship that’s now old enough to vote and smoke cigarettes, and for every grandma walk we’ve shared and all the ones still to come).

Ian Higuchi (thank you for being my big brother and fighting with anyone who dares cross me since high school French class).

Aaditya Kulkarni (my first and best faka, how could I ever have made it without you?)

Lindsey Nkem (thank you for living with me for five years and for all the support).

Thank you to everyone in Daniel’s huge family for making me feel like part of the family, for taking me in to celebrate holidays, and for keeping me warm through the -40°C Edmonton winters. To all his grandparents, parents, siblings (and siblings-in-law), aunts, uncles, cousins, and one very special niece: thank you. All of you mean so much to me. Thank you especially to Alison Cairns, for our daily Wordle, Connections, and Don’t Wordle; for our true crime discussions; for making my website; and for being a constant presence of generosity, support, and care. The regular check-ins

Acknowledgments

helped me get through the day, from sharing little (and sometimes not-so-little) updates to laughing about a strange encounter. Mahalo nui, I am so grateful.

Thank you to the many people I met at Green College. Thank you to Emma Cunliffe for your service as principal, and thank you to Carolina Sartor for being the first person who made me feel like I could belong here. Thank you for our thrifting dates, and for gifting me my first jacket when I moved here and looked cold. Thank you to Peter Law for the dimsum parties and for inviting me to that parade where we all got to shake hands with Justin Trudeau. Thank you to all the friends I've gotten to spend time with throughout my time living at Green. I wish I had space to say more about each of you and the ways each of you mean to me, but for now, thank you so much: Michael Carelse, Jaya Scott, Angel Bella, Dhanya Indraganti, Sabrina Draude, Anna Navarro, Jeremy van der Heijden, Seon Shim, Sara Li, Carmen Gohn, Alex Banks, Morris Prosser, Keegan Colwell, Eduardo Jimenez, Ben Zammit, Reina Nadler, Clara Kiyoko Kumagai, Jasmine Yao, Claire Hein-Salvi, Michael Banerjee, William Huang, Erin Evans, Alejandro Rojas, Yingxiang Li, Shawn Nirody, and Zachary Couture. I am sure I am forgetting people unintentionally but thank you all so much for making Green College such a wonderful experience.

I would like to acknowledge the people who showed me what being a PhD student could look like and who helped make this path feel possible. There are many people I have met who fit this description (you know who you are), but I would especially like to give a shout out to my friends from USC: Sahand Pirbadian, Pouye Sedighian, Arash Roshani, and Juni Sengupta. From you I really came to understand what graduate life would be like first hand when I was an undergraduate. Thank you for being so nice to me and for all our adventures living in California together. Thank you also to Chad Shomura, for volunteering your time to mentor me when I was still in high school and for including me in your PhD thesis acknowledgments years before I ever imagined pursuing one myself. (Ho'omaika'i on your well-deserved tenure!)

Thank you to my family: my mom Karen Baron, my (late) father Daniel Baron (miss you), and my sister Danielle Baron. To my cousins, Lyla Kamigaki and Ciana Roldan, thank you for our facetime calls, and thanks Cristian for the soccer tickets! Thank you to my Hilo Family for throwing me dinner parties during my field work visits, Uncle Hal Kamigaki and auntie Lil,

Acknowledgments

Ashley, Mark, and kids! To my auntie Susan Kamigaki and my uncle Daryl Baron for frequently checking in on me. Thank you to my grandparents and great-grandparents for bringing me to Honaunau Beach, for teaching me Pidgin, and for instilling the values that I carried throughout this work. This work is inspired by you and the lessons you passed down to me.

Me ke aloha.

Dedication

This dissertation is dedicated to:

my mother, da Akamai kid

to my late father who told everyone after I won a silly award in middle school
that he knew I would one day become a doctor (well, probably wrong kind
of doctor but hi Dad!)

and to Hawai'i nei (its people and the land, mahalo nui I am forever indebted)

Chapter 1

Introduction

1.1 Research Focus

This dissertation examines the psycholinguistic connections between two local languages spoken in Hawai‘i, Pidgin (Hawai‘i Creole) and ‘ōlelo Hawai‘i (Hawaiian), in comparison to Hawai‘i English, the local prestige language. Given prior evidence that Pidgin and ‘ōlelo Hawai‘i overlap across multiple linguistic, social, and contextual dimensions, this dissertation investigates how such overlap may be leveraged in speech processing by examining psycholinguistic connections between these two languages, relative to English. Grounded in ‘ōiwi epistemology, this dissertation approaches these connections not merely as isolated psycholinguistic phenomena, but as processes embedded within community, shaped by intergenerational histories, and understood in relation to community priorities. My analytic approach is guided by my own kuleana as a researcher with personal ties to this work. Accordingly, the central research question investigates how shared linguistic and sociocultural knowledge across Pidgin and ‘ōlelo Hawai‘i is leveraged during speech processing, and what comparison with English reveals about these patterns.

“E lei kau, e lei ho‘oilo i ke aloha.”

Wear aloha like a lei, throughout all seasons.

I begin this thesis with this ‘ōlelo no‘eau as a guiding reminder throughout my work to continually return to aloha at all stages of this research: in working with participants, in speaking about this work in public, in receiving mana‘o from community members, and in documenting this work now.

As this ‘ōlelo no‘eau is also included in the introduction to Mary Kawena Pukui’s *‘Ōlelo No‘eau: Hawaiian Proverbs and Poetical Sayings*, I hope to begin this thesis following the lead of one of the most influential Hawaiian language and cultural scholars in Hawaiian history. This stands as an intentional acknowledgment that this work is not mine alone, but is crucially built upon foundations laid by the kumu who committed their lives to work

involving the Hawaiian language. As I, along with five generations of my family before me, was born, raised, and educated in Hawai'i, the education that informs this dissertation is grounded in teachings passed down through elders such as kumu Mary Kawena Pukui. I honor kumu Mary Kawena Pukui from the outset to recognize how her teachings have shaped my intergenerational education and to foreground the foundational teachings that much language work in Hawai'i continues to build upon.

Like a lei, this dissertation weaves together many parts, beginning with mo'olelo of events which transpired while conducting this research, and following with three separate experimental studies with distinct methodologies. The common thread throughout is the return to aloha. What is the meaning of aloha? The term has been so heavily marketed that it has taken on a new meaning. In the Hawaiian Dictionary (Pukui and Elbert [1986]), also written by Kumu Mary Kawena Pukui, the term aloha is a concept that encompasses meanings we might translate into English as love, affection, compassion, kindness, mercy, and charity, in addition to being a term used to greet, welcome, or express feelings of regard.

At the forefront of the expression of aloha is aloha 'āina, defined as aloha for that which nourishes. Central to Hawaiian epistemology, aloha 'āina is a perspective which captures, in essence, that a vital part of receiving is giving back. This reciprocal relationship encompasses not just the literal land but also in pilina, 'ike 'āina, mo'okū'auhau, lāhui, and to 'ōlelo. All are understood as living entities that must be cared for to sustain life in the past, present, and future.

Under this framework, returning to aloha as a common thread throughout the process meant producing research that supports the well-being of Hawai'i as a whole, including the land, the language, the public education system, and our local communities. It meant centering mutual pilina with lāhui, mo'okū'auhau, 'ike 'āina, and 'ōlelo. My intention with this work is therefore not only to answer my research question but also to challenge how we design and carry out experimental research when we conduct work that is community driven. As a member of the community, it is important for me to not just claim identity and then call my work community-driven. It meant thinking more critically about my place in this community, the linguistic methodologies I selected, and the relationships I built with people across a range of professions, fields, and roles. Throughout this process, I sought to elevate the concerns and priorities expressed by community members, while deferring to the goals and priorities that consistently emerged from these interactions and allowing them to guide this work. At times, admittedly, the commitments of traditional academia and the practice of community cen-

tered work are in conflict. This highlights the fact that institutions were historically not built to align with community needs, values, or ways of prioritizing knowledge and goals. In fact, it was largely built to reject, exclude, objectify, exploit, and profit from communities, including but not limited to the language communities at the center of this work. I describe this friction, and consider the implications of this work in community, predominantly in my mo‘olelo section, but also in the conclusions of each experimental chapter. As a scholar now living and receiving my education in British Columbia, I found it important to carry forward the commitments that guide my work in Hawai‘i while learning from and engaging with the communities where I am situated, including First Nations language communities.

As Candace Fujikane, a prominent kama‘āina scholar and activist cautions, “There are political responsibilities to claiming any identity,” especially when “local”¹ identity is often used as a means of self-legitimation at the expense of peoples who face ongoing political struggles in Hawai‘i” (Fujikane [1997]). Although I am an intergenerational kama‘āina, and am recognized as such when returning to Hawai‘i, I am not Kānaka Maoli. I am mixed sixth-generation Japanese, with ancestral ties to Hiroshima. While my work on the relationship between Pidgin and ‘ōlelo Hawai‘i is very personal to me and is largely rooted in my ancestors’ lives as Hawaiian Kingdom subjects who spoke and lived these languages on the land, I also approach questions of kuleana and ‘āina-based practice with an awareness of my position as a non-Kānaka Maoli scholar.

Noelani Goodyear Ka‘ōpua (Goodyear-Ka‘ōpua [2013]) argues that aloha ‘āina for non-Kānaka Maoli kama‘āina means taking responsibility for ‘āina by actively i) supporting Kānaka Maoli who have been detached from ‘āina in reestablishing their connections to ‘āina and ii) helping to rebuild ‘ōiwi structures that create the conditions for transforming settler colonial relations. Furthermore, as Hōkūlani K. Aikau (‘Aikau [2019]) argues, “there is no state of being *hoa‘āina* (friends of ‘āina) at which one arrives, only a continual process of learning from and giving back to the ‘āina.” Working within a framework of non-Kānaka Maoli kama‘āina positionality, I ground

¹A social category used in Hawai‘i, narrower in scope than the English term *local*, referring not just to residency but to individuals who have typically lived in Hawai‘i across generations, especially as Kānaka Maoli or as descendants of plantation laborers (e.g., Miyares [2008]). Trask (2000) criticizes the term ‘Local’ as applied to settlers, whose socio-economic and political success, she argues, both reinforces the American ‘nation of immigrants’ myth and reproduces a new foreign hegemony at the expense of Native Hawaiians. Through using the term ‘Local’ here, I do not imply that anyone but Native Hawaiians have a legitimate claim to Hawai‘i. Instead, we use the term in line with its widespread use in the islands.

this work in the mo‘olelo of my genealogy as non-Kānaka Maoli alongside an ongoing huaka‘i. This huaka‘i includes coming to terms with what complicity in settler colonialism means across the past, present, and future for plantation settlers in Hawai‘i in conducting research with languages. This discussion therefore runs as a central thread throughout the thesis.

It is also necessary to situate Pidgin within the broader historical conditions of settler colonialism in Hawai‘i. The relationship between ‘ōlelo Hawai‘i, Pidgin, and English cannot be reduced to a binary distinction between haole and Kānaka Maoli, but instead reflects a more complex set of historical and sociolinguistic relations shaped by settler colonialism and multilingual contact. As scholars such as Candace Fujikane and Haunani-Kay Trask have emphasized (e.g., Fujikane [1997], Trask [1999]), the social and linguistic landscape of Hawai‘i has been profoundly shaped by the displacement of Kānaka Maoli from their land, language, and systems of knowledge. Within this context, while Pidgin has functioned and continues to function as a language of resistance and local identity, particularly among Kānaka Maoli and settlers of color whose communities experienced exploitation under the same colonial systems that brought indentured labor to Hawai‘i, the language is also entangled in the historical processes that contributed to the marginalization of ‘ōlelo Hawai‘i. Indeed, for many Kānaka Maoli, particularly during periods when access to ‘ōlelo Hawai‘i was actively suppressed, Pidgin emerged as a common language not by choice, but rather because colonial conditions reshaped the linguistic landscape of Hawai‘i. On one hand, Pidgin continues to exist within structures that stigmatize its speakers, reflecting broader systems of racism and the colonial privileging of English as the dominant language of power. At the same time, however, Pidgin is neither equivalent to ‘ōlelo Hawai‘i nor English in its social position, a distinction shaped by Hawai‘i’s colonial history. For many speakers, the rise of Pidgin as the majority first language occurred under conditions whereby the intergenerational transmission of ‘ōlelo Hawai‘i had been disrupted. In this sense, Pidgin’s emergence and uptake are inseparable from the broader project of colonial assimilation, even as its use may simultaneously encode resistance to it. Pidgin may therefore also carry affective traces of this history, reflecting not only resilience and local identity, but also the enduring impacts of the displacement of Kānaka Maoli, ‘ōlelo Hawai‘i, and ‘ōiwi epistemologies.

A part of my huaka‘i in carrying out this research is acknowledgment of valid criticism of adopting aloha ‘āina as a land-based framework for non-Kānaka Maoli kama‘āina. As Kānaka Maoli scholar Uahikea Maile critiques, to be of the ‘āina means to be of aloha ‘āina, for the relationship between

them is inherently one of ‘āina to ‘āina (Maile [2024]). Aloha ‘āina, from this perspective, is not simply an ethic but a physical and affective force that draws Kānaka Maoli to land. Maile argues that when non-Kānaka Maoli claim or express aloha ‘āina, whether they are kama‘āina or not, it risks obscuring its meaning. To be clear, this thesis does not incorporate the framework of aloha ‘āina as a way of dismissing or disputing these critiques. Instead, I operate from aloha ‘āina because it is the orientation I was raised with by my kumu and kūpuna to navigate the world, while remaining attentive to its complexities within ongoing discussions about kuleana for non-Kānaka Maoli. The question of kuleana inherently underlies the investigation of the connection between Pidgin and ‘ōlelo Hawai‘i, as Pidgin is the language which inherently contains the makeup of Kānaka Maoli and intergenerational plantation migrants. However, there is on-going debate, in efforts of ‘ōlelo Hawai‘i revitalization, on kuleana. These reflections are discussed throughout the dissertation and are addressed most fully in Chapter 2, Mo‘olelo, where the broader self-reflective framework for this work is established.

1.2 The Linguistic Landscape of Hawai‘i

‘ōlelo Hawai‘i, also known as Hawaiian, is an Austronesian language of the Polynesian subgroup. For centuries, it functioned as the single language of governance, education, and everyday life in Hawai‘i (Wilson [1998a], Wilson [1998b]). Beginning with contact in 1778, the Hawaiian Islands underwent sustained linguistic and sociopolitical change, with the arrival of missionaries from the United States in 1820 marking a significant expansion of English-language presence and the introduction of colonial ideologies (Reinecke [1938]). This shift had direct consequences for the educational system: while many schools serving maka‘āinana continued to provide instruction in ‘ōlelo Hawai‘i, elite schools were established with English-only policies to educate ali‘i and the children of white Americans (Shi [1990]).

During the nineteenth century, the Kānaka Maoli population sharply declined due to the introduction of foreign disease. Estimates place the pre-contact population between 800,000 and over one million (Stannard [1989]), a number that had fallen to 37,656 by 1900 (Nordyke [1989]). American missionaries framed this mass death as a moral failure of Hawaiian culture, attributing disease to cultural practices including oli, hula, and the use of ‘ōlelo Hawai‘i. Within this colonial logic, English-only education was established as a solution. Missionaries argued, “with this knowledge of English

will go into the young American republican and Christian ideas; and as this knowledge goes in, kahunaism, fetishism and heathenism generally will largely go out" (McArthur 1895, cited in Schütz [1994]).

These ideological shifts, including those centered on language, set the stage for the illegal overthrow of the Kingdom of Hawai'i. In 1893, the Kingdom was overthrown by U.S. businessmen with the support of U.S. Marines. In 1896, the Republic enacted Act 57, mandating English as the exclusive medium of instruction in all schools except those on the island of Ni'ihau (Territory of Hawai'i [1905]), despite 'ōlelo Hawai'i being spoken by a large majority of the population at the time (Schütz [1994]). The use of 'ōlelo Hawai'i on school grounds was prohibited (Wist [1940]), and schools that resisted were denied government funding (Lucas [2000], Oliveira [2014]). By the early 1900s it was nearly impossible to find work without being a speaker of English (Schütz [1994]).

There are well-documented accounts of students being punished for speaking 'ōlelo Hawai'i on school grounds (Kauhi [2000], Kamanā and Wilson [2012], Wilson and Kamanā [2006], Lucas [2000], Woodside et al. [2011]). Students were typically punished by being hit, kept after school, or forced to stand in corners while holding heavy stones, and were subjected to psychological ridicule. Further, students were intimidated by teachers to stop using Hawaiian and to use English at home. Teachers that were native speakers of 'ōlelo Hawai'i were also threatened with dismissal for using Hawaiian in school" (Kimura [1983]).

As the United States extended its colonial control over Hawai'i, the expansion of sugar plantations, fueled by American business interests, reinforced the dominance of English in the islands and prompted large-scale labor migration from Asia and other Pacific Islands (Roberts [1998], Roberts [2000]). These plantation contexts created the conditions under which the creole language Pidgin emerged (Sakoda and Siegel [2003], Siegel [2008]).

Pidgin initially developed as Pidgin Hawaiian, a Hawaiian-based pidgin (Roberts [1998]). Over time, the language *creolized*, defined as the process of nativization of a pidgin, whereby a previously non-native contact variety undergoes systematic structural development, grammatical stabilization, and functional expansion as it becomes the primary language of a community through intergenerational transmission (Bickerton [1984]). Alongside this creolization process in Hawai'i, 'ōlelo Hawai'i speakers community experienced increased contact with the English language; thus English increasingly became the dominant lexifier for the creole language, ultimately forming what is now known as modern Hawai'i Creole, locally referred to as Pidgin.

Although English became the dominant source of vocabulary, Pidgin re-

tained strong structural ties to 'ōlelo Hawai'i, but also became influenced by more newly introduced languages including Japanese, Portuguese, Cantonese, Ilocano, Tagalog, Korean, Sāmoan, Spanish, and others (Sakoda and Siegel [2003], Siegel [2008], Roberts [1998], Roberts [2000], Roberts [2005]).

In this sense, 'ōlelo Hawai'i functions as the *primary substrate language*, meaning it contributed many of the underlying structural features of Pidgin, even as English serves as the main lexical source (e.g., Bickerton [1984], Roberts [1998], Roberts [2000], Roberts [2005]). This linguistic connection is expanded upon in the following section (Linguistic Description of 'ōlelo Hawai'i, Pidgin, and Hawai'i English).

Today, Hawai'i exhibits an especially rich linguistic landscape relative to its small population. Over 80% of the local population reports speaking a language other than English in the household, including Ilocano, Tagalog, Japanese, Mandarin, Cantonese, Spanish, Korean, Vietnamese, Sāmoan, Chuukese, Marshallese, Visayan, and Tongan, among others. Reflecting this linguistic diversity, Hawai'i has been reported to have the highest proportion of residents who speak languages other than English of any U.S. state, including both those who speak only a non-English language and those who speak a non-English language in addition to speaking English (Hawai'i Leads Country [2016]). Within this multilingual ecology, Pidgin occupies a prominent position as a majority language, with an estimated 600,000 speakers in Hawai'i, or roughly half of the state's population (Siegel [2008]). Alongside Pidgin, 'ōlelo Hawai'i is estimated to be spoken by approximately 18,610 people, based on census data from 2009–2013 (Hawai'i State Data Center Department of Business, Economic Development and Tourism [2016]).

Despite being the majority language in Hawai'i, Pidgin is widely stigmatized, with negative attitudes such as “rude, low-class, and unintelligent” commonly associated with speakers (Drager [2012], Marlow and Giles [2010b]). Official attempts to ban Pidgin in classrooms occurred as recently as the late 1990s (Higgins [2010], Lockwood and Saft [2016], Tamura [1996]). Today, teachers and schools commonly discourage or punish its use (Romaine [1999], Hargrove and Sakoda [1999]), framing the language as a barrier to the “proper” acquisition of English (Yokota [2008]). Employers also have legal precedent to prevent employment on the basis of Pidgin-accentedness (Kahakua et al. v. Hallgren [1987]), which remains current and unchallenged to this day. Evaluations of Pidgin as “broken English” and a “liability in the job market” persist (Sato [1985]) into the present, reflecting the widespread assumption that discouraging children from speaking Pidgin at home is necessary for their future economic success (Sato [1985]).

While the scale and historical trajectories of 'ōlelo Hawai'i and Pidgin

differ, the two languages share commonalities in how they have been socially marginalized and ideologically positioned in relation to English. Both languages exist within a context of language diglossia in Hawai‘i, where English has historically been imposed as the prestige language to support colonial and economic hierarchies, often at the expense of Kānaka Maoli. Today, most Kānaka Maoli are speakers of Pidgin and most speakers of ‘ōlelo Hawai‘i are Kānaka Maoli (Sakoda and Siegel [2003]). Thus, the impact of restrictions of use and minoritization of the languages across particular domains especially affects Kānaka Maoli in negative ways.

Community activism has played a crucial role in elevating the status of both Pidgin and ‘ōlelo Hawai‘i. The Hawaiian language revitalization movement has taken significant steps toward both the “(re)teaching of the language and the (re)normalization of Hawaiian within everyday social life” (Saft [2022]). These efforts include the establishment of Hawaiian immersion education, including notably the ‘Aha Pūnana Leo preschools in 1983 (Wilson and Kamanā [2001]), and Ke Kula o Nāwahīokalani‘opu‘u in 1994 (the largest Hawaiian immersion school on the island of Hawai‘i) offering K-12 (Wilson and Kamanā [2007]). The latter includes a program called Hui Kīpaepae (family enrichment) which brings families of students in the school together once a week for cultural events that include Hawaiian language lessons for families of students attending the school. In addition, Ka Haka ‘Ula o Ke‘elikōlani at the University of Hawai‘i at Hilo offers the ability to receive bachelors to PhD degrees in ‘ōlelo Hawai‘i. Due to such efforts it is possible now for a child to receive an education in ‘ōlelo Hawai‘i from preschool to PhD (Silva [2017]).

Other efforts outside the domain of formal education include media, literature, and entertainment. In television, ‘Ōiwi TV broadcasts some English-language programming but produces substantial content exclusively in ‘ōlelo Hawai‘i. More recently, a new television series, *Chief of War*, stars Kānaka Maoli actor Jason Momoa and is produced by Kānaka Maoli filmmaker Thomas Pa‘a Sibbett, with dialogue presented exclusively in ‘ōlelo Hawai‘i. In literature, major works such as *Harry Potter*, *Alice in Wonderland*, and *The Hobbit* have been translated into ‘ōlelo Hawai‘i by authors including Keao NeSmith. Hawaiian language media also includes radio programming, such as *Ka Leo Hawai‘i* (1972 to 1988) and its later revival (1991 to 2000), which featured interviews with L1 Hawaiian speakers conducted in ‘ōlelo Hawai‘i. Contemporary radio continues this work through stations such as KTUH, affiliated with the University of Hawai‘i, where hosts include Paige Okamura, known as DJ Mermaid, and Ha‘alilo Solomon offer music and commentary in ‘ōlelo Hawai‘i.

1.3. Linguistic Description of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English

Grassroots activism has similarly led to the promotion of Pidgin, most notably through the nonprofit organization Da Pidgin Coup (Da Pidgin Coup [2008], Sakoda and Siegel [2003], Saft [2022], Higgins [2010], Saft [2019]), affiliated with the University of Hawai‘i at Mānoa. Da Pidgin Coup is a collective of community members, university faculty, and students who meet biweekly with the goal of addressing language discrimination through public outreach and engagement with research on Pidgin. Leaders in the Pidgin advocacy movement, such as Kent Sakoda, promote the recognition of Pidgin as a legitimate language both through teaching university-level courses on Pidgin linguistics and maintaining consistent use of Pidgin in academic contexts. More broadly, activists work to reduce discrimination and negative attitudes toward Pidgin by demanding space for its use in various domains of society, especially in education.

Other forms of activism include the promotion of Pidgin in media and the arts. A leader in the movement is Lee Tonouchi, also known as “Da Pidgin Guerrilla,” who advocates for the use of Pidgin in creative and academic writing, as well as in television and theater. Tonouchi argues that Pidgin use in creative and performance domains legitimizes the language through the re-inscription of cultural pride (Tonouchi [2021]), a claim echoed by Dr. Tammy Haili‘ōpua Baker’s analysis of Pidgin theater as both advocacy and cultural identity work (Baker [2020]).

As language activism movements gain momentum, the promotion of both Pidgin and ‘ōlelo Hawai‘i has served as political resistance to threats of United States assimilation (e.g., Saft [2022], Grama et al. [2022], (Saft [2022], Baker [2020], Silva [2017]). Furthermore, there is evidence that ‘ōlelo Hawai‘i and Pidgin are socially co-associated in opposition to English. For instance, qualitative research shows that the use of ‘ōlelo Hawai‘i and Pidgin can serve as a linguistic means of aligning with “local” Hawaiian values, signaling in-group belongingness (Sato [1991b], Marlow and Giles [2010a], Marlow and Giles [2008]), and are mutually positioned as markers of identity distinct from English.

1.3 Linguistic Description of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English

Creoles emerge through the interaction of superstrate and substrate languages, each contributing distinct linguistic features (Holm [2000]). The superstrate, typically the socially dominant language (such as English, French, or Portuguese), contributes most of the vocabulary to the creole. In contrast,

1.3. Linguistic Description of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English

the substrates, which are languages spoken by subordinate populations such as enslaved or indentured laborers, shape the phonology, semantics, and syntax of the emerging language. This creates a dynamic interaction in which the creole reflects the surface words of the superstrate but follows structural patterns influenced by the substrates.

In the case of Pidgin, Hawai‘i English serves as the superstrate while ‘ōlelo Hawai‘i serves as the primary substrate language. Thus, Pidgin shares more linguistic overlap with Hawai‘i English than with any of its substrate languages (Sato [1993]). Pidgin and Hawai‘i English also exhibit a nuanced social distinction. Pidgin is often labeled as “broken English” or, in surveys asking how many languages people speak, is sometimes identified as English (e.g. Bickerton and Odo [1976]). Nevertheless, Pidgin speakers demonstrate awareness of the distinction between the two varieties and can switch between English and Pidgin when prompted (e.g., Labov [1971]).

Hawai‘i English is the primary lexifier of Pidgin and shares overlap in core syntactic patterns, as well as in aspects of phonological organization and phonetic realization (Sato [1993]). The relationship between two languages has been characterized as a creole continuum (DeCamp [1971]), with Hawai‘i English as the acrolect (least creole-like) and Pidgin as the basilect (most creole-like) and the mesolect representing variation in between the two (Sato [1993]). Accordingly, rather than forming discretely distinct varieties, Pidgin and Hawai‘i English are best understood as points along a continuum, with intermediate varieties showing varying degrees of ‘Pidgin-like’ and ‘English-like’ features.

Note that throughout this dissertation, the term *Hawai‘i English* is used to refer specifically to the local variety of English spoken in Hawai‘i. When the term *English* is used, it refers more broadly to the English language, including both regional varieties and standardized forms of the language. This broader usage encompasses both English as a linguistic system distinct from Pidgin and English as a socially constructed idealized standard code that is often positioned in contrast to Pidgin. In the Hawai‘i context, the term “proper English” is understood primarily as an ideological construct, often associated with Standard American English norms in institutional and educational settings. However, Hawai‘i English itself functions as the de facto spoken standard variety for many local speakers. Importantly, the term English is used in a general linguistic sense in this dissertation, including in community contexts and in participant recruitment, because the term Hawai‘i English it is not treated as a distinct named language in everyday usage. The term *English* more generally is also used in this general sense when referring to English as a lexical system or when defining English terminology,

1.3. Linguistic Description of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English

without further specification of variety. For example, in Mo‘olelo, English is used more generally, reflecting both colloquial usage and its more idealized, conceptualized form. In Experiments 1-2, when referring to English translation equivalents or dictionary definitions, the term English is used in this broader sense; however, when referring to the auditory variety produced by participants and therefore included in the analysis, Hawai‘i English is used, as the regional variety specified in the recordings may have linguistic implications. In Experiment 3, written stimuli are used, so English is again used in this more general sense.

The Pidgin Density Measure (PDM) was created to capture the linguistic differences between basilectal and mesolectal Pidgin (e.g., Sakoda and Siegel [2008]), based off the Dialect Density Measure (e.g. van Hofwegen and Wolfram [2010]) which quantifies dialect use. The PDM score represents the ratio of basilectal to acrolectal words based on the sum of morpho-syntactic features used by a speaker. Higher PDM scores have been shown to strongly predict basilectal phonological and phonetic patterns (Grama [2015]).

For instance, these sets of sentences represent four translation equivalents along the Pidgin basilect-acrolect continuum.²

Basilect: Da buta ste kaukau.

Mesolect1: Da pig ste eating.

Mesolect2: Da pig Ø eating.

Acrolect: The pig is eating.

Basilect: Dakine wen go make die dead.

Mesolect1: Dakine wen go die.

Mesolect2: Dakine died.

Acrolect: That thing died.

Across the Pidgin continuum, basilectal varieties tend to exhibit greater influence from substrate languages like ‘ōlelo Hawai‘i in both lexicon and syntactic structure, whereas acrolectal varieties more closely approximate the superstrate language, English.

‘ōlelo Hawai‘i: ‘Ono ka laulau.

Pidgin (Basilect): ‘Ono da laulau.

Pidgin (Mesolect): Da laulau is ‘ono.

Acrolect: The laulau is delicious.

²The data presented here are drawn from the author’s own linguistic intuitions as a speaker.

‘ōlelo Hawai‘i: Hū ka wela!
Pidgin (Basilect): Hū da hot!
Pidgin (Mesolect): Wow it hot!
Acrolect: Wow it’s hot!

While superstrate and substrate driven influences are evident across many levels of linguistic structure, I will now examine in further depth the phonetic and phonological patterns, as these constitute the primary focus of the central research question.

Hawai‘i English (HE) is described as differing from General American English (GAE) in several ways (e.g., Kirtley et al. [2016]). These differences include fully realized vowels where schwa might be used in other dialects (e.g., ‘today’ in HE: [tudei], in GAE: [tədeɪ]); stopping of the voiced interdental fricative /ð/ → /d/ (e.g., ‘bathe’ in HE: [beɪd], in GAE: [beɪð]); and vocalization of syllable-final /ɪ/ (e.g., ‘car’ in HE: [khaɪ], in GAE: [khaɪɪ]) (Sato [1993]). Kirtley et al. [2016] further describe Hawai‘i English vowels as showing systematic centralization of front vowels. Compared to General American English, long vowels in Hawai‘i English are produced with less peripheral vowel qualities, meaning that they tend to occupy more centralized positions within the vowel space (e.g., ‘hand’ in HE: [hɛnd], in GAE: [hænd]). Hawai‘i English diphthongs are also less dynamic, showing smaller changes in formant trajectories over the course of the vowel. As a result, diphthongs in Hawai‘i English may sound more monophthongal relative to their General American English counterparts, with less movement between the beginning and ending vowel qualities (e.g., ‘smoke’ in HE: [smo:k], in GAE: [smoʊk]). Overall, the attributes of Hawai‘i English likely reflect both substrate influence from Hawaiian and contact effects from other local languages, resulting in a distinct regional variety of English.

‘ōlelo Hawai‘i, as the primary substrate language of Pidgin, has linguistic influence across several structural dimensions (Sakoda and Siegel [2003]). These include shared characteristics at the syntactic level (e.g., Sakoda and Siegel [2003], Siegel [2008], Schütz et al. [2005]), in the lexicon (e.g., Sakoda and Siegel [2003], Reinecke [1969], Carr [2019]), as well as at the segmental (Grama [2015]) and suprasegmental levels (Murphy [2013]). Consequently, there is well-documented evidence of significant overlap between Pidgin and ‘ōlelo Hawai‘i.

The ‘ōlelo Hawai‘i phonological inventory consists of eight contrastive consonant segments, /p, k, ʔ, h, m, n, l, w/, and a vowel system comprising five short monophthongs /i, e, a, o, u/, five long monophthongs /i:, e:, a:, o:, u:/, nine short diphthongs /ae, ai, ao, au, ei, eu, iu, oi, ou/, and six long

diphthongs /a:e, a:i, a:o, a:u, e:i, o:u/. The segment /ʔ/ corresponds to the orthographic character <‘>, which in ‘ōlelo Hawai‘i is called ‘okina. Long vowels are represented orthographically with the macron symbol (e.g. <ī, ē, ā, ō, ū>) (e.g., Schütz [1994], Parker Jones [2018]).

Basilectal Pidgin is described as having a consonant inventory derived primarily from English and is influenced by substrate languages like ‘ōlelo Hawai‘i. The consonant inventory of Pidgin includes the following segments: /p, b, t, d, k, g, ʔ, m, n, ŋ, f, v, s, z, ʃ, tʃ, dʒ, ɹ, r, l, j, w, h/. The vowel system consists of monophthongs /i, u, e, o, ɔ, æ, a/, and three diphthongs /aἰ, ɔἰ, aṽ/ (Bickerton and Odo [1976], Grama [2015], Sakoda and Siegel [2008]). Basilectal Pidgin does not distinguish high lax vowels from high tense vowels, and thus does not distinguish between /i/ and /ɪ/, nor between /u/ and /ʊ/. Furthermore, /e/ and /o/ are typically realized as monophthongal (Sakoda and Siegel [2008]). Dental fricatives /θ, ð/ are typically replaced by /t, d/. Liquids /l, ɹ/ are often vocalized. Consonant clusters are often simplified, with final consonants sometimes deleted, particularly in syllable-final position. Additional sounds which occur in English appear variably in mesolectal Pidgin.

While a full comparative acoustic analysis has not yet been conducted, existing studies indicate that the Pidgin vowel system shows measurable influence from the Hawaiian substrate. Corpus analyses of Pidgin speech show that duration distinguishes vowel categories that overlap spectrally (Grama [2015]) (e.g., ‘beat’ [bi:t] and ‘beat’ [bit]). Similarly, ‘ōlelo Hawai‘i maintains a contrastive distinction between long and short vowels (e.g., ‘wāhine’ (woman.PL) [wa:hine] vs. ‘wahine’ (woman.SG) [wahine]). Thus, in both ‘ōlelo Hawai‘i and Pidgin, vowel duration is likely a more important linguistic cue than in English. There has also been limited research on the influence of ‘ōlelo Hawai‘i lexical items on Pidgin, suggesting that Pidgin speakers realize /a/ higher in the vowel space in words of Hawaiian origin than in words of English origin (Grama [2016]).

In addition to the well-established influence of ‘ōlelo Hawai‘i on Pidgin as its primary substrate language, the three languages examined in this dissertation are also presumed to exert mutual influence on one another, motivated by evidence of cross-language interaction in multilingual speakers (e.g., Grosjean [1989, 2001], Cook [2003]) as well as sustained contact at the community level (e.g., Thomason and Kaufman [1988], Winford [2003]). To my knowledge, this mutual influence has not been explicitly investigated. Nonetheless, descriptive accounts identify differences between Hawaiian spoken in contexts of English and Pidgin influence and non-contact ‘ōlelo Hawai‘i, identifying specific sound changes attributed to influence from English and especially

Pidgin (e.g., Nesmith [2002]). In this sense, Pidgin in particular has been theorized to exert influence on Hawaiian as spoken today outside of Ni‘ihau communities.

1.3.1 Research Context

Prior work has demonstrated a facilitating effect between two linguistically associated languages. Eye-tracking literature (e.g. Jacobs et al. [2016]), for instance, has found that code-switched English-to-Spanish sentences facilitated faster correct responses, and faster divergence in gaze proportions to the target item when the English speech preceding the switch incorporated Spanish-like phonetic cues. Fricke, Kroll and Dussias (2016) argue based on English-to-Spanish code-switched corpus data that the English phrase, anticipating the code-switch, naturally carries signatures of the upcoming switch to Spanish, creating a linguistic association between the two languages. This association appears to be utilized to cue listeners to an upcoming code-switch. Thus, due to the fact that Pidgin carries phonetic elements of ‘ōlelo Hawai‘i, we may expect trilinguals of Pidgin, English, and ‘ōlelo Hawai‘i to anticipate the switch to ‘ōlelo Hawai‘i when Pidgin precedes the switch, as opposed to when English precedes the switch. From a cue-weighting perspective, differences in facilitation may be explained by how knowledge of a second language may affect the process of mapping phonetic cues to phonological contrasts when comprehending speech (e.g. Lisker [1986], Holt and Lotto [2006], Francis et al. [2000], Chandrasekaran et al. [2010a]).

Bilingual listeners in particular have consistently demonstrated that individuals may rely on certain cues over others in a manner which is predictable based on the first language of the listener (e.g. voice onset time (VOT) vs onset f0 for voicing (Schertz et al. [2015]); formant frequency vs duration for vowels (Kim et al. [2018])). Due to the linguistic relationship between Pidgin and ‘ōlelo Hawai‘i, it may be argued that listeners with more Pidgin experience may be better adapted at distinguishing vowels using duration when the contrast is spectrally ambiguous, a skill which is required for the comprehension of ‘ōlelo Hawai‘i. Experience with English, a language which primarily uses spectral cues to distinguish between the same vowel pairs, is not expected to activate the same sensitivity to duration cues as we may expect for Pidgin.

Similarly, research has demonstrated a facilitating effect between languages that are socially linked. Szakay, Babel, and King (2016), for instance, found that Māori English (the variety of English associated with those of Māori descent) facilitated the priming of Māori lexical items over Pākehā

English (the variety of English associated with those of European descent) (Szakay et al. [2016]). While Māori English differs from Pākehā English in both phonetic and phonological features, the priming effects were interpreted as reflecting the activation of social category information rather than mere phonetic similarity. By analogy, due to the social and cultural relationship between Pidgin and ‘ōlelo Hawai‘i (e.g., Marlow and Giles [2010b]), we may similarly expect Pidgin to facilitate priming of ‘ōlelo Hawai‘i lexical items over English, consistent with a socially mediated effect rather than a purely phonetic or phonological one.

Moreover, additional language (L2+) acquisition theory suggests that the statistical knowledge used to learn the sound categorization and acoustic patterns of a first language (L1) may play a role in the learning of an L2+. When learning an L2+ that is similar to an L1 along some linguistic dimension, we might expect that the statistical regularities in the two languages would be similar enough to allow the learner to bootstrap their knowledge of one language to facilitate learning of the other (e.g. Escudero [2009], Ingram and Park [1997]).

This dissertation reports on three experiments that bring linguistically informed research questions to bear on the Hawaiian language space. Returning this work to the community, the ultimate goal is to use these results to inform decisions about language teaching and the use of Pidgin as a medium of instruction for ‘ōlelo Hawai‘i. Due to its status as a stigmatized language, as discussed above, the language is largely suppressed in the domain of education. As Pidgin is spoken as a first language by a substantial portion of the population, this places an undue burden on the community, particularly for Kānaka Maoli, who represent a large proportion of Pidgin speakers. Regardless of this stigma, ‘ōlelo Hawai‘i teachers have increasingly been adopting the practice of using Pidgin as a medium of instruction as opposed to English, often citing the cultural and linguistic connections between the two languages (e.g. Martin [2018], Cabral [2018], Rozet [2017b], Rozet [2018]).

By empirically examining whether shared linguistic dimensions between Pidgin and ‘ōlelo Hawai‘i are leveraged in speech processing, this dissertation seeks to contribute evidence that can help ground these pedagogical practices in psycholinguistic research. In opposition to the stigma Pidgin is associated with, this work reframes it as a potential benefit that learners and teachers already draw upon in the Hawaiian language classroom. In doing so, the dissertation aligns with community driven efforts to challenge linguistic stigma and supports approaches to ‘ōlelo Hawai‘i education that consider the lived linguistic realities of Kānaka Maoli and kama‘āina.

1.4 Outline of Dissertation

Chapter 2 presents mo'olelo from conducting this work in community, reflecting on research experiences, methodological choices, and the role of pilina in the research process. Chapter 3 examines 'ōlelo Hawai'i word recognition in the context of code-switching to test whether participants recall 'ōlelo Hawai'i lexical items more effectively when preceded by Pidgin rather than English. Chapter 4 employs a translation priming task to assess whether Pidgin and 'ōlelo Hawai'i are perceptually co-activated. Finally, Chapter 5 investigates whether Pidgin speakers demonstrate implicit knowledge of 'ōlelo Hawai'i phonotactics in a well formedness rating and word identification task, offering an initial step toward understanding how experience with Pidgin may support aspects of 'ōlelo Hawai'i L2+ acquisition.

Some degree of repetition in the contextualization and framing of the research questions throughout this thesis is intentional. Each chapter is designed to function as a self-contained unit, providing sufficient theoretical, social, and historical grounding for readers approaching the work from different entry points. At the same time, these chapters are unified by a shared through-line that repeatedly returns to a central question concerning the connections between Pidgin and 'ōlelo Hawai'i, both in community contexts and in the psycholinguistic literature. By returning to this framing across chapters, each empirical study is placed within a broader argument, making clear how distinct methodologies and research paradigms work together to build a cumulative account of the psycholinguistic relationship between Pidgin and 'ōlelo Hawai'i.

Chapter 2

Mo‘olelo, “talk story”, from work done in-community

2.1 Introduction

Both traditional and contemporary Kānaka Maoli scholars have documented mo‘olelo (story or story-telling) as a form of identifying people, events and the story-tellers’ mana‘o and ‘ike ku‘una. In Pidgin, mo‘olelo can be roughly translated as “talk story” and functions as both a noun and a verb. However, the term conveys a more casual sense of storytelling or sharing stories than the original Hawaiian term. In ‘ōiwi epistemology, the usage of mo‘olelo allows the scholar to not only explain but also to reproduce the lāhui. In my education in scientific research best practices, I was trained to be wary of subjectivity and taught to avoid emotional investment, maintaining distance from the work in order to strive for objectivity and eliminate bias. However, I intentionally employ storytelling as a methodology to interrogate claims of true *objectivity* in empirical studies. Regardless as to whether we, as researchers, do work which is quantitative or qualitative, there is always an element of subjectivity and frame of reference based on our personal positions in relation to communities. According to ‘ōiwi practice, mo‘olelo provides a framework through which this interrogation can take place.

In this chapter, I include a series of mo‘olelo, stories in the form of short vignettes, from doing this research in-community. These mo‘olelo are from my point of view, and are composed of experiences I have as a person carrying several layers of positionality.

Each subsection represents community engagement in different ways: first with ‘ohana and hoa, second with kumu and haumāna, third with the lāhui in Hawai‘i and abroad.

2.2 ‘Ohana and Hoa

2.2.1 Mo‘olelo I: Bridging the sacred space between ‘ike ku‘una and the present

“Our kūpuna join us in our work, whether we recognize them or not.” - Noelani Goodyear-Ka‘ōpua

My great-grandfather Tadashi Yoshida was trilingual, a speaker of Hawaiian, Pidgin and Japanese. He was a fisherman. I have fond memories of sitting with him outside in my grandmother’s garden, watching him slice a marlin on a makeshift desk, barely held together, as the marlin’s blood ran down and mixed with the dirt. He would let me have a bite of the raw fish, feeding it directly into my mouth without saying a word. When he was finished, he left the fish spread out to dry so we could eat it as jerky. I remember thinking it was the best thing I had ever eaten. It probably was. To me, Kona (a district located on the western, leeward coast of Hawai‘i Island) has always carried a faint smell of marlin. Kona smells different from the city, Honolulu, where I am from.

Great-grandpa was an extremely quiet man, yet oddly social and had many friends. Most of his friends happened to be kanaka (Pidgin word for Kānaka Maoli), because of his profession and where he was raised. While not kanaka himself, he learned to speak Hawaiian to socialize with his friends and to better develop his knowledge surrounding his craft as a fisherman. He would never have admitted it, but he was so well known as a fisherman that he was featured on a popular local television show in Hawai‘i at the time, called Let’s Go Fishing! (1969–1998).

When I am asked why I decided to pursue a PhD, the true story begins with my great-grandfather. He was well known for his many talents, but among those from out of town and new haole that moved to Kona, his worth was often diminished. I believe this was largely because he was illiterate. His mother died when he was very young, and he left school in the second grade to work in order to support his siblings. The absence of formal education did not diminish his value to his community. He learned languages not to display knowledge or seek praise, but to better support those around him through his craft and to show his love for others.

My family’s survival was inseparable from their bonds with Kānaka Maoli and their ties to the ‘āina. Their pilina ensured sustenance. Through these relationships, my ‘ohana learned Hawaiian ways of fishing, catching chickens, and hunting wild boar, and of cooking them in an imu. Learning to mālama the ‘āina involved relational practices of land stewardship, including planting

fruit trees, caring for the lo‘i, understanding how rocks and soils affected harvests, and observing proper protocols of respect toward Pele, the goddess of the volcano, especially during seasons when her eruptions were impending.

In life and in friendship, relying on each other on the plantations was all we had. For my family, Hawai‘i could be seen as a physical *pu‘uhonua* (place of refuge), leaving Hiroshima for Hawai‘i long before the Second World War. I think about how little my family had, but how they never felt poor because of the relationships they shared with others, how mutual generosity and care helped everyone survive, and what a profound blessing that was. I am the first generation in my maternal lineage to grow up in a house with plumbing and electricity, with floors made of carpet rather than dirt, without the embodied experience of plantation labor, and living with an abundance of food I can just buy at a 24 hour 7-11. I remember my grandmother, even in her older age, coming home with dirt-stained overalls and bumpy calloused hands from years of hard work, while I spent the whole day running around her yard in a frilly pink ballet dress, chasing bugs and lizards and stomping on crunchy lava pebbles without a care in the world.

Despite knowing the hardship and sacrifice that shaped my upbringing of privilege, growing up in Honolulu with more resources than anyone in my family’s history, I grew up deeply embarrassed by the belief that I came from “not a smart family.” My first memories of this shame was before I ever went to school. My father, white and from California, often felt the need to correct the way my mother spoke. “It’s not shut da light, it is turn off the light.” A story I heard many times growing up, and one that still makes my family laugh decades later, comes from the first time my dad’s family met my mom’s family at my parents’ wedding. At one point, someone from my dad’s side exclaimed, “Stop speaking Japanese in front of us. That’s rude!” My mom’s side of the family found this absolutely hilarious because, as they would always say, “So lolo, yeah? We don’t even speak Japanese. We was speaking ENGLISH!” They were speaking Pidgin. There was always a feeling that, with certain people, we had to speak in a particular, non-Pidgin way just to be treated as normal, complex people with thoughts and ideas. Unfortunately, that feeling sometimes came from within my own family.

My earliest memories of school include teachers criticizing my accent when I read aloud. The way I spoke marked me as coming from a family of Pidgin speakers, and I internalized this as shame. Much of my huaka‘i in formal education was spent chasing external validation through academic accolades, a pursuit that meant running away from Hawai‘i and from my roots in an attempt to establish my worth. Ironically, it was in pursuit of graduate education where I found myself again, by reclaiming Pidgin and

reorienting myself as a person raised by people like my great-grandfather with the ethic that prioritizes pilina. I now see this as a strength in my work as a scholar. When I faced hardship conducting this research in community, I often returned to a simple question: what is the point? For me, the answer returns to pilina. In conversations with community members who expressed shame about speaking Pidgin, or who felt that learning Hawaiian was too difficult, I heard deep insecurities about worth being tied to language ability. In those moments, I returned to the lessons I learned from my great-grandfather. I shared his story and posed a question of my own: why did any of our ancestors learn language in the first place? For languages such as Hawaiian and Pidgin, our kūpuna did not learn or speak them for ego, status, or monetary gain. In fact, it was often the opposite. Speaking these languages was frequently devalued, sometimes seen as less prestigious than being monolingual in English. For generations, my family was mocked and excluded from positions of prestige, and a big part of this was due to language. Yet our kūpuna continued to speak these languages, whether as first speakers or as learners, in order to build relationships. Speaking Hawaiian, Pidgin, or any other language was an act of care in community. It was about making friends, learning from friends, and meeting others with humility. It was about being on the same level as those around you, across many communities, allowing everyone the space to speak their own languages. It was about catching five fat ‘ahi and eating together in good company.

Although my great-grandfather has long passed into the ao ‘aumakua, today I carry his DNA, his ethic of practice, and his personal commitment to the community into this work. With this work in gaining ‘ike from other places, I weave in the ‘ike ku‘una I have inherited. While my great-grandfather did not get the chance to go past the second grade, together we have been in school for a long time.

2.2.2 Mo‘olelo II: Kapu and Kuleana

As Morris Young puts it, there is an “anxiety” among locals due to the uncertainty around “what claims on citizenship they may have” (Young [2004]).

One day, I was talking to my mom about some things I learned about Pidgin during my masters program. I thought she would be excited but instead she said to me, “don’t you think that YOU talking about Pidgin is annoying? You da least *local* there is.” There are many layers to what “not too local” means in Hawai‘i, but when put in this content the question is about kapu. Knowing the context, she was referring to the kapu surrounding the use of Pidgin in Hawai‘i. As a half white city girl, I grew up with the

option to speak Pidgin or not, and throughout my upbringing I often chose not to. With the ability to speak both languages, I was able to choose when to use Pidgin, avoiding much of the discrimination that Pidgin monolinguals face. This was a privilege not afforded to many people in Hawai‘i whose maternal family backgrounds were similar to mine.

While my mom had a point, of course I did not like what she said. It felt especially unfair because, at the time, I was the only person with personal ties to Pidgin at all in my cohort at the Linguistics Department at the University of Hawai‘i. Indeed there were others that studied Pidgin, but I found myself repeatedly confronted with interactions with the larger student body that felt condescending and entitled. Haole classmates would ask me, again and again throughout the day, at lunch or at the bar while drinking beer, for “grammaticality judgments,” as though access to the language was something I owed them. Others demanded that we “collaborate,” while not having been bothered to learn about Hawaiian history nor the social implications surrounding speaking Pidgin or ‘ōlelo Hawai‘i, treating the language merely as a puzzle to be solved. I have witnessed trusted colleagues treat the language as a spectacle, using it to display to friends back home the strange and exotic things they were learning in Hawai‘i, recounting to me that their friends said “it sort of sounded like caveman speak”. I have never once been given the impression that any of the others that were not from the community at the time have ever thought about being the “right person” to do the work. Yet, despite all of this, I was being confronted with the question of whether this work was kapu for me to do.

It made me mad because it was right. I am not the most local. If we are being honest, “da least local there is” is probably a pretty accurate description. I do understand that we carry added kuleana when we claim connection to a place and community. While we may know ourselves to be part of a community and speak from our own perspectives, claiming group identity is a powerful assertion that is, and perhaps should be, continually questioned. It is certainly the case that working our way into academic spaces often means becoming among the most privileged members of our communities, particularly within communities that have experienced marginalization. Further, these realities are not inherently oppositional. The same power structures I critique in this thesis are also ones that benefit me relative to others in my community, including my mother. For many people, having one foot in academia and another in the community is taken to mean that you are at least partially part of the problem. Navigating these issues involves the difficult task of defining one’s identity when identity itself shifts through experiences such as entering a graduate program.

Although this initial discussion focused on Pidgin, it prompted me to consider how framing a research question around Pidgin in its interaction with Hawaiian broadens the question of kapu as it relates to the Hawaiian language. Although many of us grew up learning basic Hawaiian in school, along with our parents and grandparents for several generations back, some do hold the perspective that the language is “not for everyone”. This understanding emerges from a long history in which Hawaiian has been commodified by haole for profit, while Kānaka Maoli have been exploited by haole, whether white or otherwise. I do not want others to assume that I am Kānaka Maoli when I use Hawaiian in my interactions. This led me to question whether speaking or studying these languages is kapu to me.

From lived experience, I have come to the conclusion that, at least in Hawai‘i, there is nothing worse than claiming local identity for personal gain. There is a long running joke that tourists will try to get kama‘āina discount at local businesses but cannot pronounce the word correctly (“Can I please have the Kamala Harris discount?”). There is a long history of outsiders using local identity to exploit Hawaiians. Speaking Pidgin “like a haole” is seen as repulsive to many people, a sly attempt for outsiders to enter spaces meant for locals-only. Many of these issues revolve around what could be understood as cultural appropriation, making the act itself socially gray at best. Even for members of the community, speaking Pidgin or ‘ōlelo Hawai‘i can be viewed as a political act of claiming belonging to identity. Speaking one of the languages is then sometimes treated as an act of deception if you are not perceived as belonging to those communities.

While such questioning of group identity and kapu is understandable, given the misuse of claims to local identity by outsiders, it can nonetheless be painful for community-member researchers to repeatedly confront questions of belonging, sometimes even from their own family members.

There are several options in which a researcher in this kind of position may choose to respond. One option is to speak English and reserve Pidgin or ‘ōlelo Hawai‘i for situations where mutual comfort is established, thereby avoiding the use of these languages in ways that might feel like a forced display of belonging.

This approach, however, is not universally understood as pono. In my work in-community, I have encountered perspectives that are directly oppositional, framing the withholding of language and self-identification not as ethical restraint from that which is kapu, but as a failure of kuleana. When visiting Ka Haka ‘Ula O Ke‘elikōlani in 2024, I was met with the perspective that if you were not actively perpetuating ‘ōlelo Hawai‘i and Pidgin by speaking it as much as possible in your daily life, then you are doing a disservice to

the movement and to Hawai‘i. One particular person, a haole individual who holds significant power and leadership within the community, went so far as to tell me that I was a “fraud” for not speaking Hawaiian or Pidgin with him. “Your mom will die one day and then who will you speak the language with if you refuse to speak it,” this person told me. This interaction affected me deeply; I lost sleep for a long time and contemplated withdrawing from the work entirely. Yet, I was provided with a valuable perspective, “if you are a part of the community, and are benefiting from it, it is your kuleana to learn, speak, and use the languages you study.” This perspective certainly is one that carries mana (strength) in the community at Hilo, leading to much more Hawaiian heard out in public than any other comparatively large city in Hawai‘i. This perspective has directly shaped the prevalence of Hawaiian being heard in public in Hilo, a result that may, in fact, be central to the aims of any language revitalization research.

Reexamining the question of kapu, I began having more conversations and doing some reading about this myself. As portrayed in her own mo‘olelo featured in the book *Kanaka ‘Ōiwi Methodologies: Mo‘olelo and Metaphor*, Summer Puanani Maunakea retells a conversation she had with Dr. Jonathan Kamakawiwo‘ole Osorio on the context of exploitative academia, “If you do not have a kuleana to something, you should not be going there. If you have the kuleana it means being from a community in which you have a stake that everyone else recognizes; it is not something you claim... It does not need to be about koko (blood). It has to be about: does the community recognize you? If they do then you have a kuleana” (Oliveira [2015]).

When visiting Ka Haka ‘Ula O Ke‘elikōlani again in 2025, United States President Trump declared English as the “official language” of the country, undermining the status of Hawaiian as one of the official languages of the state of Hawai‘i. During a panel, a Hawaiian language kumu stated in response to this policy being passed, “it is our kuleana now more than ever for all subjects of the occupied kingdom to learn Hawaiian in the face of budget cuts and impeding hurdles on the language.”

Throughout my time in doing this research, especially in doing this work publicly, I have taken every experience that I have encountered to reevaluate my kuleana. Kuleana comes in shapes and sizes and forms, and it does not always need to regard to language. My kuleana exists in this position of in-between. While there are several reasons why I am not the best person to do this work, in every sense, both directly and through means of others, I am personally impacted by wrongs done on the community. In my training with linguistics, my kuleana is to ensure my research is pono and properly assesses its limitations. This is why, when I describe who I am, I tell you also

my community, where I am from, alongside where my ancestors are from. In certain spaces in Hawai‘i, my name is enough to know exactly where I am from. I carry this with me to describe the body of people I represent in this work and why it is necessary to do it right by the community.

In 2025, I attended a workshop in Hilo that addressed the kuleana of lāhui members from various backgrounds to learn and use language. One kumu stated, “I tell my students that their personal kuleana is their identity, including who they are, where they are from, their genealogy. If they feel called to kuleana, act on that kuleana, but clearly state who you are. Your committed effort in relation to your identity can pave the path to inspiring others who share aspects of yourself, extending influence outward like a web. Whether we want to or not, every decision we make is itself an action. Shying away is also an action, and who then does that serve?”

Throughout these experiences with kapu and kuleana, I have come to the conclusion that the answer simply may lie in sitting with the discomfort of continuously reevaluating and explicitly stating one’s ever-changing positionality. Further, the burden of questioning kapu and kuleana should not solely lie on community members. Every community is different, and harm can take many forms. I argue that identifying one’s own kapu and kuleana is an important research protocol that all researchers must actively engage with, regardless of belonging to a particular community or not.

2.2.3 **Mo‘olelo III: Pōhaku at the Pu‘uhonua O Hōnaunau**

The Pu‘uhonua O Hōnaunau “city of refuge” is a sacred area of ocean and land that held the purpose traditionally as a space where people that broke a kapu (here, meaning “law”) could find refuge. If reaching this enclosed space, offenders and prisoners of war could avoid the death penalty, as killing someone on this sacred land was forbidden. My mother and grandparents and great grandparents were born and raised a 10 minute drive from this place of refuge. It is one of the only federally protected areas under the authority of the United States government that does not have an American flag and displays only a Hawaiian flag.

In addition to this historic significance, this place was sacred to my childhood. As a child, I remember going there to learn how to fish and slap um on a pan later for dinner, pluck tsubu and ‘opihi off the rocks to eat as a snack. My sister and I would collect fully preserved bright red crab exoskeletons off the beach to “scare” our elders (they put on a good show for our sake), and search for hours for the best puka shells to add to our growing collection of necklaces, hung at my grandma’s house alongside ancient dried tī leaf leis

that were never thrown away after countless past occasions. During the day the sand was often covered with basking sea turtles, and at night the shore will come alive with night mollusks and hermit crabs. There are eel in the water you need to swim away from, and wana on the seafloor that got stuck in my foot when I was a teenager which stayed in my foot for five years (eek). The ocean water appears vibrant cobalt, sometimes even lavender, against the black lava floor. Schools of butterflyfish appear as glowing orbs of yellow in the water. A tide pool connected to a wai source through a narrow opening creates a suction that swirls swimmers through a deep pool, functioning as a free “natural water slide”. There is a photo of me sitting on a bench there on my first birthday, wearing a stupid birthday hat and sporting the same haircut as my grandma, great-grandma, and great-aunt. It was also the place where my mother, her mother, and her mother’s mother celebrated their first birthdays. My mom would always remind me, “Back when we was kids, we never had fancy things like toys. We wen blow up da grasshoppahs as our toy.” Ok, Mom.

There is a special pōhaku at the Pu‘uhonua O Hōnaunau where in ancient times, if you reached, it meant you officially have reached the city of refuge and were safe from harm. When I went back in 2025 to recruit participants for my research, my mom pointed out the pōhaku, which was not roped off or marked as kapu (here, forbidden) at all. It was strange. There was a very small sign with a number next to the rock, which you could use with a legend on a map to discover that this was indeed the location of a place with such profound historical and cultural significance.

Standing beside the sacred pōhaku, a tourist’s sunburned bare butt entered my peripheral vision as they tossed their goggles toward the rock, treating it like any ordinary ledge for their belongings. I sat beside the pōhaku and put my arm around it feeling the need to protect it. As I sat hugging the pōhaku, I thought about the people who once reached the stone to escape punishment in ancient times. What the relief it must have felt to finally touch this rock. I reflected on how much of a relief it was for me to return to Hawai‘i on this trip, after many years of trying to abandon where I came from where I took this place for granted. As I sat, I also reflected on my relation to ‘āina, my pilina with ‘āina as *hoa*. I sat and thought these things as if I was praying to the pōhaku, asking for permission to continue doing this work.

Days later a friend was showing me a deck of cards she was gifted called the “mana cards” that includes a companion book (Becker [1998]). I am not entirely sure about the backstory of the cards, nor the source of the information on Hawaiian epistemology, but you were meant to pull a card, to

2.2. *‘Ohana and Hoa*

represent “the culmination of the past, present, and future” sort of like tarot cards, and coincidentally, I pulled the “pōhaku” card, which read “pōhaku (stones), representing relationships”. “Stones represent the gods, earth, and relationships. They symbolize the close relationships that Hawaiians have with the earth and with each other. Stones can be used to represent a relationships with one who is either alive or passed on, with one’s ‘aumākua or with one’s god. Stones contain mana and should be treated with respect.” An ‘ōlelo no‘eau was quoted on the card, “He ola ka pōhaku a he make ka pōhaku.” (There is life in the stone and death in the stone). A proverb carrying the meaning that the pōhaku, representing relationships, must be nurtured as they are living things that can die if they are not cared for.



Figure 2.1: The Pōhaku card from the Mana Card Deck (Becker [1998]). Pōhaku, meaning stone, evokes our connection to the source, to others, and to our communities, offering strength, inspiration, and purpose. Photo taken by Katrina Karl.

“Pōhaku reminds us that we are not alone, but are connected to the source and to others. Through our relationships, we develop a sense of who we are, where we came from, and where we are going. This connection can be a source of incredible inspiration and strength.” It goes on to describe how your relationships could be familial, but also in any connection based in shared “spiritual, environmental, political, economic, or artistic concerns.” When nurturing any community, it says, it “takes effort and can be painful”, but connecting to community in relationship to others and the land, “you will find a renewed sense of inspiration and purpose.” “As you relate to others,

pay close attention to what is needed. You have an important role to play and a contribution to make that will allow you to strengthen your connection to the source.”

In reflecting on what it means to do work in-community, it occurred to me that this must also include attending to the perspectives of the physical ‘āina, including the act of asking permission. In sitting with the pōhaku at Pu‘uhonua O Hōnaunau, I felt a deep resonance with what Dr. Uahihea Maile describes in *Becoming Hoa* (Maile [2024]). My relationship with the pōhaku does not feel like being of the stone or belonging to the long history of its meaning, but rather like a long-standing family friendship, shaped by the mo‘olelo the stone tells and by my personal and familial pilina with the stone as *hoa*.

In thinking through questions of pilina, ‘āina, kuleana, and the practice of asking permission in different ways to different sources, I spoke with Kumu Larry Kimura, often described as “the grandfather of the revitalization movement.” I asked him directly, “Kumu Larry, if you thought I was doing something wrong, would you tell me?” He responded simply, “You said you are doing this for your family and friendship. That is why I am doing this too.” I thought about this conversation again when looking at the pōhaku card, as I had so many moments of wrestling with whether I was the right person or whether I was good enough to do this work, a process I would describe as painful. This part, written on the pōhaku card, especially resonated with me: “It takes effort and can be painful, but you will find a renewed sense of inspiration and purpose... As you relate to others, pay close attention to what is needed. You have an important role to play and a contribution to make that will allow you to strengthen your connection to the source.”

Of course, the wisdom of the pōhaku, formed through its long presence across generations, extends beyond what any human or book could have offered me, and beyond the scope of this research. It did, however, offer me friendship throughout my life and, again in 2025, reassured me that the lesson was not simply in the work itself, but in being in pilina. It was a meaningful lesson for me at a time when I needed it most.

2.2.4 Mo‘olelo IV: Ka‘aloha’s Super J’s Laulau

Ka‘aloha’s Super J’s Laulau is a family-owned restaurant that serves traditional Hawaiian food, located in Hōnaunau. My mom and I stopped there to eat lunch and see long-time family friends, the Ka‘aloas. When we opened the door to the shop, Auntie J., the owner, sat behind the counter as if she

had been expecting us, shaking her head, smiling with tears in her eyes. My mom joked, “Oh, you remember me?” Auntie J. stood up to hug her and said, “I had a dream two nights ago that you were here. We were kids again, and our fathers were watching us. I woke up thinking, weren’t our fathers just such good friends? And now you’re here.”

I sat back like a fly on the wall while they “talked story”, their Pidgin slipped into an older and older dialect as they recalled the past. When I was a kid, my grandfather worked as a water delivery man, going from house to house to deliver water to homes because there was no plumbing. Everyone knew he loved his big red truck, and the way people talked about him, it almost felt like the truck was literally an extension of his body. “Eh, I wen see you the oddah day, I see da red!” (“I saw you the other day. I saw the red (truck)”). There are photos of him plopping me on his truck as if they were his shoulders. Auntie J. talked about how funny it was for her father, who was a big Hawaiian man, to see how my grandfather, who was of Japanese descent and exactly 5’2, could operate such a huge truck, using everything he could to make it work for him. She recalled that my grandfather kept one of those coffee picker sticks in his truck so he could manually adjust one of the back mirrors, concentrating so hard that he would stick his tongue out and close one eye whenever he needed to reverse the truck to park, otherwise he would have been too short to see. My mom recalled how Auntie J.’s father taught her dad how to bake a Hawaiian-style coconut cake, and Auntie J. responded, “And then your father would make it for every occasion for us. Do you remember that time Dad got sick and your father brought a cake every day and let us stay with you folks for months. Such good friends they was.”

When we were done eating, she refused to let us pay, and they went back and forth until my mother hid the money in Auntie J.’s pocket. “Dass just our way,” (“That’s just our way”). Auntie J. said to me, acknowledging that maybe I would find this foreign, being from Honolulu.

Academic work on Pidgin sometimes describes, for the sake of simplicity, that the language developed on the plantations as a lingua franca that allowed various ethnic groups to speak to one another. Sitting there listening to Auntie J. and my mom talk story, I realized that Pidgin was more than a common language; it was kind of more like a love language carrying friendship, laulau, and coconut cake.

2.2.5 Mo'olelo V: Akamai Kid

In high school, my mom wrote an anonymous column published in her high school newspaper called *Akamai Kid* (The smart kid), where she portrayed a gregarious, sporty, slightly dim-witted alter-ego, entirely in Pidgin. At her high school graduation, my mom won an award for the column, and it was revealed that *Da Akamai Kid* had been my mother all along, a shock to everyone, including her own siblings who also attended the school and enjoyed reading the column every week.

In speaking with secondary school teachers in Hawai'i about whether there is a place to develop reading or writing in Pidgin, some say they like the idea but would find it difficult to make room for it because state curriculum standards keep them busy with other learning material. The goal of preparing students for later success makes using class time for work that does not explicitly meet these standards feel like a risk, one that most are understandably unwilling to take.

In Hawai'i, when people talk about “preparing for the future,” what is often meant is providing students the tools to succeed in life outside of Hawai'i. There is something sad in how this goal is repeatedly framed, positioning life beyond Hawai'i as the benchmark of success. It seems that in Hawai'i we remain largely stuck in an inferiority complex when it comes to education, one in which everyone else is assumed to be smarter than us.

I understand the valid fear of being poorly equipped for a world in which English is the dominant language, and I do not blame anyone for wanting what is best for the youth. Yet, I think about the joy that writing in Pidgin could bring some students if it were fostered in the same way that English writing is, and about how few opportunities students have to celebrate Pidgin in the classroom.

In my mother's case, while the adults around her were demanding that she learn to “speak properly,” to the point that anxiety turned into swallowed words and silence, *Akamai Kid* allowed her to be herself. In fact, when I read this now this person sounds a lot more like the person I know to be my mother than the shy person people describe her to be as a high school student. When doing this research, another thing that gave this work a sense of purpose was what felt like the opportunity to champion the little *Akamai Kid* in us all, advocating for Pidgin speakers to find our voices in language rights efforts in Hawai'i and to recognize ourselves as *akamai* in our own right. My message is not that everyone should go out and learn Pidgin if they do not have a connection to it, nor is it a criticism of parents or educators who prioritize English education. Rather, this work is an advocacy for those who have

2.3. Kumu and Haumāna

been told their whole lives that their native language makes them less than akamai.

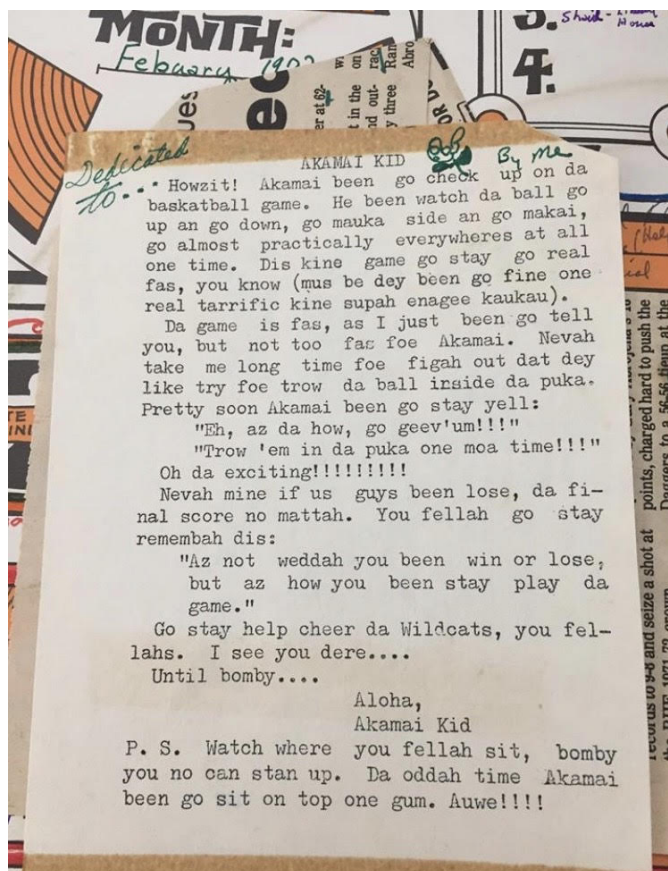


Figure 2.2: Akamai Kid, published in 1974 at Konawaena High School, written by Karen Kamigaki

2.3 Kumu and Haumāna

2.3.1 Mo'olelo I: Language Loss, Shame, and Collective Responsibility

In keeping with my commitment to conducting this research with kuleana, I engaged in extensive one-on-one interaction with participants both before and after the study. I wanted to take accountability for any emotions the study may have left them with.

2.3. *Kumu and Haumāna*

One of the most common types of feedback encompassed concerns of inadequacy in Hawaiian language proficiency.

Here are some examples of responses:

“OK, all pau. Man, way to make me feel inadequate. HAHHAHAHA. I better go practice my ‘ōlelo!”

“Please keep the money and you might to throw out my study. I tried my best to distinguish if ‘ōlelo Hawai‘i words were real or not, but since I am not quite fluent and have only seriously begun to expand my vocabulary this semester I may have incorrectly identified some words from my own lack of knowledge.”

“I definitely was second-guessing myself because I know I don’t know all Hawaiian words.”

In conducting linguistics research, a real potential harm is that individuals may feel inadequate as a result of participating. I would argue that this risk is even higher when working with Indigenous languages.

To all such individuals, I reassured them that this was not a test of knowledge, but rather a way of comparing groups of people with different language backgrounds to one another. Usually, however, their response was that, while they understood the study wasn’t a test, the task itself still led them to feel insecurity about their language abilities.

One participant responded, “I know, I just did not want to disappoint you with the infancy of my Hawaiian language abilities”.

In retrospect, I could have done a lot differently to prevent this. The process was often, make a stimuli list myself, with the help of a dictionary, and hire Hawaiian language consultants for help checking for issues. Being too eager to continue on with a timeline I needed to follow, I sometimes accepted a perhaps more hesitant than firm “yes” when asking whether they thought this could work with students at the Hawaiian 102+ level. I also did not pilot the stimuli with Hawaiian language students, I asked available non-Hawaiian speaking lab members to pilot as well as Hawaiian language consultants but not newer speakers of Hawaiian. This decision was influenced by concerns about losing potential participants during piloting, as the language community is already quite small. I believed I had done my due

diligence in ensuring the task was of high quality, but could have been more careful.

In hindsight, had I known that participation might leave individuals with such feelings, I would have approached the study design and recruitment criteria differently. (Specific details on future studies are included in discussion for Experiments 1 and 2).

I would like to note that all participants who reached out to express such feelings *were* proficient speakers, as evidenced by their high accuracy on the task. No matter the case on their “task accuracy,” however, I do not wish to retraumatize participants who have experienced language loss and leave them feeling negatively about themselves as speakers. The issue at hand is not whether the participants knew enough Hawaiian to take part in the study, but rather how we can conduct linguistics research involving such communities with adequate care.

Linguistics literature formally defines a phenomenon called “linguistic insecurity” which encompasses a speaker’s personal anxiety, self doubt, or lack of confidence about their language use. However, I personally do not believe this term quite fits in this context. “Insecurity” takes upon implication of individuality, an uncertainty of “oneself”. In the case of ‘ōlelo Hawai‘i, and languages like it, the feeling of loss is a collective experience.

For example, although the speakers were themselves proficient, collective experience of knowing this is still not enough in itself may lead to feelings of shame. I also recognize this in conversations with fluent Hawaiian speakers, who discuss their belief in shared kuleana for those who possess knowledge of ‘ōlelo Hawai‘i to teach it to others when they can. It appears the perpetuation of the language is a shared burden, held by the community of those invested in the life of ‘ōlelo Hawai‘i.

One participant wrote to me, “while I had the privilege of going to the Kamehameha schools, I feel shame that I had the opportunity while my cousins did not. I feel shame that I have no one else in my family to converse with. I feel shame that despite my experiences I am not better than I am.” This lends a further example to the notion that certain dynamics simply do not fit the label “linguistic insecurity”.

Moving forward in efforts to fix the problem, I spoke to some additional Hawaiian language teachers and received helpful feedback. One Hawaiian language teacher who conducts linguistics research in ‘ōlelo Hawai‘i told me that he believes such reactions are, to some extent, an inevitable part of this work, as some individuals may experience a sense of heaviness or leave with difficult feelings with anything that feels like a language test at all. He noted, however, that the burden can be made less heavy by strongly

emphasizing before the task that it is not an evaluation, clearly articulating what the study is doing, and intentionally holding space for participants' feelings after the study is complete. Another kumu said it is probably better to be more rigorous in the criteria than to lower the standards and risk leaving students feeling discouraged or harmed.

Taking all of this into account, I revised my recruitment protocol midway to specify that the experiment sought language teachers, translators, and native speakers. Some individuals who did not fall into these categories were still permitted to participate if they explicitly indicated a desire to do so.

Surprisingly to me, one of the most fruitful "solutions" I found for addressing feelings of shame involved changing the compensation structure by creating the option for participants to donate their participant compensation to the 'Aha Pūnana Leo Hawaiian immersion schools. This suggestion emerged in conversation with my friend, Lindsey Nkem, a PhD candidate at the Institute for Gender, Race, Sexuality and Social Justice, who is also an exceptionally solution-oriented community organizer. Although this option was originally introduced to encourage participation among individuals who were not motivated by the \$15 gift card compensation, its implementation resulted in a notable shift in participant feedback.

"No need money for me, donate um! I just happy to contribute to the lāhui."

"I don't know if I did well enough on the study or if it was helpful to you, but I am very happy to donate my money. At least I know I am helping somebody. Actually, I did not even know about the 'Aha Pūnana Leo schools before now. I just wanted to say I am glad I did it because for me now it feels like the start to a larger journey myself of doing what I can for the (Hawaiian) nation. I kind of did not think about how I could take small steps myself in helping the community until now."

"I am a lingual Pidgin speaker and do not yet speak Hawaiian at all. I want to donate the money to the schools so keiki can get a chance to learn the language natively and grow the nation of Hawaiian speakers."

Although this change did not fully eliminate feedback expressing feelings of shame or inadequacy, many of the comments following the decision to make the donation option available reflected a shift in participant affect or tone. While some days of community work are undeniably difficult, witnessing the openness, generosity, and commitment expressed by community members

fills me with hope and gratitude for the future steps we can take together as a community.

In total, over \$1000 was donated to ‘Aha Pūnana Leo immersion schools through participant donations.

2.3.2 **Mo‘olelo II: If you can, please kōkua me**

During an extended period of recruitment, I engaged in extensive public outreach of the research through talks across multiple islands in Hawai‘i, media appearances, flyers, utilizing teacher and community networks to promote the study, yet participation remained in single digits. When the study was posted publicly online, including in Hawaiian language groups or through public “resharing” on social media, it attracted a large number of bots and individuals from unrelated regions who falsely claimed eligibility and completed the task by entering random responses in order to receive monetary compensation. People genuinely seemed excited to speak with me about the work and share with me their thoughts, which I deeply appreciated. Nonetheless, I began to worry whether I would ever be able to recruit enough participants to complete the study on time.

I questioned whether this reflected insufficient monetary incentive, too small a size of individuals that spoke the required languages but also had the time to spare to take the study, or even worse, perhaps everyone hated me and only pretended to be enthusiastic about my ideas so as to not hurt my feelings and thought my idea was stupid or that I was a *fraud*.

In desperation, I began reaching out directly to friends and posting on social media, writing messages such as, “If you can, please kōkua me. I really need to graduate”. Within a single week, I received nearly one hundred responses from friends and acquaintances across many stages of my life, as well as from friends of those friends and acquaintances. People I had not spoken to over a decade ago reached out to say that they were not only happy to participate but also eager to share the study with others they knew. I also received emails from people whom I am still uncertain how they came to know about the study, expressing a desire to help, including messages such as, “I love this and what you stand for. I want to help you graduate too, girl!”

Some others who had seen me struggling with recruitment earlier later reached out, “I didn’t respond earlier because I didn’t want to take your money, I assumed this was just a thing for people who needed it. Now that I know it’s to help you I am happy to help!”

Again, this process left me profoundly grateful for the care people showed, not only toward the research but also toward me, whether as a friend, acquaintance, or complete stranger. That many participants were motivated less by the compensation than by the opportunity to *kōkua* someone they identified as part of their community in need was a touching lesson to learn.

2.3.3 **Mo‘olelo III: Tūtū lauhala lei hulu**

After giving a public talk in which I argued for the value of Pidgin within the Hawaiian language revitalization movement based on historic facts that set up my research question, a revered Hawaiian language teacher raised her hand and shared that, having been raised in a household where the use of Pidgin was strictly prohibited, she had never questioned the view that Pidgin was bad. As a result, she explained, she had also made it a point to discipline students for using Pidgin in her classroom.

Wearing a wide brimmed lauhala hat encircled with lei hulu, attire often associated with respected elders, this tūtū carries the kind of presence that quiets a room the moment she enters it. I felt embarrassed by the fact that she had to listen to anything I said at all.

She continued, “I was raised to think Pidgin would hold us down as people, but today I think it’s not fair to say we need to speak English to deserve respect. It simply has never occurred to me before. So thanks, I just wanted to say, you changed my mind”.

I didn’t know what to say in response. If I am being honest, in recounting this now I am not even sure exactly what I said. I was struck by the humility of her choosing to share this, given her depth of wisdom and decades of experience as a teacher at the forefront of the movement, alongside my own position at the time as a naive and idealistic student.

I was also humbled by the realization that linguistics work has little impact unless we are willing to move beyond comfort zones, exchange perspectives out in public, and understand that any work we do must defer to and support those already doing this work in the community. Limiting work to academic circles ultimately undermines its capacity to serve the people it is meant to benefit.

2.3.4 **Mo‘olelo IV: Pidgin as Resistance in the Hawaiian Language Classroom**

I had a long conversation with a Hawaiian language teacher, Kalikoaloha Martin, who has publicly shared his philosophy of why he personally chooses

to use Pidgin as a medium of instruction when teaching Hawaiian. His perspective differed from many of the conversations I had prior, as it was explicitly motivated by political activism. “In 1999, Haunani Kay Trask said that 75% of students within the University of Hawai‘i system are local, while 75% of teachers are non-local. I think this tells our students a message, that in order to be successful they need to emulate their teachers, including their language, value system, and perspective on the world”.

When looking up this statistic today, I found that indeed it is the case that a majority of students are local (something more like 68% today) and the vast majority of faculty are non-local. An article in 2019 (Hofschneider [2019]) problematizes the fact that a vast majority of the faculty at the University of Hawai‘i are white while the student body is largely non-white. Some departments are worse than others. At the University of Hawai‘i at Mānoa, Linguistics and Second Language Studies Departments, for instance, in 2025, 0% are local, and 0% are Kānaka Maoli. In referencing this statistic, Kalikoaloha states “This is a hegemonic force from the institution saying who belongs in the academy and what kind of knowledge is supreme”.

Kalikoaloha argued that within the Hawaiian language classroom, the proportion of local and native students is even higher than the general student population, “its more like 90% local, and even among the 10% I get some kānaka aunties and unkos from da continent. When considering that 90% along with the diaspora kānaka, they all know Pidgin”.

“In my classroom, for those that do not know Pidgin, I tell them if you need me fo translate um in English I will, but teaching Pidgin is my pedagogical choice to make the classroom feel like Hawai‘i”.

“This is beneficial to everyone. For non-locals, they often tell me they get the two for one deal! I think for them as well it is important to know about the two languages and their dynamic as they are really spoken during their time spent in Hawai‘i”.

Kalikoaloha described two types of criticism he typically receives. One comes from positions of power that assert Pidgin is not professional. The second comes from students who express discomfort with the feelings that arise when they are surrounded by the language. He identifies both attitudes as rooted in internalized colonialism, *haole mentality*.

The point about those in positions of power demeaning the use of Pidgin as a marker of lacking professionalism was a sentiment that emerged repeatedly and unprompted across conversations with multiple teachers. Dr. Keao NeSmith, a prominent Hawaiian language activist, linguist, translator, and language kumu, for instance, raised this point in a separate conversation. He told me, “I take the stance that Hawaiian language education needs to be

grassroots. It is from the people, and it is not meant to be gate-kept by the elite. I always taught Hawaiian using Pidgin, but at the institutional level I was scolded and told off by my supervisor that it was ‘not an appropriate way to conduct oneself.’ When I told her, ‘hey, that is some elitist b.s. Are you even interested in actual student learning or just the image (of the school),’ they retaliated and didn’t rehire me the following year”.

Kalikoaloha explained how he addresses students who have internalized negative attitudes toward Pidgin. “If students ever have a problem with Pidgin, I have a heart-to-heart with them. Who made you da Pidgin police? I talk to them on a personal level about it. I understand them, they’ve been told Pidgin is bad their whole lives. But for every student still working through that, we get the student too who will say wow I don’t have other classes that say Pidgin is ok. That’s pretty cool, mister. I can be me”.

As with everyone I speak to about this research, whether students, teachers, activists, or others, I ask what they hope to see in my work. When I asked Kalikoaloha this question, he said he hopes to see the work shared. For now, he noted, we have only anecdotes. To have something that empirically shows that “hey, Pidgin is not just okay, but it is helpful in acquiring Hawaiian” would mean so much to so many people. It would matter for everyone who has been silenced, made fun of, or corrected for speaking Pidgin. “We are all going to read your manuscript and we are all going to tell everyone about it”.

We ended the conversation discussing how to keep going, despite threats coming from every direction to stop. He told me, “For me, I remember da unko from Wai’anae who finally learning how fo ‘ōlelo Hawai’i (to speak Hawaiian). Tailoring the classroom experience for him by speaking his language. We gotta remember to do this for unko”.

2.3.5 Mo’olelo V: Pidgin as Access, Care, and Relational Pedagogy

Ali Rozet, another Hawaiian language teacher, described how using Pidgin helps her teach, “One of the reasons I use Pidgin in the Hawaiian classroom relates to bringing students back to childhood, the language is a familial tie”. It is true, for many of us that grew up with the language, we are tied by blood to the originators of the language. Pidgin is the language that brings us back to the comfort of home.

After discussing what it meant to make students more comfortable, Rozet also noted, “I also would say Pidgin was always a language about coming together to get something done, but remembering to do so with humor and

lightness”.

Expanding on this, Rozet related how Pidgin was used by our ancestors to hana (work) together, sharing the same inside jokes across workers no matter what ethnic group we came from, and how that legacy continues into the present. “In Hawai‘i today, Pidgin still represents lightness in the face of difficulty. I never want a student to feel stupid or shame in the classroom, or to leave feeling worse about themselves than when they entered.”

Rozet noted the utility of using Pidgin when students are struggling to concentrate on the lesson. “If I say, like, ‘hui, e listen,’ (hey come listen) I can ask a student for their attention and remove that aspect of some boss laying down the law. For me it’s about prioritizing student comfort, bringing myself to the same level as the student, and allowing a familial purpose back into the classroom.” She emphasized her awareness that if a student has a bad experience, they may not want to continue learning Hawaiian, which she identifies as the greatest disservice to the movement.

Throughout our conversations, Rozet emphasized the importance of intentionally leveraging Pidgin’s structural relationship to Hawaiian as a pedagogical tool. “For many of us who grew up in Hawai‘i, our first Hawaiian words were learned at home, through ‘ōlelo Pa‘i‘ai. I like showing students those similarities, so they can recognize how much knowledge they already have.”

At the same time, Rozet was clear that, for her, the heart of this practice lies in connection. Above all else, the use of Pidgin in the classroom is about creating a sense of ease, familiarity, and connection so that students feel they belong in the classroom and believe they are capable of succeeding.

Similarly, Dr. Larson Ng, who teaches statistics in the Department of Education at the University of Hawai‘i, shared with me his decision to use Pidgin as a medium of instruction in his own discipline. He explained, “Pidgin is the first language for the majority of my students. I do it because it relieves anxiety and fear, it is a way for me to assert the Pidgin worldview into the subject matter, and I believe it builds confidence. A student that sees me talking Pidgin, it is a strong statement.”

Like Rozet, Ng also discussed the feeling of connection as central to his choice to use Pidgin in the classroom. He explained, “Speaking Pidgin signals that I wasn’t educated at Punahou. I am just one braddah from Kalihi Valley. If I can do it, my students definitely can. Pidgin also speaks to a value system and a way of life. There are concepts in statistics that mirror what local students already encounter in their daily lives. I often explain everything through local food metaphors, ‘ahi, poke, crab, one can ‘o spam (a can of lunch meat), and weave Pidgin into those examples as a

way of talking about overcoming hardship.”

The accounts shared by both kumu illustrate that the use of Pidgin in the classroom is a deliberate pedagogical methodology. Across the two disciplines, Pidgin functions as a connection to home, to prior knowledge, to a shared value system, and to a sense of belonging that many Pidgin speaking students have been historically denied. By drawing on a shared cultural identity, both kumu adopt the use of Pidgin as a way of resisting shame and creating the conditions for students to feel they belong in the classroom. In doing so, they affirm that academic success does not require abandoning one’s language, worldview, or place, but rather, embracing them. In this way, Pidgin becomes not only a tool for creating a more comfortable learning environment, but also an assertion of who the classroom serves and to what end. The classroom is for the community, and its teaching is oriented toward uplifting that community.

2.4 Lāhui, in Hawai‘i and abroad

2.4.1 Mo‘olelo I: Pidgin Stigma

I was unprepared for the emotional weight of the feedback I received through this community-engaged work, as my public-facing research elicited strong responses, both positive and negative.

Most responses were overwhelmingly supportive. Many people shared that they had never been told that speaking Pidgin was “okay.” Others expressed hope for a future in Hawai‘i grounded in greater self-acceptance, noting that they felt motivated to advocate for Pidgin, deepen their study of Hawaiian, and support this work in whatever ways they could.

Others were extremely critical and personal, arguing that promoting the use of Pidgin was actively harmful to the educational, career, and economic futures of our keiki.

“Telling people to keep talking wrong and uneducated is a horrible message and is damaging. It is anti-productive and is keeping us down education-wise and career-wise.”

One Facebook comment on a repost of our episode of Hawai‘i Public Radio entitled *Linguist explores if Pidgin speakers have an advantage in learning ‘ōlelo Hawai‘i* asked, “Does she want us to be held back forever? What about the international market and tourism?” Another comment writes, “Ok so let our kids speak wrong. Then how will our kids have a future at all?”

What is interesting is that both types of feedback, supportive and critical, point to the same underlying issue: Pidgin stigma persists in the present.

This is actually difficult to find empirically.

I have come to find that those who are most eager to participate in a formal study are often already inclined to view Pidgin favorably, or they recognize an ethos of advocacy for Pidgin reflected in the study itself that they might not wish to challenge. Further, from my experience, people are also just so much more willing to be candid in their direct interactions with me or on a Facebook comment than in a formal study, especially if the study is affiliated with the university.

A high school classmate of mine, for instance, shared with me personally, “I have one word for what I think about Pidgin. Yuck!” However, in a formal interview wished to keep her feelings more neutral.

Self-identifying as a “Pidgin speaker” also biases the sample toward a particular subset of individuals who are willing to associate themselves with the label.

Negative feedback often comes from speakers of the language themselves, or from those who grew up with it. This is unsurprising, as these individuals have directly or indirectly experienced the consequences of stigma and may wish to protect themselves and their loved ones from further harm.

I understand this response to stigma deeply. While others have experienced far worse than I have, I have nonetheless encountered the consequences of Pidgin stigma firsthand, particularly in academic settings. As early as elementary school, teachers would ask, “Do we talk liddis or do we speak properly?” In high school, classmates openly expressed bigoted views about peers who spoke Pidgin in class, treating the language as evidence of lower intelligence or being low-class. As an undergraduate at the University of Southern California, a white student once overheard me speaking Pidgin with a classmate and remarked that he found it “disgusting” to hear the English language “get butchered.” The list of such experiences, unfortunately, goes on.

There were moments when my mother would recall from her childhood when everyone would sit around the television watching Daniel Inouye, the late senator from Hawai‘i, deliver an address in English. The adults in the room would say, “Hū dis guy talk real good. E try listen, bumbai one day you go be like him,” (This guy is well spoken. Listen to him, so one day you will be like him). Growing up, whenever I saw Daniel Inouye on television, my mom would repeat the same sentiment she had heard growing up: “He talks so good.” Without saying it directly, they were comparing Pidgin, the language spoken at home, to the English he spoke and positioning Pidgin as the lesser variety. People like Daniel Inouye, especially those with family histories similar to ours, were held up as models of who we could be, if only

we learned to speak the prestigious code.

So while nothing I personally experienced, whether directly or indirectly, comes close to what others have experienced, I understand why, even in my own family, Pidgin could not always be defended despite holding what I would consider relatively positive views of the language. Pidgin was seen as a force that held us back.

It would be inaccurate to say that Pidgin stigma comes solely from outside forces or from those in positions of power. Internalized shame operates differently from externally imposed stigma and carries its own consequences. I find internalized views of Pidgin to be much more difficult to challenge and unpack. Internalized stigma, I have come to learn, is often shaped by a desire to protect and care for others. It emerges as a response to real social consequences and the fear that challenging the norms held by those in power may lead to further harm.

A Kanaka Maoli political activist once shared with me that while he “really agreed with the message” of discussing how Pidgin activism can be utilized to support ‘ōlelo Hawai‘i revitalization, he believed it was my due diligence to also “share the nuance that Hawaiians and locals are also capable of causing harm, it is not just a haole issue”. He remarked that, in the case of minoritization of both languages, “elite Hawaiians played a hand in the decline of the Hawaiian language, and still improperly claim kuleana to dictate what words we collectively decide to use in our language.” Likewise, with Pidgin, he argued that “when you talk about who hates Pidgin the most and who is responsible for the narrative that Pidgin is low-class, stupid, or whatever, you cannot honestly say it’s been your experience that haoles are solely to blame for that.” As he put it, “no one hates Pidgin more than elitist Hawaiians do,” and he urged me to “hold Hawaiians accountable for the role they have to play in this story you share about language in Hawai‘i.” He urged that activism that “truly makes a difference” must also critically engage with the stigma that emerges from within the community.

I believe it is important to acknowledge that, like any nation or community, Hawaiians and other locals can also cause harm to fellow community members, especially when dynamics of power and privilege are at play. No group is monolithic in their beliefs, of course, and it would be dehumanizing to suggest that any group lacks the agency to cause harm or participate in the perpetuation of harmful narratives.

While it remains my stance that many of the issues surrounding language stigma in Hawai‘i are both directly and indirectly rooted in dynamics embedded within American imperialism, it is also important to acknowledge that community members, both Kānaka Maoli and other locals, have agency

and are not uniquely exempt from the human capacity to cause harm.

Claims made by those external to the community that Pidgin is no longer stigmatized in Hawai‘i stand in tension with the realities I encountered through this work, and in lived experience. For those of us living in-community with the language, stigma remains strongly felt. The question, therefore, is not whether harm persists, but how we move forward. Acknowledging harm sometimes requires holding multiple truths at once. Language stigma in Hawai‘i is inseparable from American imperialism and racial hierarchy, yet it is also maintained through internalized shame and in-community dynamics of power and class.

To name this complexity is not to assign blame, but to insist on responsibility. As a scholar, it is important to resist both romanticizing Pidgin and dismissing the negative feelings shared with me, including fear, while instead making space for honest reckoning. I have found that this form of community-engaged scholarship not only illuminates truths that often remain unseen in more formal studies, but also creates opportunities for dialogue that can actively challenge stigmatizing attitudes and beliefs.

Most effective in disrupting such stigma, I have found, has been engaging through story. I share the mo‘olelo I have learned through this work, including how Pidgin emerged as a means of serving our community, how it enabled workers to organize across differences in home languages and demand better wages from haole plantation owners. I relay the stories of how Pidgin continues to function as a language of resistance to American assimilation, and can be used to make space for ‘ōlelo Hawai‘i to (re)emerge. I also draw on linguistic work to emphasize that Pidgin is a language in its own right, rather than a deficient or broken form of English. I discuss how monolingualism is an American myth, and how research in language acquisition shows we do not need to fear Pidgin in preventing the acquisition of any language, including English or ‘ōlelo Hawai‘i. Although it is uncomfortable, I make it a point to share this nuance and make space for the negative emotions that others may direct toward me in the sharing of these mo‘olelo.

In these conversations, I have also found it useful to draw upon perspective learned from other communities beyond Hawai‘i. I have shared how people I have met from Nigeria describe their own Pidgin as no longer stigmatized, noting that it is broadcast on platforms such as BBC Radio. I also share how Indigenous leaders in Canada have told me that they look to the ‘ōlelo Hawai‘i movement as a source of inspiration, citing the achievements of Hawaiian leaders and community members as evidence of what language revitalization can accomplish. In many of these moments, I have been surprised by how effective this approach has been. In temporarily decentering

Hawai‘i, some individuals are able to reconsider their assumptions.

I leave this section with a message shared by a participant, one that serves as a reminder that, while more difficult, doing work in this way can promote active change:

“Thank you for working so hard on this. I am so proud and I can say wholeheartedly that my girls are too, to have been a teeny, tiny part of this by way of you. Thank you for stepping up and doing the work to help our keiki. For defending us, our culture and our people and for helping to acknowledge that we are more than what we have been recognized for. Thank you on behalf of my keiki and for our future generations to come and the generations before us. I know it was and is not an easy task, but we are grateful for your work. Keep up the awesome job. I love this.”

2.4.2 Mo‘olelo II: Hālau-Style Approaches to Revitalization

I spoke with a community leader that works in food production to discuss her mana‘o on learning ‘ōlelo Hawai‘i as not only an effort that needs to be led by language teachers, but also one that is led by community leaders in other areas.

“In traditional Hawaiian education, we had hālau-style learning, it was holistic, you learn math, science, the Hawaiian language, by participating in an activity, and center the land, center family, center food... The language was practiced as it lived through life itself.”

Hālau, broadly defined as a place for gathering, is often associated with canoes or with instructional spaces for practices such as hula.

She continued, “When I think of hālau-style learning, I think about how language is encountered and lived rather than documented in a book. Pidgin makes sense because it lives alongside ‘ōlelo Hawai‘i, that is, out in public rather than in textbooks. When I work with kalo I can say we all ‘ōlelo (speak) Pidgin, and then we will randomly switch to ‘ōlelo Hawai‘i when it fits. So I think in the natural setting, in the daily life of your typical Hawaiian today, cutting out Pidgin and implementing a strict regimented haole classroom goes against our way.”

Similarly, Keao NeSmith quotes a Hawaiian language teacher respondent that took his study, “(In the study of ‘ōlelo Hawai‘i) I’d like to see our people go back to hālau. In ancient times people would gather at these types of structures to learn hula and/or other things... We need to go back to... family styles of learning ‘cause it’s... the way to learn. ‘Cause... our language is holistic...and that’s the way it should be measured: through your kumu and through your hālau.”

A third person I spoke with, a political activist involved in Hawaiian rights advocacy, shared how he uses hālau-style learning to push back against what he described as *elitism* in ‘ōlelo Hawai‘i revitalization circles. He shared this from a deeply personal place, saying, “I am saying this as a Hawaiian speaker, as the first one to go to college. Currently, most of those who know how to speak Hawaiian are among the privileged. Free online Hawaiian classes exist, but without pilina embedded in the course material, meaningful access to education is limited.” He shared that, as a citizen, he personally works to address this issue by perpetuating the language through Pidgin outside of the classroom, in spaces where it is freely spoken among locals. The language, he says, allows him to demonstrate his relationality within the family structure.

Providing an example, he shared, “My ‘ohana members constantly ask me about different Hawaiian words they remember from the plantations. They go, ‘eh boy boy, what does this word mean?’ From my experience, Pidgin recalls memories of experiencing Hawaiian through plantation work, through discussions of land and how to mālama, or care for it. I can use these moments as opportunities to help perpetuate Hawaiian in ways that connect to what they already know.”

What stood out to me in these conversations about hālau-style learning was how solution-oriented they were in addressing ‘ōlelo Hawai‘i revitalization beyond the classroom, as part of everyday life. Across all three perspectives, there was a shared emphasis on language revitalization as lived practice, framing pedagogy as a collective pursuit attentive to questions of access, lived experience, and the realities of learning alongside others.

Perhaps in this way, the distinction between language acquisition and language revitalization hinges on whether the pursuit of language (re)emergence is oriented toward individual or collective goals. While language revitalization depends on individual commitment and processes of language acquisition, it is fundamentally a community-driven process. As such, it must ask how language acquisition can function within community contexts, with deference to authority extending beyond the language classroom to leaders in other domains, and with attention to how language is lived across different sectors of community life.

In the context of endangered and Indigenous languages such as ‘ōlelo Hawai‘i, language learning is necessarily a collective endeavor. By contrast, for globally dominant colonial languages like English, learning can function solely as a personal goal, guided by the question of how the language serves the individual learner. English will continue to thrive regardless of whether any single person learns it well or takes responsibility for its perpetuation.

For ‘ōlelo Hawai‘i, however, an individualistic pursuit does little to serve the language, which comes alive only through shared practice.

This collective pressure, coupled with an intense desire to support the community, can become a heavy burden shared across the community, often manifesting as shame. For some, this is precisely why alternative ways of accessing ‘ōlelo Hawai‘i education are necessary, approaches that move beyond Western-styled classroom structures and institutional elitism toward everyday, lived practice. Not every individual thrives in the same rigid environment, and more importantly, the language is lived outside the classroom. When exploring hālau-style practice as a framework for linguistic research in service of language revitalization, the persistence of shame demands careful attention.

Returning to my conversation with Kalikoaloha Martin, a Hawaiian language teacher, he spoke about next steps for actively addressing the shame often experienced by ‘ōlelo Hawai‘i learners. For him, mitigating shame requires both naming its origins and redistributing its weight. As he explained,

“I think reducing the shame comes from two things. First, identifying the origins of that shame as haole. They shamed us in the schools for generations, mocking us for our Hawaiian and our Pidgin. Second, the burden is shared between you and me, teacher and student. Just because I am now a Hawaiian speaker does not mean it is no longer my burden to share in your learning. Speaking Pidgin is my way of demonstrating humility to my students. I use it with the intention of asserting that the burden, and everything that comes with it, is mine and yours.”

2.4.3 Mo‘olelo III: On languages as living beings

During a public talk with members of the community at Ka Haka ‘Ula o Ke‘elikōlani, I was asked several direct questions about the utility of Pidgin, often framed in terms of whether it deserved space at all. One undergraduate student remarked,

“English and Hawaiian have financial value. I can write them both on my resume, and both are seen as assets that make me more employable. With Pidgin, this is not the case. I would be laughed at if I wrote that I speak Pidgin on my resume. Even though it is the language I speak with my parents, should we bother teaching our children the language?”

This question stayed with me, in part because on a different day, an Uber driver addressed the same underlying concern, though attached to the opposite language.

He told me that he had just graduated from Kamehameha and shared he

“regrets” spending so much time studying Hawaiian. He explained, “Pidgin is still financially valuable to me because people speak the language all over the place. If I become a doctor, I will have patients who speak Pidgin to me every day. Knowing Hawaiian has never helped me since graduating high school. Only my friends who have gone on to work in cultural preservation have ever used their Hawaiian.”

Truthfully, I don’t blame either of them. Each of these young people raised valid and practical concerns about their futures. Their comments nonetheless underscored how a language, and by extension the culture of a people, can be reduced to capital value alone, a condition produced by English language imperialism as it is currently structured.

I posed the question to both young people the value of language, does this boil down to value of cultures and value of people? Is the purpose of knowing a language only to place it on a resume or to use it at work? And if so, is that enough reason to allow a language to disappear?

In response to the question at Ka Haka ‘Ula o Ke‘elikōlani, I asked the audience whether they had ever felt more familiar, more comfortable, or closer to home when meeting someone for the first time who spoke to them in Pidgin. People in the audience nodded their heads. I followed by asking a quieter question: what is the value of that feeling?

A Hawaiian language teacher responded, “It occurs to me that perhaps we should come to see Pidgin, Hawaiian, and English each as living people, each with their own role in a shared community. After all, are these languages not living beings?” Others in the audience were nodding and writing notes.

A student chimed in, “If Pidgin were a person, he would be the easygoing one. Maybe he makes less money than the other guys but he does not need to be all show off, he is the one who invites everyone to dinner, people from all over the world.”

Conceptualizing languages as living beings reflects an ‘ōiwi way of knowing, and is seen across Indigenous communities. This perspective moves beyond speaker counts and economic measures, making space for other forms of value that are less easily quantified. For many, this is why the language is not described as “dead” or “dying,” but rather as “sleeping,” still living.

Back in the Uber, I spoke with the driver about the kinds of work he understood to be encompassed within “cultural preservation.” He responded that he never really challenged this terminology before. He began, “Well, I think of work like for ‘ōlelo, doing something academic like Hawaiian history, or at the Bishop Museum or doing music or hula...” Then trailed off, “But you know, I think it does expand, like I have friends that work for Hawaiian airlines and they really are in high demand for those that know how ‘ōlelo

there.”

When he mentioned wanting to become a doctor, I shared the true story of one of my closest friends, Lucia, a Kanaka Maoli physician who had just begun her practice in Hawai‘i. I told him a story she had shared with me about how traditional Hawaiian medicine remains deeply woven into everyday practice in Hawai‘i, alongside research supported by the National Institutes of Health that points to tumor-fighting properties in traditional medicinal plants used in ‘ōiwi practice, including noni. I found myself wondering aloud about the many ways “cultural preservation work” might live outside the narrow frames we often place around it, especially when ‘ōlelo preservation work is understood as a daily practice sustained through engagement across different domains.

The driver told me he was grateful for the conversation, that maybe his time had not been wasted, and that perhaps his knowledge of Hawaiian was an asset after all.

Before I left, the driver asked me, “Hey, is Pidgin called Pidgin because people think it’s disgusting, like the pigeon bird?” I told him no, I did not think so. He nodded and said, “Well, good. That would make me sad.”

2.4.4 Mo‘olelo IV: Kānaka Maoli in the Diaspora: What happens when there are no speakers around you?

The He ‘Ōlelo Ola conference held at Hilo featured four pānela (workshop panels) that discussed various aspects of the concept of ola ka ‘ōlelo Hawai‘i (continuing the life of the Hawaiian language) in diaspora. This panel discussed the reality of the current situation: the majority of Native Hawaiians now live outside Hawai‘i. According to recent U.S. Census data, about 53 percent reside on the continental United States, while approximately 47 percent live in Hawai‘i. Much of this panel focused on how Kānaka Maoli in the diaspora are sometimes viewed as having done something shameful by moving away, even though many leave because they have no other choice.

Prior to attending the panel, I conducted surveys with Kānaka Maoli about their experiences learning ‘ōlelo Hawai‘i in diaspora. Many described feelings of isolation when they did not have others nearby with whom to practice the language. For some, this lack of community felt like an insurmountable barrier.

Taking the opportunity to pose a question to the panel, whose members were well versed in the lived experience of perpetuating ‘ōlelo Hawai‘i while in the diaspora, I asked, “What advice would you give to Kānaka struggling to find community in their effort to use and practice the Hawaiian language?”

One of the panelists grew very still, her expression stern, as if the question carried some pain. After a pause, she spoke carefully,

“I would say, this is really important for us to consider. How often many of us have found ourselves alone. Alone in our commitment, alone because we are away, we still do what we do but we need to do it alone. Genealogy has always been important to us. Tell them to take the time to research their ancestors, learn things about them, get to know them. Then when you have gotten to know them, call upon them and practice with them.”

Her response reminded me of Noelani Goodyear Ka‘ōpua’s mo‘olelo on the importance of genealogy in language and life: “(When one) return(s) to the past, (one) can actually breathe that life, that ea, into our culture and traditions again. So it’s not relics of the past but rather a living and breathing function of who we are as people.” Through the act of calling upon members of our genealogy, then, it becomes possible to breathe life into the language.

Since this panel, in my engagement with Hawaiian teachers in the diaspora, I have met more than one person deeply invested in teaching Hawaiian outside of Hawai‘i. One teacher, who works at a daycare, shared that she teaches Hawaiian on the continent and understands this work as her kuleana. She said this work helps promote exposure to the Hawaiian language outside of Hawai‘i and makes it possible to experience immersion while geographically removed from the ‘āina.

One message we were left with by the panel was that there may also be unexpected possibilities in diaspora for the vitality of ‘ōlelo Hawai‘i. One speaker reflected, “While many of us did not choose this, how special is it that now we can speak out all over the world? May the world hear us ‘ōlelo Hawai‘i everywhere and know who we are and demand that space.”

2.4.5 Mo‘olelo V: Inter-community Work

At the He ‘Ōlelo Ola (the living language) conference, an Indigenous woman from Vancouver Island in British Columbia raised her hand and spoke. She shared that she had attended the same conference two years earlier, and that at the time she had posed a question to Kumu Larry Kimura. “I told him, my language is so small and the situation is so dire. I am not sure I am a good enough speaker and I am not sure I trust in myself. I want my kids to speak the language, and I want to create a language school to help the other families in the community come learn. How can I possibly do this without the right resources?”

In her story, the woman recounted that Kumu Larry replied, “You are

good enough. Just try your best. That is more than enough. You are doing right by your language and you are doing the language proud with your efforts.”

The woman then shared that after that interaction, she went right home and started a small school with her children along with children in the neighborhood.

Her eldest child stood up and fluently spoke the language, and said she hopes to make language games for kids and elders to spontaneously speak the language together.

I was struck by how a few words of encouragement could ripple outward, having a profound impact far beyond Hawai‘i.

You are good enough.

I was also moved by how a man I knew, along with so many others in the movement, has not only done immense work for our own community, but has also extended aloha, care, and encouragement to other communities across the world. Work that seeks to address the effects of colonization can feel so immense and impossible to take on alone. To me, Kumu Larry’s advice to the woman made me think of how in CPR, even one breath can be enough to resuscitate a life. Even a small step can help breathe life into a language and allow it to flourish.

The concept of breath is deeply significant within Hawaiian epistemology. The word *ea* means breath, but also means life, to rise up or emerge, and it also means sovereignty. As Noelani Goodyear-Ka‘ōpua writes, “like breathing, *ea* cannot be achieved or possessed; it requires constant action day after day, generation after generation.” As spoken language requires breath, this can be seen both as literal and metaphorical. Language requires breath.

Through my own engagement in community work beyond the *lāhui*, including fieldwork with Secwépemc communities in the interior of British Columbia and Anishinaabe communities across regions such as Ontario and Manitoba, as well as public talks across Canada, I have come to see how these conversations extend far beyond Hawai‘i. Inter-community conversations have felt much like the sharing of medicine between communities, learning from what has been healing for others and carrying those practices home to be taken up, adapted, and shared again.

To return to *ea*, breath is not something that can be stored or secured once and for all. It must be taken in continually. It must be shared and circulated. It must be renewed. In the same way, *ea* as life, as emergence, and

as sovereignty exists through ongoing action. Languages live on not through protection alone, but because community members continue to breathe life into them. Sustaining that breath through the exchange of ideas, practices, and encouragement between communities is itself can be seen as a form of healing ea.

2.5 Conclusion: Weaving the Lei

2.5.1 Kulāiwi

Mai ka pi‘ina a ka lā i Ha‘eha‘e	From where the sun rises at its gateway at Ha‘eha‘e
A i ka mole ‘olu o Lehua	To the pleasant tap root foundation of Lehua Islet
E ia au ko kama ē	Here I am your offspring
He mamo aloha na Hāloa	A beloved descendant of Hāloa
‘O Hawai‘i ku‘u kulāiwi	Hawai‘i is my native land
Mai nā kūpuna mai	From my ancestors
‘A‘ohe mea nāna e kū‘ai	No one can purchase
I ke ewe o ku‘u mau iwi	The homeland of my lineage
E ola au i ku‘u lāhui	I will live on in my cultural identity
Ke kuleana o ka ‘ōiwi	The right and responsibility of a native son or daughter
‘O ka ‘ī ma kāna ‘ōlelo	To declare in his or her own language
He Hawai‘i au mau a mau	I am a Hawaiian for now and forever

– Mele Na Larry Kauanoe Kimura

For the audio recording of this mele, see this link.

I conclude this mo‘olelo chapter with this mele, written by Dr. “Kumu” Larry Kimura. When I visited Ka Haka ‘Ula O Ke‘elikōlani, Kumu Larry

2.5. Conclusion: Weaving the Lei

performed the mele for a group of us during lunch. Before singing the song, he shared with us a brief mo‘olelo, “You see, when I visited some of our Native American friends on the continent they asked me where Hawaiians live and whether we also live on reservations. I thought about how, in Hawai‘i, we are not confined to one space but rather you can find us living all around Hawai‘i. All of Hawai‘i is our reservation. Mai ka pi‘ina a ka lā i Ha‘eha‘e (from where the sun rises at its gateway at Ha‘eha‘e) a i ka mole ‘olu o Lehua (to the pleasant tap root foundation of Lehua Islet) ‘o Hawai‘i ku‘u kulāiwi (Hawai‘i all around is my home)”.

This experience brought together several strands of my thinking developed throughout the research process and arrived at a critical moment, as I was nearing the completion of the study phase for this thesis. Life in Hawai‘i is a privilege for many reasons, including the constant presence of ‘ōlelo Hawai‘i in the soundscape. Rather than being confined to the boundaries of a particular reservation, it permeates public space and everyday life.

Hearing the language all around us creates the conditions for developing co-associations between Pidgin and ‘ōlelo Hawai‘i, as well as for the emergence of proto-lexical knowledge, a concept I explore throughout this work.

Although many languages are spoken in the islands, ‘ōlelo Hawai‘i remains ever present in the ambient, shaping the linguistic environment even for those who do not speak it fluently. A central goal of this thesis is to understand how community members themselves may be practitioners of language revitalization. Rather than positioning revitalization as the work of a linguist observing a community from the outside, this perspective highlights the everyday acts through which people participate in revitalization simply by living ‘ōlelo Hawai‘i.

My hope is that this work helps shift a familiar narrative. Through my community work, I have witnessed how many members of the lāhui experience feelings of shame, often tied to being speakers of Pidgin rather than ‘ōlelo Hawai‘i. Grounded in service to this community need, this thesis seeks to affirm that we are already participants in language revitalization, regardless of whether we yet consider ourselves speakers of Hawaiian or the work we do to make a living. By moving beyond feelings of shame, bashfulness, and deficit-based framing, this perspective emphasizes that linguistic knowledge is continually being built through everyday exposure to ‘ōlelo Hawai‘i within the community.

With this power comes our kuleana to perpetuate ‘ōlelo Hawai‘i in the ways we can, by committing to practices that make time for the language and support its continuation, regardless of our genealogy or whether we live on ‘āina or in the diaspora.

2.5. Conclusion: Weaving the Lei

‘A‘a i ka hula, waiho ka hilahila i ka hale. When one wants to dance, bashfulness should be left at home. May we all be given the strength to resist shame and bashfulness as we perpetuate ‘ōlelo Hawai‘i, whether as an act of aloha ‘āina or as hoā aloha.

I include this song in recognition of Kumu Larry Kimura, who composed it and contributed directly to this work. He is a co founder of ‘Aha Pūnana Leo and has played a central role in Hawaiian language revitalization across media, education, and community initiatives. His leadership includes serving as director and coordinator of Ka Leo Hawai‘i, a radio program dedicated to interviews with native speakers across the island chain, and as a founding director of the Hale Kuamo‘o Hawaiian Language Center at Ka Haka ‘Ula o Ke‘elikōlani at the University of Hawai‘i at Hilo. He continues this work through his involvement with the Hawaiian Lexicon Committee, Kani‘āina, digital voice archives of native speakers, culture based science education, and inter community language revitalization beyond Hawai‘i.

For me, the song serves as a point of continuity between personal experience and the broader themes of this dissertation concerning language, community, and belonging. My association with the song begins in childhood, when my grandmother often played it on the radio in her home in Kona. Through this research, I came to understand its message more fully and its relevance to my developing sense of place within the community.

Chapter 3

Experiment 1: A Psycholinguistic Exploration of Code-switching to ‘ōlelo Hawai‘i from Pidgin versus Hawai‘i English

3.1 Introduction

This chapter presents Experiment 1, which used a visual world paradigm to examine potential perceptual advantages and costs in code-switching for trilingual speakers of English, Pidgin, and ‘ōlelo Hawai‘i. The primary goal was to determine whether the language preceding a switch affects access to the target language. The study tests whether switches from Pidgin to ‘ōlelo Hawai‘i incur lower processing costs than switches from Hawai‘i English to ‘ōlelo Hawai‘i. Code-switching is the practice of alternating between languages, within the same discourse, within a single phrase, or at the level of smaller syntactic units (Poplack [1980]). In this study, an intra-sentential code-switching advantage was assessed using reaction times (RTs) and eye-tracking data, with shorter RTs and a higher proportion of looks to competitor (relative to the target) indicating an effect of facilitation (e.g. Kranzler [2012]). Here, *increased facilitation* refers broadly to one language yielding lower reaction times, while *lower switch costs* refer to a reduced difference between unilingual and code-switched speech in one language relative to another.

Note that in the present study, “switch cost” is used as a descriptive term referring to the net difference in reaction time between code-switched and unilingual trials. This measure reflects the combined influence of processes that slow performance during language switching (e.g., suppression of the previously active language, particularly when it is more dominant or more

distinct across multiple dimensions from the target language) as well as processes that may facilitate performance (e.g., overlap in cues or priming effect across two languages). Accordingly, observed differences in switch costs may reflect variation in cost-inducing (inhibitory) mechanisms, facilitative mechanisms that enhance processing efficiency, or both. In the current study, we use the term “code-switched benefit” to refer collectively to all effects attributed to the frame language that influence processing cost or efficiency when considering code-switching to the target language.

Speech perception in a single language is already a complex task, requiring listeners to continuously map a rapidly unfolding rich acoustic signal onto probabilistic phonological, lexical, and syntactic representations, often resolving ambiguity among competing candidates (e.g. Marslen-Wilson [1987] McClelland and Elman [1986]). The perceptual system must strike a balance between generalization and specificity. On the one hand, successful speech perception requires generalizing across variation that does not affect lexical meaning. This involves filtering out external acoustic distortions, such as background noise (Mattys et al. [2012]), as well as accommodating speaker-related variation, such as differences in vocal tract anatomy, social identity, dialect, or co-articulatory patterns (Pierrehumbert [2003b], Labov [1994], Johnson and Sjerps [2018], Kleinschmidt and Jaeger [2015], Norris et al. [2003], Campbell-Kibler [2011], Hall-Lew et al. [2021]). On the other hand, listeners must also be sufficiently sensitive to finely-grained linguistic cues to accurately interpret meaning. In order to resolve lexical competition, for example, especially in the case of minimal pairs (e.g., Hawaiian /wahine/ vs. /wa:hine/, English /bæt/ vs. /pæt/, Pidgin /bit/ vs. /bi:t/), listeners must be sufficiently sensitive to subtle acoustic details to distinguish phonemic contrasts that share nearly all features (Luce and Pisoni [1998], Kleinschmidt and Jaeger [2015]), Norris et al. [2003]).

Within the executive control framework, successful unilingual speech perception relies on cognitive control to manage and resolve competition across linguistic representations. The Adaptive Control Model (Green and Abutalebi [2013]) predicts that code-switched perception introduces additional complexity by bringing cross-linguistic phonological, lexical, and syntactic systems into the perceptual field. At the lexical level, according to this model, words that are connected in meaning cross-linguistically are co-activated (Grosjean [1997a], Thierry and Wu [2007], Costa et al. [2000], Spivey and Marian [1999], Van Heuven et al. [1998]). This co-activation theoretically increases the number of lexical candidates the perceptual system must contend with, placing greater demands on cognitive control mechanisms to select the intended representation (e.g., Hartsuiker and Pickering [2008]).

These increased demands, arising from the need to manage co-activated lexical representations across languages, may make processing more effortful (e.g., Kuipers and Thierry [2010]), often reflected in longer reaction times (e.g. Grainger and Beauvillain [1987], Grosjean [1997b], Thomas and Allport [2000], Green and Abutalebi [2013], Hartsuiker and Pickering [2008]), and in eye-tracking studies, higher proportion of gazes to competitors (e.g., Barr [2008], Magnuson et al. [2007], Fricke et al. [2016]). This effect has been shown in bilinguals, as well as in trilinguals (e.g., Lemhofer et al. [2004]).

Most work on code-switching has focused on speech production, showing that multilingual speech often incurs measurable costs due to increased cognitive demands, such as selecting words from competing languages and inhibiting the non-target language (Kaushanskaya and Marian [2007], Wu and Thierry [2012]). These costs are reflected in greater pauses prior to uttering a target lexical item and phonological influence from the non-target language (e.g., Amengual [2012], Goldrick et al. [2014], Jacobs et al. [2016], Meuter and Allport [1999], Thomas and Allport [2000], Costa and Santesteban [2004], Gollan and Ferreira [2009]). However, these costs may be reduced when switches occur spontaneously, as speakers can draw on learned maintenance strategies to facilitate code-switched production (e.g., Zhu et al. [2022], Kleinman and Gollan [2016]). Likewise, while many experimental studies report that code-switched speech is more costly to process than unilingual speech, the magnitude of this cost may depend on listeners' language experience (e.g., Grosjean [1997a], Green and Abutalebi [2013]).

High exposure to code-switching in one's language background can reduce perceptual costs, especially when switches occur in predictable, naturalistic contexts (e.g., Cheng and Howard [2008], Beauvillain and Grainger [1987], Caramazza and Brones [1980]). Multilinguals that are well adapted to code-switched input are likely more sensitive to language-specific and anticipatory cues, and are better at managing interference from competing lexical representations (e.g., Martín et al. [2010]). Subtle phonetic cues that precede a switch, such as slowed speech rate or the integration of features from the upcoming language, can be utilized by listeners to help anticipate a switch and reduce costs in perception (e.g., Shen et al. [2020], Fricke et al. [2016], Krizman et al. [2017], Kuipers and Thierry [2010]).

The sociolinguistic context, including the frequency of code-switching in a particular community, plays a role in reducing code-switching costs in experimental studies. Additionally, the relative likelihood of switching between specific language pairs may affect code-switching costs (e.g., Blanco-Elorrieta and Pykkänen [2017]). Relatedly, a pair of languages may be more strongly co-activated based on being closely tied to shared identities, such

as localness, heritage, or group solidarity (e.g., Szakay et al. [2016]). A large body of research shows that listeners’ expectations about speaker identity shape higher-level processing of speech, influencing how sounds are encoded and interpreted within context (e.g. Niedzielski [1999], Drager [2011], Hay et al. [2006]), Babel and Munson [2013a], Johnson et al. [1999]). Work in this line of research suggests that listeners encode speech variation in relation to speaker identity, and that multilinguals can draw upon this information across languages to facilitate processing.

Other factors which may contribute to code-switching costs include language dominance, proficiency, age of acquisition, and word frequency. The dominant language typically exerts stronger influence during lexical access in both bilinguals (e.g., Costa et al. [2006]) and trilinguals (e.g., Aparicio and Lavour [2014]), while activation of the weaker language occurs only after a certain proficiency threshold is reached (Van Hell and Dijkstra [2002]). Moreover, regardless of dominance, the L1 often has a stronger effect than subsequent languages (L2+) (e.g., Jonathan Grainger and Holcomb [2010]). Frequency effects also play a role in code-switching such that lower frequency words in a unilingual phrase can be just as or more costly to process than higher-frequency cross-language items (e.g., Yacovone et al. [2021], Moreno et al. [2002]). Low-frequency words incur greater processing costs, especially for listeners with lower proficiency.

In the current study, I investigate whether ‘ōlelo Hawai‘i lexical items are more readily accessed when preceded by Pidgin rather than Hawai‘i English in code-switched sentences. Specifically, we ask whether switching from Pidgin to Hawaiian incurs lower processing costs than switching from Hawai‘i English. We predict that switching from Pidgin to ‘ōlelo Hawai‘i is likely to incur lower processing costs than switching from Hawai‘i English for three main reasons: (1) tighter sociolinguistic ties between Pidgin and ‘ōlelo Hawai‘i, (2) frequent co-use and code-switching between the two languages in naturalistic contexts, and (3) shared cue-weighting strategies that guide lexical access in both languages.

To illustrate the pattern that would indicate lower processing costs for Pidgin-to-‘ōlelo Hawai‘i switches compared to Hawai‘i English-to-‘ōlelo Hawai‘i switches, Figure 3.1 presents a schematic representation of the hypothesized data. The y-axis represents mean reaction time (in ms), and the x-axis represents trial type, within each panel, with unilingual trials on the left and code-switched trials on the right. Frame language is indicated by color: red for Hawai‘i English and blue for Pidgin. The plot depicts two potential cases: one in which code-switching costs are lower for Pidgin-to-Hawaiian switches (right panel) and another in which costs are lower for Hawai‘i English-to-

3.1. Introduction

Hawaiian switches (left panel). Baseline performance in unilingual trials is illustrated as equal across frame languages, indicating no overall effect of frame language. The code-switched condition is also depicted as more costly, with longer mean reaction times compared to the unilingual condition. Differences across frame language and code-switched conditions are reflected in the change from unilingual to code-switched trials: Pidgin-to-Hawaiian switches show reduced code-switching costs relative to Hawai'i English-to-Hawaiian (right panel), whereas Hawai'i English-to-Hawaiian switches show reduced costs relative to Pidgin-to-Hawaiian (left panel). The left panel illustrates the hypothesized pattern, showing that Pidgin-to-'ōlelo Hawai'i switches incur lower processing costs than Hawai'i English-to-'ōlelo Hawai'i switches.

3.1. Introduction

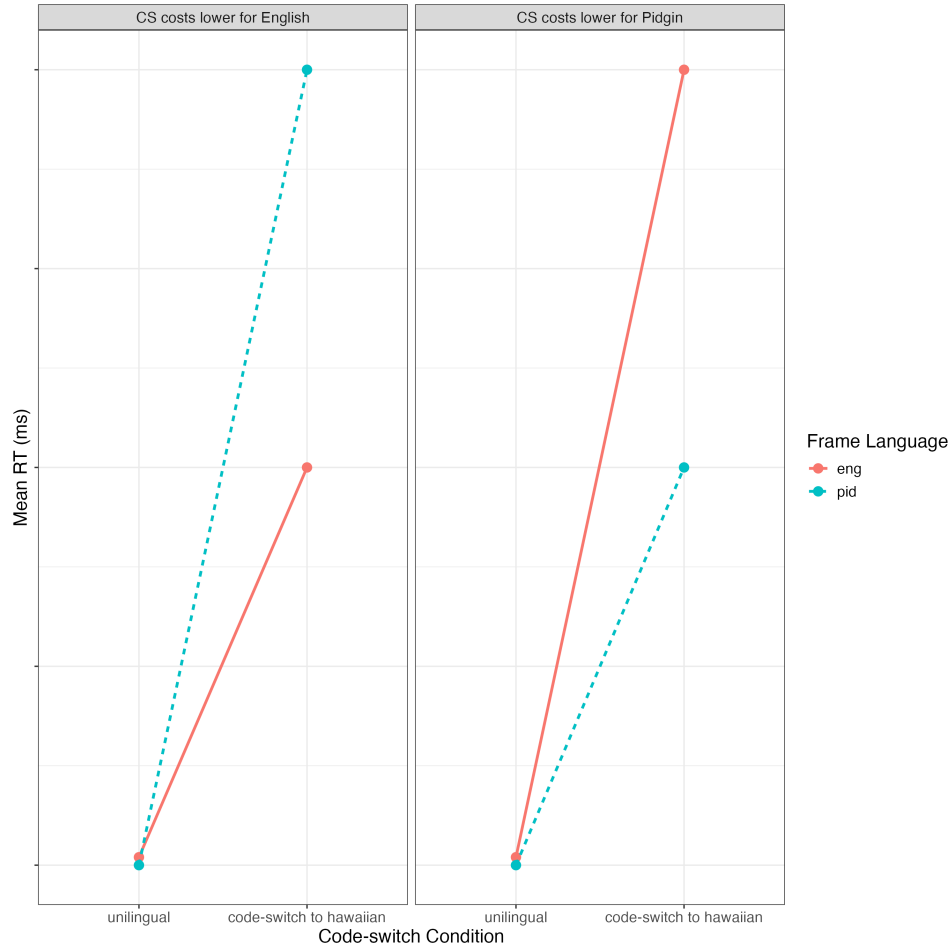


Figure 3.1: Cartoon data illustrating the interaction between frame language and trial type such that code-switching (CS) costs are lower in Hawai‘i English on the left; and code-switching costs (CS) are lower in Pidgin on the right.

There are two primary reasons to hypothesize that Pidgin facilitates access to ‘ōlelo Hawai‘i: contextual and sociolinguistic connections, and linguistic overlap. We hypothesize that Pidgin facilitates access to ‘ōlelo Hawai‘i because the two languages are conjectured to commonly co-occur in naturalistic settings and are socially linked to “local” identity in Hawai‘i. Although most speakers of ‘ōlelo Hawai‘i also speak Pidgin and Hawai‘i English, in

the context where English was historically imposed as the local prestige language, ‘ōlelo Hawai‘i and Pidgin are used as a linguistic means to simultaneously align with local Hawaiian values, establishing belongingness to the in-group (Sato [1991b], Sato [1993], Marlow and Giles [2008], Marlow and Giles [2010b]). The two languages may also more frequently co-occur, as both languages are codes used in contexts to express opposition or resistance to threats of American assimilation (Reinecke [1938]). The tighter connection between Pidgin and ‘ōlelo Hawai‘i, formed through frequent code-switching and shared group identity, motivates our hypothesis, aligning with prior findings that sociolinguistic context and speaker identity influence listeners’ expectations and reduce code-switching costs (e.g., Blanco-Elorrieta and Pytkänen [2017]; Szakay et al. [2016]; Niedzielski [1999]; Drager [2011]).

In addition to their social and contextual connections, Pidgin and ‘ōlelo Hawai‘i also share linguistic overlap, which we propose may provide an additional source of facilitation in processing when switching from Pidgin to ‘ōlelo Hawai‘i. As the primary substrate of Pidgin, ‘ōlelo Hawai‘i and Pidgin also share structural features across several linguistic dimensions, including at the syntactic, lexical, and word levels. For instance, both languages use vowel duration as a cue to distinguish between contrastive vowels. ‘ōlelo Hawai‘i has a contrastive distinction between long and short vowels (i.e. wāhine ‘woman.PL’ wahine ‘woman.SG’), while Pidgin similarly uses duration to distinguish overlapping vowel categories (i.e., green [g.ai: n] and grin [g.ain])(Grama [2015]). Extensive work has shown that multilingual listeners flexibly re-weight acoustic cues based on linguistic experience and contextual demands (e.g., Holt and Lotto [2006], Francis et al. [2000], Chandrasekaran et al. [2010b]). Multilingual listeners are able to dynamically re-weight cues in response to linguistic contexts (Bushong and Jaeger [2019]). Language features heard most recently are more likely to be activated in memory and can influence perception, lexical access, or decision-making in tasks (e.g., Goldinger [1998]). To my knowledge, no empirical research has yet demonstrated that shared cue-weighting strategies reduce code-switching costs for one language pair over another. Nevertheless, it is plausible that such linguistic overlap facilitates multilingual speech perception, since multilinguals adjust cue weighting following exposure to relevant features (e.g., Francis et al. [2000]; Chandrasekaran et al. [2010b]; Bushong and Jaeger [2019]). In the case of Pidgin and ‘ōlelo Hawai‘i, recent exposure to Pidgin may prepare listeners for processing speech that relies on the same cue-weighting strategies as the upcoming language. These shared linguistic features could thus give participants an advantage when processing ‘ōlelo Hawai‘i, creating a facilitative effect when code-switching from Pidgin to ‘ōlelo Hawai‘i. By

contrast, Hawai'i English shares fewer linguistic features with 'ōlelo Hawai'i and, for example, prioritizes spectral cues over duration, meaning duration is not a primary cue-weighting strategy, so it would not likely provide the same potential facilitative activation as Pidgin might.

Thus, in the case of Pidgin and 'ōlelo Hawai'i, potential facilitation may be stronger than for Hawai'i English and 'ōlelo Hawai'i for three reasons: first, due to the social link between the two languages; second, because of the frequent contextual co-occurrence between Pidgin and 'ōlelo Hawai'i in naturalistic settings; and third, due to their linguistic similarity. Hearing Pidgin prior to a switch to 'ōlelo Hawai'i may allow trilinguals to pre-activate features shared between the two languages and readily apply shared cue-weighting strategies, anticipate language use in socially familiar settings, and predict usage in naturalistic contexts, thereby facilitating the processing of the 'ōlelo Hawai'i item. This facilitation is less likely to occur when switching from Hawai'i English to 'ōlelo Hawai'i due to limited linguistic, social, and contextual overlap.

Alternatively, from an inhibitory control perspective (e.g., Meuter and Allport [1999]), reduced switch costs may reflect a diminished need to inhibit the previously active language, thereby reflecting a *reduction in inhibition* rather than facilitation. From this perspective, because 'ōlelo Hawai'i and Pidgin share more overlapping cues than 'ōlelo Hawai'i and Hawai'i English, switching between Pidgin and 'ōlelo Hawai'i involves less change and thus reduced inhibitory demands. Although the facilitation and inhibition accounts propose different mechanisms, they make the same prediction: greater overlap between Pidgin and 'ōlelo Hawai'i should result in reduced switch costs.

Research on code-switching perception seeks to understand not only the real-time processing demands of switching between a particular pair of languages but also the broader organization and accessibility of multiple languages in the mind. Building on this, the present study examines whether two languages within a trilingual community are more closely linked in processing than another language pair shared within the community, thus expanding our understanding of how different language associations shape asymmetries in multilingual language use. By testing whether switching from Pidgin to 'ōlelo Hawai'i incurs lower processing costs than switching from Hawai'i English, we examine how more tightly knit sociolinguistic ties, frequent co-use, and shared cue-weighting strategies may jointly facilitate multilingual comprehension. Code-switching remains understudied in trilingual populations, particularly in speech perception, where interactions among all three languages may uncover patterns that do not emerge in bilingual or bidialectal contexts. Moreover, due to Pidgin being conceptualized as existing

along a creole-continuum from the acrolect (which is more Hawai‘i English-like) to the basilect (which incorporates a greater proportion of ‘ōlelo Hawai‘i linguistic features and lexical items), this study may help bridge the gap between prior research on bidialectalism and bilingualism by examining how a linguistic variety that sits between a dialect and a distinct language interacts with other languages in the mind in multilingual processing. In other words, if Pidgin leads to notable differences compared to Hawai‘i English in perceptual behavior concerning access to ‘ōlelo Hawai‘i, it could reveal mechanisms underlying multilingual comprehension more broadly.

Beyond linguistic implications, this work addresses community concerns regarding whether Pidgin may serve as an effective instructional medium for ‘ōlelo Hawai‘i language learning. As discussed in Chapter 1, Pidgin is an L1 for the majority of the population in Hawai‘i (Sakoda and Siegel [2003]); however its use is stigmatized and is largely suppressed or even punished in the domain of education locally. Negative attitudes associated with Pidgin, especially when spoken in the classroom, include “unintelligent”, “uneducated”, “rude”, and “low-class” (Sato [1991a], Marlow and Giles [2010a], Choy and Dodd [1976], Day [1980], Yamamoto and Hargrove [1982], Yokota [2008]). This poses an undue burden on local youth. Regardless of the stigma associated with Pidgin, ‘ōlelo Hawai‘i language instructors have increasingly been adopting the practice of using Pidgin as a medium of instruction as opposed to English, often citing the cultural and linguistic connections between the two languages (e.g. Martin [2018], Cabral [2018], Rozet [2017b], Rozet [2018]). Understanding how code-switching from Pidgin versus Hawai‘i English affects access to ‘ōlelo Hawai‘i has important implications in the community, as this research addresses questions actively raised by local teachers and students surrounding the value of Pidgin in the classroom, particularly with respect to supporting ‘ōlelo Hawai‘i revitalization efforts. The current study also reflects the linguistic reality in Hawai‘i today, where many learners aiming to improve their Hawaiian proficiency identify to varying degrees of dominance as trilingual speakers of English, Pidgin, and ‘ōlelo Hawai‘i. More broadly, this work has implications for second or additional language education in revitalization contexts, particularly where there is flexibility in choosing the medium of instruction that best supports learners. In other communities aiming to revitalize a language, this raises the question of how using a creole or mixed language as the medium of instruction might affect students’ perceptual access to the target language.

3.2 Participants

3.2.1 Pidgin, English, ‘ōlelo Hawai‘i trilinguals

This experiment investigates trilingual speakers of ‘ōlelo Hawai‘i, Pidgin and English. Participants were identified as speakers through self-reporting. There were several pathways through which participants were recruited. 1) Through the Linguistics department at the University of Hawai‘i at Mānoa “Linguistics Beyond the Classroom” program, 2) via physical flyers which were posted on campus at the University of Hawai‘i at Mānoa and were distributed by Hawaiian language teachers to students taking Hawaiian 202 at the University of Hawai‘i at Mānoa, 3) through Facebook ‘ōlelo Hawai‘i advocacy groups including “Hoa ‘ōlelo Hawai‘i”, 4) by the investigator through media appearances including Hawai‘i Public Radio and the Moana Nui Podcast.

A total of 35 participants completed the study. Of these, 19 participants indicated in their Language Background Questionnaire that they did not meet the minimum qualifications, leaving 16 participants for inclusion in the analysis.

Among participants, substantial variation was observed across languages, indicating that trilinguals differed in their self-reported age of acquisition and proficiency across the three languages. These scores were obtained from a Language Background Questionnaire (see Appendix A). For age of acquisition, English was predominantly reported as having been learned from birth (age 0) across participants, with relatively little variation. Pidgin was also most commonly reported as being acquired from birth; however, responses showed greater variability, with some participants indicating later ages of acquisition. In contrast, reported ages of acquisition for ‘ōlelo Hawai‘i were widely distributed, reflecting substantially more variation across participants (see Figure 3.2).

Proficiency ratings, measured on a scale from 1 to 10 (1 = lowest proficiency, 10 = highest proficiency), showed a similar pattern. English was consistently rated as highly proficient across all participants. In contrast, Pidgin and ‘ōlelo Hawai‘i exhibited greater variability. Self-reported proficiency for ‘ōlelo Hawai‘i ranged from as low as 1 to as high as 10, indicating heterogeneous experience with the language. Pidgin was generally rated as relatively high in proficiency, though typically lower than English, with scores spanning a broad range from 3 to 10 (see Figure 3.3). Lower Pidgin proficiency scores likely reflect the characteristics of this particular participant pool, which was recruited from the university. Participants in this context

3.2. Participants

may spend more time speaking English than Pidgin in their everyday lives due to the exclusion of Pidgin from certain domains, including in higher education.

Overall, these differences reflect substantial variability in participants' linguistic experience across the three languages, with English consistently dominant in both age of acquisition and proficiency, Pidgin showing intermediate and more variable patterns, including slightly greater variability in age of acquisition and wider variation in proficiency compared to English, and 'ōlelo Hawai'i characterized by later acquisition and lower, more dispersed proficiency ratings. This pattern may give rise to differences in expected performance in language processing across the three languages, with an advantage for English relative to Pidgin, and the weakest performance expected for 'ōlelo Hawai'i.

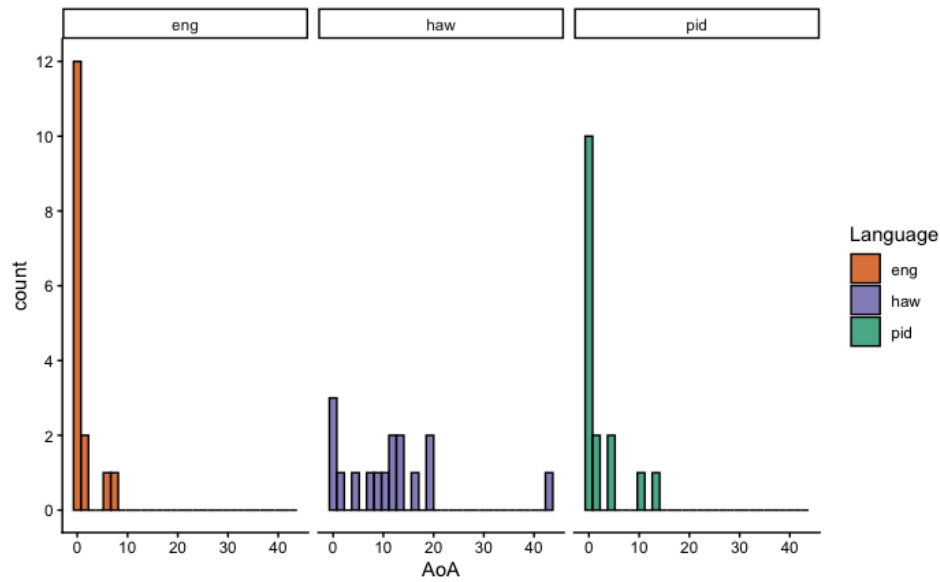


Figure 3.2: Distribution of age of acquisition (AoA) ratings for Pidgin, English, and 'ōlelo Hawai'i across participants. Values are plotted separately by language, allowing comparison of AoA distributions across the three language groups.

3.2. Participants

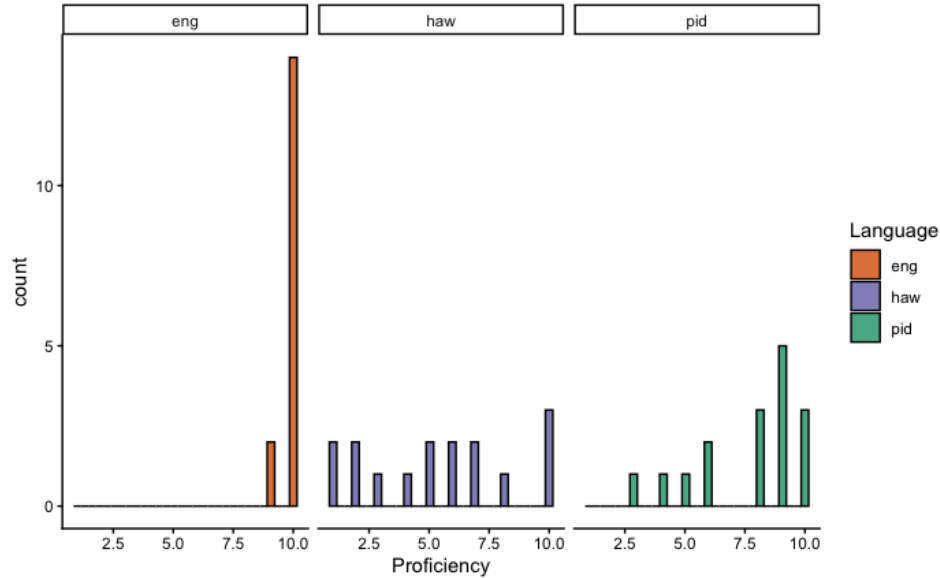


Figure 3.3: Distribution of self-reported proficiency ratings for Pidgin, English, and ‘ōlelo Hawai‘i across participants. Values are plotted separately by language, allowing comparison of proficiency distributions across the three language groups.

3.2.2 Inclusion and exclusion criteria

Participants completed a Language Background Questionnaire (see Appendix A) to characterize their experience with English, ‘ōlelo Hawai‘i, Pidgin, and any additional languages spoken. All participants self-identified as trilingual in English, ‘ōlelo Hawai‘i, and Pidgin, though the degree of exposure, proficiency, and dominance varied across individuals. The questionnaire included items on language history, frequency of use, and self-rated proficiency, which were used to estimate relative language dominance across the three languages.

Participants were required to be at least 18 years old and to report experience with English, ‘ōlelo Pa‘i‘ai (Pidgin), and ‘ōlelo Hawai‘i. Although no formal verification of language proficiency was required during recruitment, participants were asked at the start of the study to confirm these eligibility criteria and provide detailed information about their language backgrounds in the questionnaire.

3.3 Materials

This study employed a within-participant design, in which each participant experienced all experimental conditions. This was decided due to i) expected between-participant differences in baseline reaction time (RT), ii) variation in eye tracking and button press behavior, iii) differences in computer hardware (e.g. size of computer screen, keyboard), processing capacity, etc, and iv) variation in language dominance. In this experimental design, unilingual phrases appeared in Hawai'i English or Pidgin, while code-switched phrases would switch from either Hawai'i English or Pidgin into 'ōlelo Hawai'i.

3.3.1 Phrases and Wordlist

All words were placed in either a Pidgin or Hawai'i English carrier phrase. The Pidgin carrier phrase was, "E try pick da word [target]" and the Hawai'i English carrier phrase was, "Please select the word [target]." All [target] items were in Pidgin or Hawai'i English in the case of unilingual phrases or in Hawaiian in the case of code-switched phrases.

There were 36 trials in total. Each trial included a target word and a competitor. The target-competitor pairings were either morphologically related or morphologically unrelated. Morphologically related pairs contrasted by number, either in Hawaiian (e.g., wahine 'woman' vs. wāhine 'women') or in Pidgin/English (e.g., doctor vs. doctors). Morphologically unrelated pairs contrasted near-minimally, either by vowel length (e.g., waha 'mouth' vs. wāwae 'foot') or by glottal stop presence or absence (e.g., 'ahi 'tuna' vs. ahi 'fire'). Interlingual distractor pairs began with phonetically similar onsets but did not match by language (e.g., English 'nose' vs. Hawaiian 'niho' 'tooth').

Of the 36 trials, 18 were unilingual (9 Pidgin-Pidgin, 9 English-English) and 18 were code-switched (9 Pidgin-Hawaiian, 9 English-Hawaiian). Among these, 6 trials were morphologically related with Hawaiian targets, 6 were morphologically related with Pidgin/English targets, 6 were morphologically unrelated by vowel length, 6 were morphologically unrelated by glottal stop, and 18 were interlingual distractors.

To counterbalance, two word lists were made, such that all Hawaiian items were preceded by a Pidgin or Hawai'i English phrase in even counts. The lists differ in the pairs for the code-switched trials: each Hawaiian target word is paired with a different frame language such that items framed in Hawai'i English in List 1 are framed in Pidgin in List 2, and vice versa. All unilingual trials (English-English and Pidgin-Pidgin) are identical across

3.3. Materials

the two lists. This counterbalanced design allows us to test whether the frame language (Pidgin vs. Hawai'i English) preceding the Hawaiian target predicts responses, while ensuring that each item appears in both frame conditions across participants without repetition within a single list. The two experimental lists for code-switched trials are included in Tables 3.1 and 3.2 below. Unilingual trials that are the same in both lists are included in Appendix B.

Target Word	Target Language	Frame Language	Choice A	Gloss A	Choice B	Gloss B	Contrast Type	Pair Type
kanaka	haw	eng	kanaka	person	kanaka	people	vowel length	morphologically related
wahine	haw	pid	wahine	woman	wahine	women	vowel length	morphologically related
makua kāne	haw	eng	makua kāne	father	makua kāne	fathers	vowel length	morphologically related
kūpuna	haw	pid	kūpuna	elder	kūpuna	elders	vowel length	morphologically related
kaikamahine	haw	eng	kaikamahine	girl	kaikamahine	girls	vowel length	morphologically related
'elemakule	haw	pid	'elemakule	old man	'elemakule	old men	vowel length	morphologically related
waha	haw	eng	waha	mouth	wāwae	foot	vowel length	morphologically unrelated
hoku	haw	pid	hōnu	turtle	hoku	star	vowel length	morphologically unrelated
aniani	haw	eng	aniani	mirror	ānuenuē	rainbow	vowel length	morphologically unrelated
maka'i	haw	pid	maka	eyes	maka'i	policeman	vowel length	morphologically unrelated
paniolo	haw	eng	paniolo	cowboy	paluna	balloon	vowel length	morphologically unrelated
kalo	haw	pid	kalo	taro	kāma'a	shoe-slipper	vowel length	morphologically unrelated
'ahi	haw	eng	'ahi	tuna	ahi	fire	glottal stop	morphologically unrelated
apo	haw	pid	'aka'akai	onion	apo	circle	glottal stop	morphologically unrelated
'eke	haw	eng	'eke	bag	elelo	tongue	glottal stop	morphologically unrelated
i'a	haw	pid	'ilī'ilī	pebble	i'a	fish	glottal stop	morphologically unrelated
'ulu	haw	eng	'ulu	breadfruit	uaki	watch	glottal stop	morphologically unrelated
iwi	haw	pid	'iole	rat	iwi	bone	glottal stop	morphologically unrelated

Table 3.1: Code-Switched Stimuli List 1

Target Word	Target Language	Frame Language	Choice A	Gloss A	Choice B	Gloss B	Contrast Type	Pair Type
kanaka	haw	pid	kanaka	person	kanaka	people	vowel length	morphologically related
wahine	haw	eng	wahine	woman	wahine	women	vowel length	morphologically related
makua kāne	haw	pid	makua kāne	father	makua kāne	fathers	vowel length	morphologically related
kūpuna	haw	eng	kūpuna	elder	kūpuna	elders	vowel length	morphologically related
kaikamahine	haw	pid	kaikamahine	girl	kaikamahine	girls	vowel length	morphologically related
'elemakule	haw	eng	'elemakule	old man	'elemakule	old men	vowel length	morphologically related
waha	haw	pid	waha	mouth	wāwae	foot	vowel length	morphologically unrelated
hoku	haw	eng	hōnu	turtle	hoku	star	vowel length	morphologically unrelated
aniani	haw	pid	aniani	mirror	ānuenuē	rainbow	vowel length	morphologically unrelated
maika'i	haw	eng	maka	eyes	maka'i	policeman	vowel length	morphologically unrelated
paniolo	haw	pid	paniolo	cowboy	paluna	balloon	vowel length	morphologically unrelated
kalo	haw	eng	kalo	taro	kāma'a	shoe-slipper	vowel length	morphologically unrelated
'ahi	haw	pid	'ahi	tuna	ahi	fire	glottal stop	morphologically unrelated
apo	haw	eng	'aka'akai	onion	apo	circle	glottal stop	morphologically unrelated
'eke	haw	pid	'eke	bag	elelo	tongue	glottal stop	morphologically unrelated
i'a	haw	eng	'ilī'ilī	pebble	i'a	fish	glottal stop	morphologically unrelated
'ulu	haw	pid	'ulu	breadfruit	uaki	watch	glottal stop	morphologically unrelated
iwi	haw	eng	'iole	rat	iwi	bone	glottal stop	morphologically unrelated

Table 3.2: Code-Switched Stimuli List 2

All words selected were imageable nouns considered likely familiar to Hawaiian 102-level speakers³, based on input from two Hawaiian language instructors serving as consultants for the study. While Pidgin lexical items

³Participants with Hawaiian 102-level (HAW 102) experience would have completed two elementary courses (or the equivalent) in the Hawaiian language and were estimated by Hawaiian teacher consultants to have acquired the vocabulary required for the language tasks included in this study. However, completion of Hawaiian 102 does not necessarily in-

may sometimes overlap with Hawaiian, none of the Pidgin selected words were of Hawaiian origin. While some words from Hawaiian may be found in Pidgin, the decision could never have been made with knowledge of Pidgin alone. For instance, while the word *wahine* is also used as a word for “woman” in Pidgin as well as in Hawaiian, the plural form “*wāhine*” is not used but is rather omitted e.g. “two *wahine*” or adopts English -s plurality e.g. “two *wahines*”. Some Pidgin items were of English origin and some Hawaiian words were also of English origin.

3.3.2 Auditory stimuli

Auditory stimuli were recorded in a quiet room using a Zoom H4n Pro 4-Track microphone. The speaker, a self-identified male trilingual of ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English from Kāne‘ohe, Hawai‘i was asked to read sentences as they appeared on a computer screen, in a clear but natural speaking style. All Hawai‘i English target words were recorded in the English carrier phrase, “Please select the word [target].” All Pidgin target words were recorded in the Pidgin carrier phrase, “E try pick da word [target].” Each ‘ōlelo Hawai‘i target was recorded with one Pidgin and one Hawai‘i English carrier phrase. All stimuli were spliced and pasted into a single selected Pidgin or Hawai‘i English carrier sentence, to prevent the introduction of additional variables for participants during the experiment. The Hawai‘i English and Pidgin frames were recorded to have similar duration (selected frame for Pidgin: 1.40 seconds; selected frame for Hawai‘i English: 1.33 seconds). To ensure the interval from frame onset to target onset was identical across trials, 0.07 seconds of silence was added at the offset of the Hawai‘i English phrase, resulting in a total duration of exactly 1.40 seconds for each frame.

In contrast to the Hawai‘i English frame, the Pidgin frame “e try selek da” is composed of the “e” vocative which the language adopts from ‘ōlelo Hawai‘i (Roberts [1995]), the Pidgin “try” polite imperative (Sakoda and Siegel [2003]), Pidgin intonation, and Pidgin-specific phonology throughout, for instance *th*-stopping (e.g. /θ/ and /ð/ to /t/ and /d/, respectively), and final-/t/ deletion (Sakoda and Siegel [2003]).

To reduce low-grade background noise (e.g. hum of the air conditioner) in the audio recordings, a spectral subtraction technique was applied using the *librosa* library in Python. The technique utilized a Short-Time Fourier

dicade conversational fluency. As an elementary-level course sequence, proficiency among learners at this level varies substantially depending on factors such as immersion experience, family and community language use, and continued engagement with ‘ōlelo Hawai‘i.

Transform (STFT), which took the experimental audio file(s) and a noise-only segment extracted from a quiet portion of the same recording to capture the low-level background noise, measuring their time-frequency representations. First, the average spectral magnitude and phase components were identified for experimental audio and noise files across time, in the separate files. The original phase of the experimental recording was preserved so that the sound remains natural for a reconstruction step. Next, the magnitude of the audio was modified by subtracting the noise mean per frequency bin. The cleaned magnitude spectrum was then recombined with the original phase data, and an inverse STFT was performed (using the `librosa.istft` function in the `librosa` library) to convert the signal back into the time domain. The resulting denoised signal was scaled and saved as a 16-bit WAV file. The repository for this spectral subtraction algorithm can be found here: <https://github.com/shun60s/spectral-subtraction> (Shun [2020]).

3.3.3 Visual stimuli

For each trial, participants would see two images on their screen. For each trial, one image depicted the target corresponding to the presented audio stimulus, while the other image depicted the competitor that did not match the presented stimulus. All images were hand-drawn by a hired artist from Hawai'i.

3.3.4 Language Background Questionnaire

Participants filled out a language background questionnaire (LBQ) that asked questions specific to their experiences with English, Hawaiian and Pidgin, as well as other languages which the participant may speak. The full LBQ is included in Appendix A. The linguistic landscape of Hawai'i positions most 'ōlelo Hawai'i speakers as at least trilingual. The degree of exposure and dominance for each language may vary vastly by individual within the trilingual population. The LBQ provides a quantitative assessment of language dominance across these three languages (and potentially other languages spoken) and includes questions on qualitative descriptors of each participant in regard to their experience with languages in the context of Hawai'i. The LBQ was comprised of distinct themes: basic information (e.g., What year were you born?), language history (e.g., If applicable, list how many years you studied Hawaiian?; How many years of exposure to Hawaiian?), language use (e.g., In an average week, what percentage of the time do you use Hawaiian?; When choosing a language to speak with a person who

is equally fluent in all your languages, what percentage of time would you choose to speak Hawaiian?), language proficiency (e.g., On a scale from 0-10 (with 0 being the lowest, 10 being the highest), please select your level of proficiency in understanding Hawaiian?). Certain questions were language-specific, with reference to Pidgin, English, ‘ōlelo Hawai‘i, and other more general questions were asked in reference to any other language(s) spoken (e.g., Rank all languages spoken in order of dominance). The responses to these questions were then used to estimate the degree of language dominance between the three languages for each participant; details of this estimation are included in the descriptive analyses section. Responses from the LBQ will be used to describe and better understand potential dominance-based variation occurring in results.

All materials which advertised the study stated the requirements, which were that participants must be: “1. Minimally 18 years of age, 2. Speak ‘ōlelo Pa‘i‘ai (Pidgin) (at least 10 years of exposure), 3. Speak ‘ōlelo Hawai‘i (completion of Hawaiian 102 level+ or equivalent), and 4. Speak English. No proof of qualification was required by participants, but after many participants began signing up for the study without reading the qualifications, participants were all asked prior to the start of the study to confirm their qualifications, then describe their language background thoroughly in the language background questionnaire.

Language background questionnaire answers were analyzed: key questions were reported in the Results to understand the relationship between a participant’s language background and any observed Pidgin or English advantage. The 15 questions comprised: 1) self-reported order of dominance between Pidgin, English, Hawaiian, and any other languages spoken; 2) whether the participant considers themselves a “speaker” of each language (where “speakerhood” may differ from actual ability and instead reflects one’s identification with the label of being a speaker); 3) whether their parents consider themselves speakers of each language; 4) whether their grandparents consider themselves speakers of each language; 5) whether the participant was exposed to each language before the age of five; 6) whether they are Kānaka Maoli; 7) whether each language was spoken at home; 8) list of ‘ōlelo Hawai‘i classes taken (if applicable); 9) whether the participant speaks each language with friends if given the opportunity; 10) whether each language is spoken at work or school; 11) participation in activities with other speakers of each language; 12) self-reported order of language acquisition; 13) percentage of daily life spent using each language; 14) self-reported degree of foreign accent in each language; 15) self-reported frequency of being identified as a non-native speaker based on accent.

These responses were not combined into a composite measure; they were used qualitatively as a reference to ensure participants met inclusion criteria and to contextualize any potential variation seen in responses between speakers.

3.4 Procedure

In this study, I adopted 1) button-press and 2) eye-tracking methodologies using a visual world paradigm to collect accuracy (button selection of the target or competitor), RT, and eye tracking information. During the experimental task, participants were asked to select the picture that matches the auditory item that was heard using a keyboard button press.

The experimental platform which was adopted in the study used Gorilla experimental builder (Anwyl-Irvine et al. [2020]), which is an online software. Gorilla experimental builder integrates Webgazer (Papoutsaki et al. [2016]) for use of the web camera for eye tracking. WebGazer works by isolating the webcam image of the eyes as detected by a facial features recognition algorithm, reducing this image to an eye feature vector, then supplying this information to the gaze prediction model (Papoutsaki et al. [2016]).

The decision to conduct the experiment online was necessary to access participant populations in Hawai'i while the research was conducted at the University of British Columbia in Vancouver, Canada. This approach was particularly practical given that the trilingual community is spread across eight Hawaiian islands.

Previously cited advantages to Webgazer-based eye tracking, as opposed to more standard infrared eye trackers include its cost-effectiveness, relative portability, and ability to reach a larger number of people due to its ubiquity and ability to be done at home rather than a central testing location (e.g. Skovsgaard et al. [2011]). However, some disadvantages include that online eye trackers are less accurate and precise compared to infrared eye trackers. This is due to increased sensitivity to non-meaningful movement, increased sensitivity to low light conditions, and relatively low frame rate and resolution compared to infrared counterparts. Nevertheless, results using Webgazer are considered comparable to infrared eye tracking options for visual world paradigm (VWP) according to prior studies (e.g. Skovsgaard et al. [2011], Vos et al. [2022]). Ample information is collected during each experimental trial to indicate whether the trial should be excluded from analysis, including information on whether no fixations were detected for an entire trial, and face-configuration accuracy data. Face-configuration information provides

3.4. Procedure

information on the accuracy of the face recognition image captured, which supplies information to the eye tracking estimation algorithm. This score ranges from 0-1, where 0 refers to “no eye detected” and 1 refers to “clear eye detected.” The control of these variables based on previous work has led to replicable results from VWP studies using infrared eye trackers (e.g. Skovsgaard et al. [2011], Vos et al. [2022]).

Prior to the experimental component of the study, participants completed a consent form, a headphone quality check (Woods et al. [2017]), the language background questionnaire, and the eye-tracker calibration. All pre-experimental components were integrated into the online platform and had to be completed before participants could continue to the experimental task.

Previous research has shown that webcam-based eye tracking is more vulnerable to reduced accuracy due to variations in lighting and head movements (Vos et al. [2022]). To address this, frequent recalibrations were incorporated into the experimental design. The calibration is a built-in component in the Gorilla Experimental software, and consisted of looks to circles at different positions on the screen for 10 seconds each. If the circle was appropriately glanced at, a number appearing in the circle will count down from 10 to 0, then a new circle appeared at a different part of the screen and the process continued until key parts of the screen were calibrated with Webgazer. If a participant’s gaze could not be read or the instructions were not followed, the calibration ended and on-screen guidance was displayed, including troubleshooting tips such as increasing room brightness, removing glasses, or cleaning the webcam. Every 12 trials, before the scheduled eye-tracker recalibration, participants were presented with a rest screen, giving them the option to take a short break. After the break (or immediately after a trial if they chose not to pause), participants completed the eye-tracker calibration before proceeding with the next set of trials. This procedure allowed breaks to be smoothly incorporated by taking advantage of the scheduled recalibration.

Participants first viewed instructions explaining how to complete the eye-tracker calibration. The study then began with a calibration procedure, which was repeated every 12 trials throughout the study. This interval was chosen based on previous work (e.g., Vos et al. [2022]), which reported relatively high rates of decay in gaze prediction accuracy over the course of an experiment and recommended recalibrating every 12 trials.

Following the initial calibration, instructions were provided to participants, “to select the word displayed on the left please press the “z” key. To select the word displayed on the right please press the “m” key. After completing these instructions, participants proceeded to the experimental

trials. Each trial began with the presentation of a fixation symbol (“+”) which remained on the screen for 200ms. Next, participants were presented with two visual primes: one target and one competitor, displayed on either the right or left sides of the screen. The occurrence of the target (correct) word and competitor (incorrect) word were counterbalanced for where they occurred on the screen (right or left sides of the screen). Next, the onset of the auditory stimulus was presented simultaneously with the visual stimuli. Each auditory stimulus consisted of a phrase instructing participants to select a particular target word. Frame sentence-target pairings occurred in one of four conditions: Pidgin-to-Pidgin (unilingual Pidgin), English-to-English (unilingual English), Pidgin-to-Hawaiian (code-switched Pidgin-to-Hawaiian), or English-to-Hawaiian (code-switched English-to-Hawaiian). At any point after the onset of the audio, participants could press a key corresponding to the word presented, allowing responses to occur before the phrase had finished playing. This design enabled participants to respond as soon as they could distinguish the target from the competitor. However, responses made prior to phrase completion were later excluded from analysis (see Analysis and Results section for details).

At any time in the trial, the participant responded by pressing the “m” or “z” key. Once a key was pressed, a new trial began. As the participant completed the trials, a progress bar displayed at the top of the screen indicated how much of the study was completed. Button responses (“z” or “m” keys) were recorded, along with RT for each button press. Only one answer per trial was allowed.

Eye tracking information consisted of: x,y zone coordinates of gaze locations given in chronological order within each trial, fixation proportions, as well as face-configuration information. The x,y zone coordinates of gaze locations given in chronological order within each trial provide time-course data of fixations either at the right or left image from the moment of the display of the images until the button press response has been made. The units of x,y zone coordinates were measured in pixels.

3.5 Analysis and Results

Data aggregation, visualizations, and statistical analyses for all experiments in this dissertation were carried out using R (R Core Team [2021]).

The 16 participants that qualified for the study performed at an accuracy rate of 89.6%. All excluded participants either did not report their highest level of Hawaiian class completed correctly (in three cases, it appeared that

3.5. Analysis and Results

the participant provided nonsense answers to all questions) or reported a level below completion of Hawaiian 102. A more detailed overview of the high rate of underqualified participants is provided in the Discussion section.

No participants who met the initial inclusion criteria were excluded on the basis of low accuracy. All included participants achieved accuracy rates of at least 80%, indicating that they met the required standards for both language proficiency and task engagement.

We begin with an analysis of mean accuracy for qualifying participants, summarized in Table 3.3 and depicted in Figure 3.4, in order to provide an overview of performance across code-switching and frame language conditions. As shown in Table 3.3 and Figure 3.4, mean accuracy appears marginally higher in the Hawai'i English frame language than in the Pidgin frame language across both code-switching and unilingual conditions. A larger difference in accuracy is apparent between code-switching conditions, such that accuracy is higher in the unilingual conditions and reduced in the code-switched condition for both frame languages. Unilingual conditions are near 100% mean accuracy, while mean accuracy for code-switch condition centers about 80%. All visualizations were made using the tidyverse package (Wickham et al. [2019]) in R.

Frame Language	Code-switch Condition	Mean Acc	SD Acc
English	codeswitch	0.815	0.389
English	unilingual	0.991	0.096
Pidgin	codeswitch	0.794	0.404
Pidgin	unilingual	0.986	0.117

Table 3.3: Mean accuracy and standard deviation by frame language and code-switching pattern. Accuracy reflects proportion of correct responses across trials.

3.5. Analysis and Results

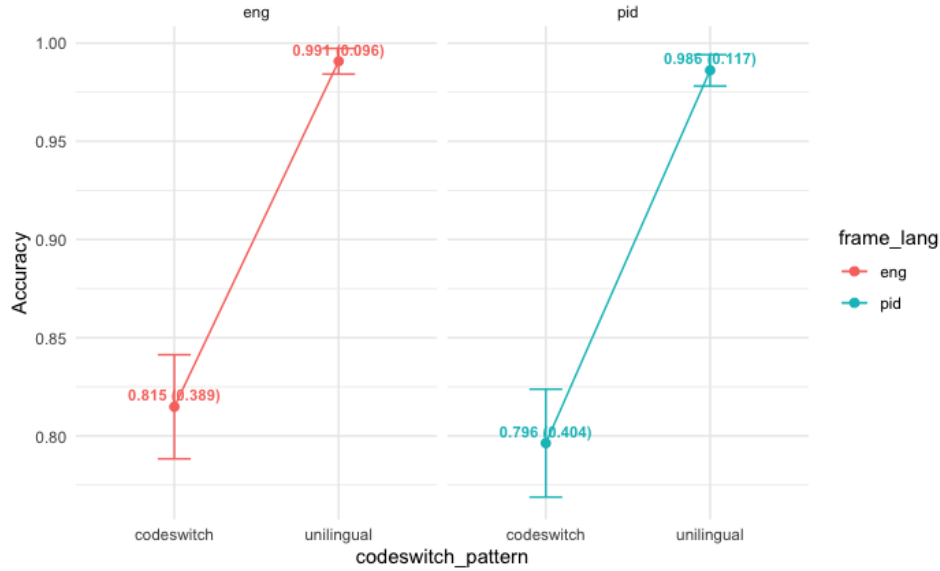


Figure 3.4: Mean proportion of correct responses (Accuracy) (0 = incorrect, 1 = correct) by code-switch condition, separated by frame language for participants among qualifying participants. Points represent condition means, lines connect means across conditions within each panel, and error bars indicate ± 1 standard error. Numeric labels show mean with standard deviation in parentheses.

Figure 3.5 depicts individual variation in accuracy among all 16 qualifying participants. The x-axis depicts frame language being Hawai'i English or Pidgin and the y-axis measures mean accuracy. The figure reveals that not all participants performed at ceiling (near 100% accuracy) overall across language and code-switched conditions. Accuracy scores suggest that listeners were more accurate on unilingual trials. Preliminary patterns suggest that accuracy varies depending on the frame language and code-switched condition. Furthermore, the slope of the line connecting performance on the code-switched trials varies considerably across participants. Notably, no participants performed at 100% across all conditions and frame languages. This will also be discussed further in the Discussion section.

3.5. Analysis and Results

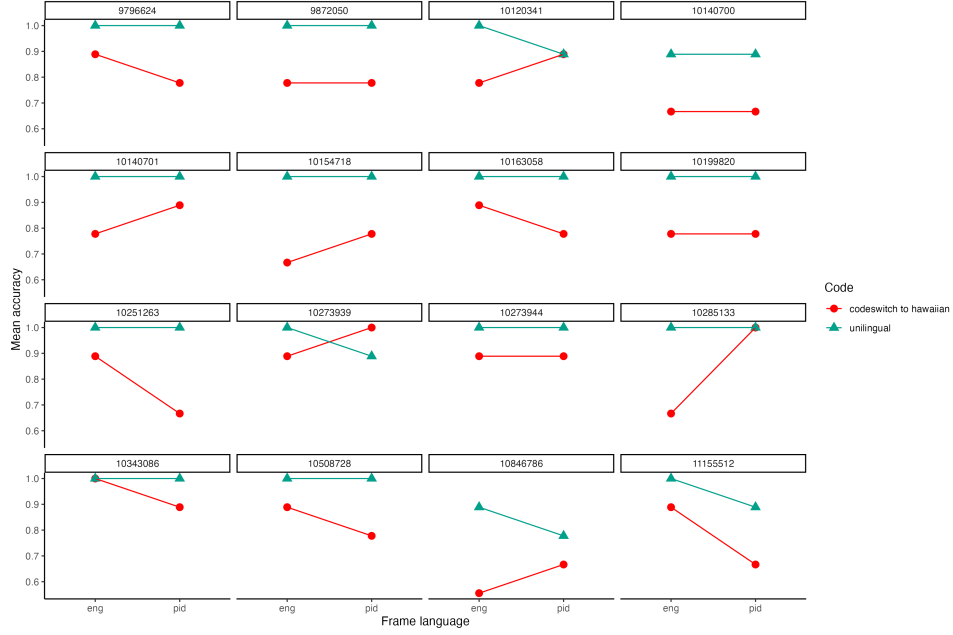


Figure 3.5: Individual variation in accuracy by frame language and trial type without removal of participants. Each point represents a participant’s mean accuracy for a given combination of frame language, Hawai’i English or Pidgin (on left, and right, respectively) and trial type (unilingual or code-switched to Hawaiian).

Tables 3.4 and 3.5 report mean accuracy across participants, along with the number of remaining trials for each participant after excluding trials with response times outside the 1400-5000 ms window. (To accommodate the length of the data, participant information is presented in two tables: Table 3.5 includes participants 9796624 to 10163058 and Table 3.6 includes participants 10199820 to 11155512.) These tables make individual level variation explicit and document differences in data availability across conditions after response time filtering. Because accuracy does not reach ceiling uniformly across participants or conditions, participant level means along with remaining trials provide important context that would be obscured by aggregate summaries alone.

3.5. Analysis and Results

Participant ID	Code-switch Condition	Frame Language	Mean Accuracy	Number of Remaining Trials
9796624	codeswitch to hawaiian	eng	0.89	9
	codeswitch to hawaiian	pid	0.75	8
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
9872050	codeswitch to hawaiian	eng	0.88	8
	codeswitch to hawaiian	pid	0.78	9
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10120341	codeswitch to hawaiian	eng	1.00	7
	codeswitch to hawaiian	pid	0.88	8
	unilingual	eng	1.00	9
	unilingual	pid	0.89	9
10140700	codeswitch to hawaiian	eng	0.67	9
	codeswitch to hawaiian	pid	0.71	7
	unilingual	eng	1.00	8
	unilingual	pid	0.89	9
10140701	codeswitch to hawaiian	eng	0.78	9
	codeswitch to hawaiian	pid	0.89	9
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10154718	codeswitch to hawaiian	eng	0.67	6
	codeswitch to hawaiian	pid	0.71	7
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10163058	codeswitch to hawaiian	eng	1.00	2
	codeswitch to hawaiian	pid	0.80	5
	unilingual	eng	1.00	9
	unilingual	pid	1.00	6

Table 3.4: Mean Accuracy for Participants 9796624 to 10163058 by Code-switch Condition and Frame Language, including number of trials completed after removal of trials with response times outside 1400–5000 ms.

3.5. Analysis and Results

Participant ID	Code-switch Condition	Frame Language	Mean Accuracy	Number of Remaining Trials
10199820	codeswitch to hawaiian	eng	0.88	8
	codeswitch to hawaiian	pid	0.88	8
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10251263	codeswitch to hawaiian	eng	0.89	9
	codeswitch to hawaiian	pid	0.75	8
	unilingual	eng	1.00	9
	unilingual	pid	1.00	7
10273939	codeswitch to hawaiian	eng	0.89	9
	codeswitch to hawaiian	pid	1.00	9
	unilingual	eng	1.00	9
	unilingual	pid	1.00	8
10273944	codeswitch to hawaiian	eng	1.00	5
	codeswitch to hawaiian	pid	1.00	7
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10285133	codeswitch to hawaiian	eng	0.67	9
	codeswitch to hawaiian	pid	1.00	9
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10343086	codeswitch to hawaiian	eng	1.00	9
	codeswitch to hawaiian	pid	0.89	9
	unilingual	eng	1.00	9
	unilingual	pid	1.00	9
10508728	codeswitch to hawaiian	eng	0.88	8
	codeswitch to hawaiian	pid	0.75	8
	unilingual	eng	1.00	9
	unilingual	pid	1.00	6
10846786	codeswitch to hawaiian	eng	0.83	6
	codeswitch to hawaiian	pid	0.67	6
	unilingual	eng	0.88	8
	unilingual	pid	0.88	8
11155512	codeswitch to hawaiian	eng	1.00	4
	codeswitch to hawaiian	pid	1.00	4
	unilingual	eng	1.00	9
	unilingual	pid	1.00	5

Table 3.5: Mean Accuracy for Participants 10199820 to 11155512 by Code-switch Condition and Frame Language, including number of trials completed after removal of trials with response times outside 1400–5000 ms.

To assess population-level accuracy empirically, the data were analyzed using a Bayesian multilevel logistic regression model implemented in the `brms` package in R (Bürkner [2017]). The model estimated accuracy as a function of code-switch condition, frame language, and their interaction. Accuracy was modeled using a binary Bernoulli distribution. The fixed effects structure included the main effects of code-switching pattern and frame language, as well as their interaction. To account for the repeated measures design, the model included by-participant random intercepts and random slopes for both predictors, allowing the effects of code-switching and frame language to vary across participants. In addition, a by-item random intercept

was included to capture variability attributable to differences across stimuli. Weakly informative priors were specified for all population-level effects using a normal distribution centered at a mean of 0 and a standard deviation of 5, reflecting the expectation that effects were plausibly small to moderate in size but allowing sufficient flexibility for the data to inform the posterior estimates (Gelman et al. [2017]).

The model was estimated using four Markov Chain Monte Carlo chains, each running for 2000 iterations with 1000 iterations discarded as warm-up. To improve sampling stability and reduce the risk of divergent transitions, the adapt delta parameter was set to 0.99. Convergence was assessed using standard diagnostics including Rhat values and effective sample sizes.

Results are presented in Table 3.6. The Bayesian multilevel logistic regression indicated a strong effect of code-switching on accuracy. Accuracy was substantially higher in the unilingual condition relative to the code-switched condition, as reflected in a large positive estimate for the unilingual contrast (Estimate = 3.84, 95% CI [1.60, 6.65]). In contrast, there was no clear evidence for a main effect of frame language, with the estimate for Pidgin relative to Hawai'i English centered near zero (Estimate = -0.04, 95% CI [-1.45, 1.40]) and a wide credible interval encompassing zero. Likewise, the interaction between code-switching and frame language was not credibly different from zero (Estimate = -0.31, 95% CI [-3.47, 2.64]), indicating that the effect of code-switching did not reliably differ across frame languages. Rhat values and effective sample sizes (Bulk ESS and Tail ESS) indicate satisfactory model convergence and sampling efficiency. Overall, these results suggest a robust cost of code-switching on accuracy that is consistent across frame languages, with no evidence for a reliable independent effect of frame language or interaction between the two predictors.

Parameter	Est.	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept	2.18	0.55	1.21, 3.37	1.00	1204	1828
Unilingual Condition	3.84	1.27	1.60, 6.65	1.00	1415	1668
Pidgin (Frame)	-0.04	0.72	-1.45, 1.40	1.00	1115	1754
Unilingual $\times$ Pidgin	-0.31	1.54	-3.47, 2.64	1.00	1652	2255

Table 3.6: Posterior estimates from the Bayesian multilevel logistic regression model predicting accuracy. The table reports fixed effects for code-switching pattern, frame language, and their interaction, as well as variance components for participant and item level random effects. Estimates are presented as posterior means with standard errors and 95% credible intervals.

Following analysis of accuracy, the data underwent further filtering prior

to analysis of reaction time. First, trials with response times under 1400 ms or over 5000 ms were excluded. Trials with response times under 1400 ms were removed because this corresponded to the duration of the sentence frame preceding the onset of the target item; responses within this window starting prior to the target item were excluded as they could not reflect accuracy for the target. Responses exceeding 5000 ms were excluded as they likely reflected distraction, confusion, or lapses in attention, rather than genuine task performance. While the 5000 ms threshold is somewhat arbitrary, similar exclusion criteria have been adopted in previous studies, where responses 4-6 standard deviations above the mean are treated as outliers unlikely to reflect task-relevant processes (e.g. Miller [2024]). Trials outside the 1400-5000 ms reaction time window were not coded as incorrect but were simply filtered out. This ensured that only responses within the defined reaction-time range were included in the analysis, without labeling such responses as inaccurate. After applying these reaction time criteria, 1 trial overall (0.02%) was excluded for responses made before 1400 ms, and 65 trials overall (11%) were excluded for responses made after 5000 ms.

Following these additional exclusions, eye-tracking data were examined but ultimately excluded from the present analysis, as the results were inconsistent and uninterpretable. Eye-tracking data revealed unexpectedly high proportions of fixations to the “no zone” and very few to the target image, including in follow-up verification trials where the researcher deliberately fixated on the correct item. The source of this issue remains unclear. Given participants’ behavioral responses in total averaging well over chance, it is likely these irregular eye-tracking patterns are attributed to data wrangling errors or potential tracking miscalculations during data collection, rather than meaningful reflections of participant attention or gaze behavior as a whole. An additional possibility is that participants could visually distinguish the images without needing to move their eyes substantially, for example due to differences in visual area or prominence associated with the singular versus plural distinction. This, combined with the nature of the required response (a keyboard press rather than a spatially directed mouse movement), may also have reduced overt eye movements toward the target. Unlike tasks involving visually guided motor actions such as mouse clicks, where gaze and action are spatially aligned, the present task required visually guided selection toward a keyboard response rather than a location on the screen, which may further explain the weak correspondence between gaze and target selection. Data wrangling errors possibly include misclassification of gaze points due to spatial binning issues (i.e., incorrect mapping of raw gaze coordinates to areas of interest). Tracking miscalculations po-

tentially include calibration failures or improper identification of gazes (i.e. over-sensitivity of the eye-tracking device: capturing small, involuntary eye or head movements as meaningful gaze shifts to “no zone”).

At earlier stages of analysis, the eye-tracking patterns appeared superficially plausible: a substantial proportion of fixations fell into the “no zone,” which is not uncommon in visual world studies where participants often blink, shift gaze between images, or briefly disengage from the screen. It was therefore not immediately evident that these data reflected systematic error rather than natural variability. Only after conducting more detailed cross-checks, such as comparing gaze patterns to participants’ behavioral accuracy, which remained well above chance, and running verification trials in which the experimenter deliberately fixated on targets, did it become clear that the observed fixation distributions could not plausibly reflect genuine attentional behavior. Previous work suggests that behavioral accuracy does not always correspond closely with eye-movement measures. In Goodrich Smith and Hudson Kam (Goodrich Smith and Hudson Kam [2012]), for example, accuracy was not consistently aligned with gaze patterns, suggesting that eye movements do not always reflect underlying decision processes. The delayed recognition of this issue reflects the iterative nature of quality control: what initially seemed like participant behavior was ultimately revealed to be a probable systematic error once multiple lines of evidence converged. Accordingly, eye-tracking data were excluded from statistical analyses but were retained for further investigation in future work.

Subsequent statistical analyses therefore focused solely on response time (RT). For the RT analysis, only trials in which participants responded correctly were retained. Restricting the analysis to correct responses helps ensure that reaction times reflect meaningful cognitive processing of the target stimuli. Including inaccurate responses could introduce noise, as RT for incorrect responses may arise from factors unrelated to the processes of interest (e.g., likely due to uncertainty, or guessing). After excluding inaccurate responses, along with trials previously filtered out with response times outside 1400–5000 ms, some participants had fewer than half of trials (≤ 4) in certain code-switch condition $\times$ frame language combinations. Such low trial counts could reduce statistical generalizability and compromise the reliability of parameter estimates in the subsequent Bayesian analysis. This step resulted in 2 participants being excluded: one participant was excluded due to having 4 or fewer trials remaining in the Hawai‘i English-to-‘ōlelo Hawai‘i condition, and a second participant was excluded due to having 4 or fewer trials remaining in both the Pidgin-to-‘ōlelo Hawai‘i condition and the Hawai‘i English-to-‘ōlelo Hawai‘i condition.

After applying all exclusion criteria, 14 participants remained for statistical analysis. With this final dataset established, we first examined mean reaction times for accurate responses at the group level (Figure 3.6). Three trends emerge: (1) unilingual trials appear to have been identified faster overall than code-switched trials (the distribution of unilingual trials is lower in mean RT than distribution of code-switched trials), (2) Hawai'i English unilingual trials appear to have been identified faster overall than Pidgin unilingual trials (the distribution of Hawai'i English unilingual trials is lower in mean RT than distribution of Pidgin unilingual trials), and (3) in code-switched trials this pattern appears to exhibit the opposite effect where Pidgin trials appear to have a lower mean RT than Hawai'i English trials. Overall, these patterns are consistent with the hypothesis that switch costs are lower for Pidgin than for Hawai'i English. The magnitude of code-switching costs is greater for Hawai'i English-to-Hawaiian than for Pidgin-to-Hawaiian, as reflected in the relative difference between unilingual and code-switched reaction times across the two languages. Together, these observed patterns demonstrated visually strongly suggest an interaction between frame language and code-switch condition, in which the relative processing advantage of Hawai'i English and Pidgin differs across unilingual and code-switched contexts. This pattern appears to reflect a classic suppression effect, in which greater inhibition is required for Hawai'i English than for Pidgin switches to 'ōlelo Hawai'i, consistent with English being the dominant language.

3.5. Analysis and Results

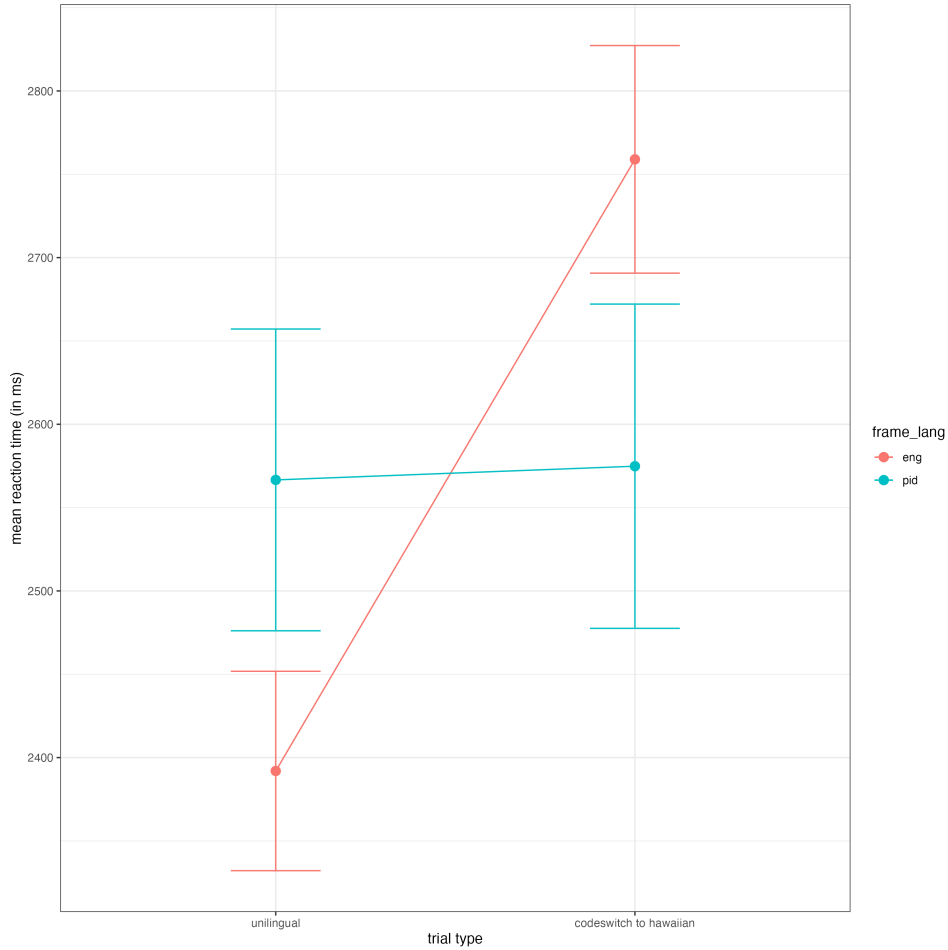


Figure 3.6: Mean response times (in milliseconds) by trial type (unilingual or code-switched speech) and frame language (Hawai'i English or Pidgin) from empirical results. Each point reflects the grand meanRT across participants for a given condition, with error bars representing the standard error of the subject-level means.

We now examine the data at the participant level, which provides insight into individual differences and reveals how each participant's reaction times contribute to the group-level patterns observed above. Figure 3.7 depicts individual variation of mean RT across trial type and frame language. At the individual-level, considerable variability is evident. Most participants

3.5. *Analysis and Results*

(10/14) were faster on average at identifying words in unilingual frames than in code-switched frames, but a minority (4/14) showed the opposite pattern; these participants are indicated near the participant number with a red triangle. Similarly, while most participants (11/14) responded faster to Hawai'i English unilingual frames, a subset (3/14) showed faster identification of Hawaiian items in code-switched trials when presented in Pidgin frames, marked with a blue circle. Finally, among participants who were faster in Hawai'i English unilingual trials, most (9/11) flipped to faster responses in Pidgin frames for code-switched trials, but two participants exhibited an overall faster response to Hawai'i English, highlighted with a purple square. Interpretation of such variation will be discussed further in the Discussion section. In general, substantial individual variability seen here suggests that individual strategies likely influenced performance, highlighting that group-level effects may not fully capture the rich variability inherent in trilingual language processing.

3.5. Analysis and Results

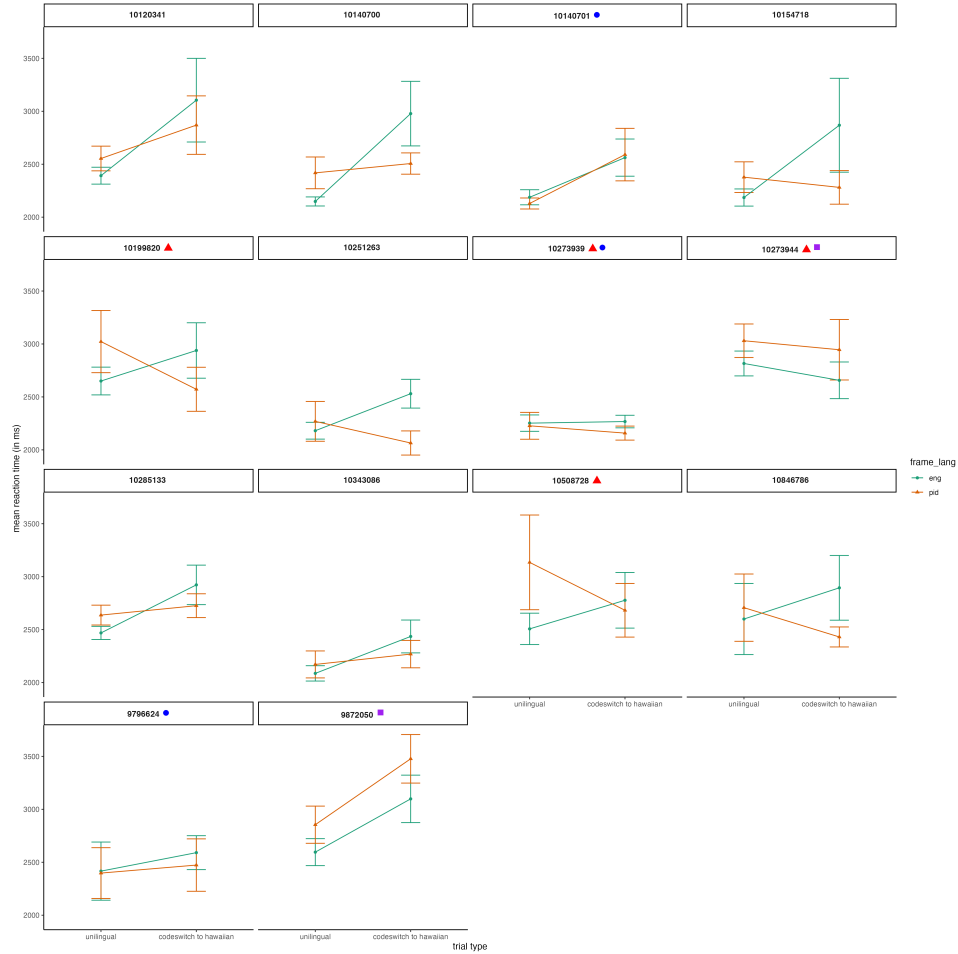


Figure 3.7: Individual Variation in Mean Reaction Times. Mean reaction times (RTs) by trial type and frame language for each participant. Error bars represent the standard error of the mean, capturing trial-level variability in response times within each condition. While group-level trends are visible, substantial individual variability is observed, particularly across language conditions. Participants who did not follow group-level patterns are marked with colored shapes next to their participant ID: red triangles indicate participants faster in code-switched than unilingual trials overall, blue circles indicate participants faster in Pidgin unilingual trials compared to Hawai'i English unilingual trials, and purple squares highlight participants who were faster in Hawai'i English unilingual trials and also faster in Hawai'i English code-switched trials.

At the individual level, self-reported proficiency from the Language Background Questionnaire varied substantially by language. English proficiency was consistently rated as high across participants, whereas ‘ōlelo Hawai‘i proficiency was generally rated low. In contrast, Pidgin proficiency showed considerable variability across individuals. This variability in Pidgin proficiency may contribute to the observed variability in accuracy and reaction time across participants.

Building upon these observations at group and individual levels, the statistical reliability of group effects was evaluated using a Bayesian mixed-effects model. To assess whether frame language (Hawai‘i English or Pidgin) facilitated processing of ‘ōlelo Hawai‘i lexical items, reaction times (RTs) for correct responses were modeled using a Bayesian mixed-effects model implemented in *Stan* via the *brms* package (Bürkner [2017]) in *R* (R Core Team [2021]). RTs were log-transformed. The model formula was:

$$\begin{aligned} \text{RT} \sim & \text{frame-language} \times \text{codeswitch-condition} \\ & + (1 + \text{frame-language} \times \text{codeswitch-condition} \mid \text{participant}) \quad (3.1) \\ & + (1 \mid \text{item}) \end{aligned}$$

This model predicts RTs as a function of: (1) Frame Language (Pidgin versus Hawai‘i English [reference level]), (2) Code-Switching Condition (code-switching to ‘ōlelo Hawai‘i versus unilingual [reference level]), (3) their interaction, (4) varying intercepts and slopes by participant, and (5) varying intercepts by item. Posterior fixed-effect estimates are interpreted relative to reference levels (Hawai‘i English for frame language, unilingual for code-switching). RTs were modeled using a lognormal likelihood. In a lognormal model, the linear predictor is on the log scale, and fixed-effect estimates therefore reflect changes in log RT relative to the reference levels (Hawai‘i English for frame language, unilingual for code-switching). Estimates reflect changes in log RT, which can be exponentiated to interpret as proportional or percent changes in raw RT.

Weakly informative priors were used in the Bayesian model to support convergence and regularize estimates, a standard practice in Bayesian modeling (e.g., McElreath [2020]). Specifically, a normal distribution with $\mu = 0$ and $\sigma = 0.25$ was applied to fixed-effect coefficients, a Cauchy distribution with $\mu = 0$ and $\sigma = 0.25$ to group-level standard deviations (class = “sd”), and a Cauchy distribution with $\mu = 0$ and $\sigma = 0.25$ to the residual standard deviation (class = “sigma”). The model was run using Markov Chain Monte Carlo (MCMC) with 6,000 iterations per chain (1,000 warm-up, thinning

factor of 2), a fixed seed for reproducibility, and an adapt-delta of 0.95 to reduce divergent transitions.

The primary effect of interest for the current research question is the interaction between frame language and code-switching condition. An interaction would mean that the size of the code-switching cost depends on the frame language. A negative estimate with a 95% credible interval (CI) that excludes zero would indicate that switching from Pidgin to ‘ōlelo Hawai‘i incurs lower processing costs than switching from Hawai‘i English to ‘ōlelo Hawai‘i, supporting the hypothesis that Pidgin facilitates access to ‘ōlelo Hawai‘i relative to Hawai‘i English. Effects with 95% CIs excluding zero are considered meaningfully different from baseline, while intervals including zero indicate uncertainty or limited evidence (e.g., Kruschke [2021], Bürkner [2017], McElreath [2020]).

Simple effects of frame language and code-switching condition are also interpreted. A simple effect of frame language reflects baseline differences between Hawai‘i English and Pidgin trials at the reference level (unilingual), while a simple effect of code-switching condition reflects differences between unilingual and code-switched trials at the reference level (Hawai‘i English). Effects with 95% CIs excluding zero are considered meaningfully distinct. Ideally, no baseline difference between frame languages would be observed, indicating similar responses across Hawai‘i English and Pidgin trials. On the contrary, consistent with previous findings, code-switching is generally expected to slow responses (e.g., Kaushanskaya and Marian [2007], Wu and Thierry [2012]); thus, a simple effect of code-switching condition would align with such expectations.

Table 3.7 provides the full output of population (fixed-effect) estimates. Population-level (fixed) effects found a model intercept of 7.77 log units (Estimate = 7.77 log units, 95% CI [7.67, 7.86]), corresponding to a mean baseline reaction time of ≈ 2380 ms ($e^{7.77}$) in raw reaction time for unilingual trials in Hawai‘i English frame.

Code-switching added about 0.13 log units to reaction time (Estimate = 0.13, 95% CI [0.02, 0.23]), corresponding to an increase of roughly 13% (≈ 317 ms) relative to unilingual trials in the Hawai‘i English frame, a credible slowing effect.

The effect of frame language was small and uncertain, with a 95% credible interval that included zero (Estimate = 0.04 log units, 95% CI [−0.07, 0.15]), corresponding to an increase of approximately 83 ms ($\sim 3.5\%$) relative to unilingual Hawai‘i English trials. Consistent with this uncertainty, the probability of direction indicated no evidence for a directional effect (PD = 0.77).

3.5. Analysis and Results

The interaction between frame language and code-switching suggested negligible evidence for reduced switching costs within the Pidgin frame relative to the Hawai'i English frame (Estimate = -0.11 log units, 95% CI $[-0.26, 0.03]$, ($\approx 11\%$ decrease, ~ 268 ms). Although the credible interval overlapped zero, the probability of direction indicated moderate directional evidence for a negative interaction (PD = 0.94), suggesting that reduced switching costs in the Pidgin frame were more likely than not, albeit with substantial remaining uncertainty.

The model converged well, with Rhat = 1.00 and high ESS across parameters, indicating stable posterior estimates.

Predictor	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept	7.77	0.05	7.68, 7.87	1.00	3889	6163
Frame Language (Pidgin vs. Hawai'i English)	0.04	0.06	-0.07, 0.15	1.00	4388	6529
Code-Switching Condition (Unilingual vs. Code-switch)	0.13	0.05	0.02, 0.23	1.00	4750	7524
Interaction (Frame Language x Code-switch)	-0.11	0.07	-0.26, 0.03	1.00	4599	6454

Table 3.7: Population-level (fixed-effect) estimates from the Bayesian mixed-effects model estimating log reaction times (log RTs). 95% credible intervals (CI) are shown alongside posterior mean estimates.

Overall, code-switching to Hawaiian significantly increased reaction times, and frame language had no reliable effect. These two findings align with our expectations: code-switching incurs costs compared to unilingual sentences, and there was no overall language bias, as reaction times did not systematically differ based on whether the frame language was Pidgin or Hawai'i English. The interaction between frame language and code-switching suggested weak evidence for modulation of switching costs, with responses showing a tendency toward reduced switching costs in the Pidgin frame relative to the Hawai'i English frame. Although the credible interval overlapped zero, the probability of direction indicated moderate directional evidence that switching costs were smaller in the Pidgin frame, with substantial uncertainty remaining.

The Bayesian mixed-effects model accounts for individual differences through participant-level random effects, estimating participant-specific variation in baseline response times and the effects of the experimental manipulations while retaining the full item-level structure of the data. To complement this analysis, we calculated a switch cost (SC) for each participant by summarizing reaction times for each condition. A drawback of this approach is that it reduces each participant's responses to a single summary value, collapsing item-level variability as well as the uncertainty and hierarchical structure captured by the mixed-effects model. However,

this approach allows for direct comparisons of switching costs across participants by providing a simplified measure of individual switching behavior. As such, this follow-up analysis allowed us to quantify the magnitude of the slowing associated with switching into Hawaiian separately for Hawai'i English and Pidgin frame language trials. Switch cost (SC) was defined as: $SC = \text{MeanRT}_{\text{unilingual}} - \text{MeanRT}_{\text{code-switch}}$, with mean RTs for code-switched trials subtracted from mean RTs for unilingual trials, calculated separately for each frame language. Mean switch costs were substantially larger in the Hawai'i English frame ($M = 367$ ms, $SD = 266$ ms) than in the Pidgin frame ($M = 8$ ms, $SD = 313$ ms).

With these SC values, a one-tailed paired-samples t-test was conducted to compare SC values between the Hawai'i English and Pidgin frame conditions across participants. The test was one-tailed (function set to alternative = "less"), based on the directional hypothesis that switching from Hawai'i English incurs a higher cost than switching from Pidgin (i.e., $SC_{\text{English}} > SC_{\text{Pidgin}}$). The resulting p-value was used to determine whether the observed difference in switch costs was statistically significant. A p-value below 0.05 was interpreted as evidence that the switch cost is significantly higher for Hawai'i English than for Pidgin. In addition to statistical significance, effect size was estimated using the Cohen's d value. The one-tailed paired-samples t-test was conducted using R's built-in `t.test` function and Cohen's d was calculated with the `cohensD` function from the `lsr` package (Navarro [2021]).

The one-tailed paired-samples t-test results revealed a significant difference between conditions, $t(13) = 4.01$, $p < 0.001$, as p was less than the conventional α level of 0.05, suggesting that code-switching costs were lower for Pidgin-to-Hawaiian switches than for English-to-Hawaiian switches. The mean difference in switch costs was 359 ms, with a one-sided 95% confidence interval excluding zero $[200 \text{ ms}, \infty]$, indicating reliably greater slowing when switching to Hawaiian from Hawai'i English than from Pidgin.

Cohen's d value ($d=1.07$) is interpreted as a large effect size ($|d| > 0.8$), where conventionally less than 0.2 represents a small effect, between 0.2 and 0.8 represents a medium effect and larger than 0.8 represents a large effect (Cohen [1988]). This indicates a large effect size for the difference in processing behavior between Pidgin and Hawai'i English speakers.

At the individual level, many participants appear to have lower code-switching costs in Pidgin compared to Hawai'i English, visually represented in Figure 3.8 by lines that slope downward from left (English) to right (Pidgin). Some variation is evident, with some participants showing relatively similar costs across both languages (flatter slopes), and one participant ex-

3.5. Analysis and Results

hibiting notably higher costs for Pidgin than for Hawai'i English (slope upward from left (English) to right (Pidgin)). Taken together, the general pattern suggests that most participants tended to exhibit lower code-switching costs in Pidgin compared to Hawai'i English, though this was not the case for all individuals.

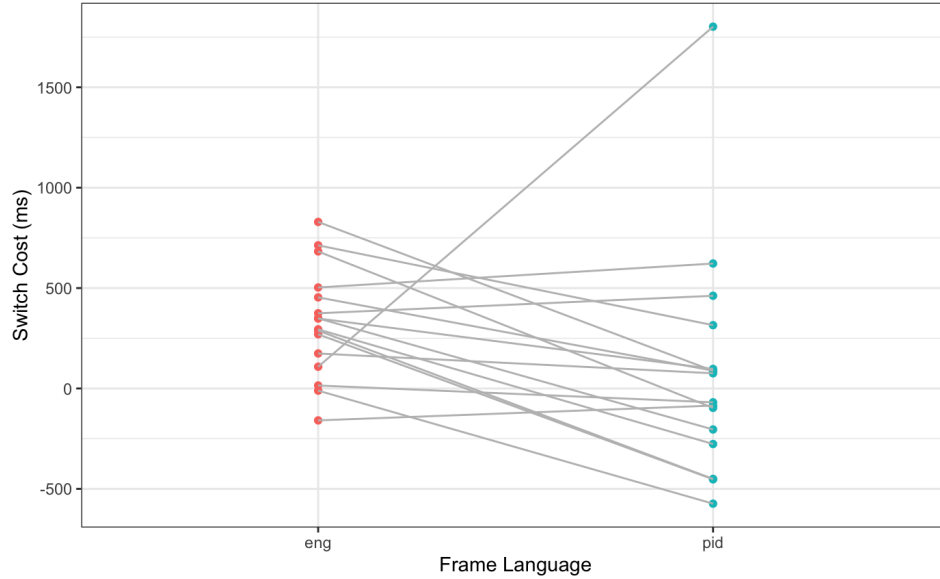


Figure 3.8: Individual code-switching costs. Each point represents the mean response time difference between code-switched and unilingual trials for each participant. Positive values indicate slower reaction times for code-switched trials relative to unilingual trials, while negative values indicate the opposite.

Note that in Figure 3.8, one participant (10199820) stands out as having substantially higher code-switch costs for Pidgin than for Hawai'i English. There is no indication from their language background questionnaire of any factors that would predict this pattern of slowing in the Pidgin frame. This participant reported maximum proficiency (10/10) in both Pidgin and English and indicated that Pidgin and English were acquired from birth.

3.6 Discussion

The primary goal of this study was to determine whether the language preceding a switch affects access to the target language. Consistent with prior work suggesting that code-switching costs can be shaped by language experience (e.g., Cheng and Howard [2008]; Beauvillain and Grainger [1987]; Caramazza and Brones [1980]), I examined whether frame language modulates the magnitude of switching costs.

In the analysis of accuracy, robust code-switching costs were observed when comparing to unilingual phrases in either Pidgin or Hawai'i English. In contrast, no credible differences in accuracy were found between frame languages, and there was no evidence of an interaction between frame language and code-switching condition. Thus, with respect to accuracy, Pidgin and Hawai'i English frame languages show comparable performance in unilingual phrases as well as when code-switching to Hawaiian.

One interpretation of the effect of accuracy found across code-switch condition is that the act of code-switching itself introduces processing costs, leading to reduced accuracy. However, because the design does not include a unilingual Hawaiian condition, an alternative explanation is that lower proficiency in Hawaiian, relative to Pidgin and Hawai'i English, contributes to the higher error rates observed in the code-switched condition.

When considering frame language, Pidgin and Hawai'i English patterned similarly in terms of accuracy, with no credible differences observed between them. This suggests that, at least in terms of accuracy, neither frame language confers a systematic advantage or disadvantage in task performance. As a result, any differences observed in other measures between Pidgin and Hawai'i English cannot be straightforwardly attributed to baseline disparities in accuracy between the two frame languages. This pattern supports the interpretation that participants were able to engage with both frame languages in a similarly effective manner when identifying the correct item heard, despite potential differences in processing demands seen in reaction times.

In examining processing demands at the population level, the Bayesian mixed-effects model provided no credible evidence for an interaction between frame language and code-switching condition, as the credible intervals included zero. Thus, when averaging across participants, there was no evidence that frame language modulates code-switching costs in the current study. However, substantial variability across participants, evident in Figure 3.7, indicated marked differences in the magnitude and direction of code-switching effects. To minimize this variability, we calculated individual

switch costs for each participant. This follow-up analysis revealed that, on average, Pidgin-to-Hawaiian switches were associated with lower switch costs than English-to-Hawaiian switches, suggesting reliably greater slowing when switching from Hawai'i English than from Pidgin for many individuals.

Overall, these findings indicate that the effect of frame language on code-switching costs varies widely across participants, including variability in magnitude, direction, and the presence or absence of an observable effect. Participant-level analyses, while more direct, are less robust than the Bayesian mixed-effects model because they aggregate responses, do not account for the hierarchical structure of fixed and random effects, and lack uncertainty estimates, making them more susceptible to noise. It remains unclear whether Pidgin or Hawai'i English as the preceding language consistently facilitates or impedes access to the target language, or whether any observed effects are strong, reliable, and generalizable. Overall, the results provide limited evidence that Pidgin facilitates access to 'ōlelo Hawai'i at the population level. However, participant-level analyses suggest that, for some individuals, having Pidgin rather than Hawai'i English as the preceding language may reduce switch costs when switching to 'ōlelo Hawai'i.

To contextualize the switching effects, we examined whether baseline processing differed between frame languages in unilingual contexts. The absence of a credible simple effect of frame language in the Bayesian model indicates that participants' baseline processing in unilingual conditions was comparable for Pidgin and Hawai'i English. In other words, participants were not systematically faster when listening to unilingual Pidgin versus unilingual Hawai'i English sentences. This equivalence in baseline performance provides a clear reference point for assessing code-switching costs and their interaction with frame language. Had group-level differences in frame language been present, they could have reflected differences in familiarity, proficiency, or language dominance, complicating the interpretation of frame-related modulation of code-switching costs (e.g., Costa et al. [2006], Aparicio and Lavour [2014]).

In sum, at the group-level, three main empirical findings emerged: (1) a clear code-switching effect, with longer reaction times and lower accuracy on switch trials than on unilingual trials; (2) no overall effect of frame language, such that Hawai'i English and Pidgin did not differ in unilingual condition in response times nor in rates of accuracy; and (3) no strong evidence indicating an interaction between frame language and code-switching condition, meaning on average, switching from Pidgin to Hawaiian did not incur reliably lower code-switching reaction times nor lower accuracy than switching from Hawai'i English. These patterns, however, appear largely

complicated by individual variation. As shown in Figure 3.7, some participants responded faster to code-switched than unilingual trials, some showed consistently faster responses in either Pidgin or Hawai'i English, and some showed an interaction effect while others did not. This variability likely contributed to uncertainty in the model, including in estimates of the interaction between frame language and code-switched condition.

Contrary to the Bayesian analysis, the follow-up t-test suggests that Pidgin reduced switch costs compared to Hawai'i English, underscoring the potential role of frame language in shaping code-switching costs. We infer that the difference between the Bayesian and t-test results likely stems from the vast variability observed across participants. This perhaps suggests that a crucial factor to consider in designing future work on this research question is identifying the sources of participant-level variability. Identifying this source of variability could help control for confounds in continued research investigating code-switching in this language context.

It is important to acknowledge that the Bayesian mixed-effects model also accounts for individual differences, but does so differently than the follow-up t-test by including participant-level random slopes, allowing the estimated effect of condition to vary across participants. Thus, the absence of a credible group-level effect in the Bayesian model does not indicate that participants showed no sensitivity to the predicted pattern, but rather that there was substantial variability in the size and direction of these effects across individuals, increasing the variance at the group-level. In contrast to the Bayesian mixed-effects model, the t-test reduces within-participant variation to a single mean point for each language condition, which allows us to interpret the overall direction of the effect. However, because this approach collapses variability across items and reduces each participant's performance to a single data point (for which the mean switch cost is calculated), it does not retain the uncertainty or hierarchical structure captured by the mixed-effects model, nor does not account for variability within participants. Thus, the t-test result should be interpreted as evidence of a potential pattern in the data rather than as a stronger or contradictory result relative to the mixed effects model.

This substantial variation may stem from several contributing factors, including listeners' code-switching experience across the three languages being variable. This was seen in analysis of participant's self-reported age of acquisition and proficiency ratings (see section 3.2.1 Figures 3.2 and 3.3). Consistent with previous work showing that reductions in perceptual costs relate to listeners' code-switching experience (e.g., Cheng and Howard [2008]; Beauvillain and Grainger [1987]; Caramazza and Brones [1980]; Costa et al.

[2006]; Aparicio and Lavour [2014]), one interpretation of the current results is that participants' experience with code-switching across the three language pairs is variable. This interpretation holds limited certainty at the group level but shows clearer patterns when individuals are assessed relative to their own baselines. Previous work suggests that baseline processing behavior is shaped by one's naturalistic experience. Accordingly, some participants may have learned Hawaiian in a Pidgin-rich environment (e.g., a home setting with frequent code-switching), whereas others may have learned Hawaiian in a more Hawai'i English-dominant environment (e.g., a classroom with a strict English-only policy). Individuals' experiences with the three languages may vary considerably, connecting Pidgin and Hawaiian in a meaningful way for some but not for others.

Additionally, participants' individual experiences with using the three languages in a laboratory setting appear to have contributed to variation in baseline RT, potentially leading to inconsistencies at the group level. As emphasized by Blanco-Elorrieta and Pykkänen [2017], ecologically valid experimental paradigms are crucial for capturing the complexities of multilingual language processing. Because Pidgin and Hawaiian are not commonly used across all domains, participants' comfort with use of these languages in an experimental paradigm may have differed from their comfort with a more widely spoken language like English. Asymmetries in comfort across the three languages is supported by differences found in base RT in the model. This variation could have contributed to uncertainty in the model overall.

The stimuli used in the current study did not represent naturalistic code-switches to Hawaiian in either Pidgin-to-Hawaiian nor English-to-Hawaiian trials. As described in Materials, stimuli were recorded by having a speaker switch by reading from a script where the switch would occur. This scripted approach, while allowing experimental control, may have also reduced naturalness; the timing and placement of switches were externally imposed rather than emerging spontaneously by the speaker. The recordings thus excluded typical speech disfluencies or phonetic cues that signal a switch is about to occur, as described by studies such as Piccinini and Garellek [2014], Fricke et al. [2016], which allow listeners to anticipate an upcoming switch. In Shen et al. [2020], it was found that withholding anticipatory phonetic cues preceding a switch negatively affects code-switched recognition, slowing reaction times and reducing attention in eye tracking. By removing these natural cues, the stimuli may have increased the cognitive demands of processing the switch, making the task less representative of real-world code-switching. Thus, any reductions in code-switching costs based on language preceding the switch may have been diminished by the lack of naturalness in the stimuli.

Relatedly, work such as Clopper and Bradlow [2008] found that standard dialects are associated with higher intelligibility in laboratory settings than non-standard dialects, even when the non-standard dialect is the listener’s native variety. Given that Hawai’i English is considered the local standard variety, while Pidgin and Hawaiian are non-standard, the laboratory settings used in this study may have further reduced the likelihood of observing any Pidgin-switch benefits. The incorporation of more naturalistic settings in future work might potentially reveal such effects.

3.7 Conclusion

This study examined whether the language preceding a switch shapes access to Hawaiian, focusing on whether the tighter conceptual, contextual and typological link between Pidgin and Hawaiian might facilitate processing. The results highlight potential directions for further research on how trilingual speakers manage their languages. The tentative indication that Pidgin-to-Hawaiian switches may involve lower costs than Hawai’i English to Hawaiian switches, while the Bayesian mixed-effects model does not show an interaction effect, raises important questions about the relationship between these three languages. As noted in the discussion of potential sources of uncertainty, future work that incorporates more naturalistic stimuli and experimental settings, investigates variables contributing to population variability, and features more comprehensive piloting may help determine whether this potential facilitation effect persists when these factors are accounted for. In addition, this study prompts questions about how the three languages are used in code-switching across different real-life domains. The development of a corpus that captures patterns in production of co-use among Hawai’i English, Pidgin, and Hawaiian would provide a valuable resource for exploring these dynamics further and for grounding experimental findings in naturalistic usage.

Furthermore, this study raises new questions about how language dominance, habitual patterns of use, and sociolinguistic context interact to shape processing in this complex trilingual population. Future research examining factors such as comfort, exposure, and the status of each language in the community will be critical for clarifying these patterns and advancing our understanding of multilingual processing.

In the following chapter, we turn to a different paradigm, cross-lingual priming, to further probe the relationship between Pidgin and Hawaiian. This approach allows us to test whether Pidgin preceding Hawaiian words

in a priming task provides greater facilitation than Hawai‘i English, offering a complementary perspective on how these languages interact in processing.

3.8 Mo‘olelo

Returning to Hawai‘i as the center of this work allows these findings to be interpreted through the lens of the community questions that motivated this study.

A longstanding belief that continues to exist within the community is that Pidgin is harmful. The appearance and persistence of this stance is reinforced through positions held by institutional authorities (Furukawa [2018]). For instance, former Hawai‘i Governor Ben Cayetano consistently argued that Pidgin disadvantaged students. During his administration (1994–2002), he identified removing Pidgin from schools as one of his policy goals (Wong [2013]). Years later, in a 2015 interview, he reiterated this position, stating, “I think it does the kids a disservice if you allow them to continue to speak Pidgin” (Wong [2015]). As demonstrated in the Mo‘olelo chapter, the belief that Pidgin is *harmful* has become internalized by some community members, shaping how they understand the costs which potentially come with allowing the use of Pidgin.

The results from Experiment 1 challenge the belief that Pidgin is inherently harmful. The absence of a clear advantage for either Pidgin or English is itself an important empirical finding, as it provides evidence against claims that Pidgin inherently disadvantages its speakers, including claims made by individuals such as Ben Cayetano, whose positions have led these views to be widely accepted as fact. Neither Pidgin nor English poses processing costs when considering code-switched speech to ‘ōlelo Hawai‘i.

Thus, in the classroom setting, ‘ōlelo Hawai‘i kumu can rest assured that they may draw upon the use of either English or Pidgin as a teaching strategy when they see fit. The finding that neither English nor Pidgin poses processing costs can be interpreted as evidence that neither language disadvantages students in the ‘ōlelo Hawai‘i classroom.

Chapter 4

Experiment 2: Investigating Cross-Lingual Priming from Pidgin and English Primes to ‘ōlelo Hawai‘i Targets

4.1 Introduction

This chapter presents Experiment 2, a short-term cross-language primed lexical decision task. Translation priming was utilized in this study, a method for investigating multilingual processing in which exposure to a word in one language facilitates recognition of its translation equivalent in another language. Experiment 2 extends the findings of Experiment 1 by shifting the focus to priming, investigating whether Pidgin facilitates processing in ‘ōlelo Hawai‘i. In Experiment 2, we pose the research question, “Do results from a short-term cross-language auditory priming indicate a stronger link between Pidgin and ‘ōlelo Hawai‘i than between English and ‘ōlelo Hawai‘i?” This question examines whether patterns of asymmetric priming reflect differences in cognitive associations between Pidgin and ‘ōlelo Hawai‘i compared to Hawai‘i English and ‘ōlelo Hawai‘i. If priming effects vary, to what extent might associations between Pidgin and ‘ōlelo Hawai‘i differ from those between Hawai‘i English and ‘ōlelo Hawai‘i?

Priming paradigms have long been used in psycholinguistics to investigate how words and sounds are represented and organized in the mental lexicon, as well as the mechanisms underlying word recognition (e.g., Tulving et al. [1982]). A priming task often begins with the presentation of an initial stimulus (called the “prime”) followed by a subsequent stimulus (the “target”), with the two stimuli relating in form or meaning. A “priming effect” is said to occur when a prime facilitates the processing of a target, often measured in faster or more accurate responses. Early accounts of priming effects include Cattell (1888), who observed faster reading times when prime and

target words were associatively related within a single language, for example semantically related pairs such as cat–dog or hand–arm. Different types of priming effects have been identified in a unilingual paradigm including semantic priming (e.g., *bread* (prime) → *butter* (target)) (e.g., Neely [1991]), phonological priming (e.g., *bed* (prime) → *bad* (target)) (e.g., Grainger and Beauvillain [1987], Clopper and Walker [2017]), morphological priming (e.g., *teach* (prime) → *teacher* (target)) (e.g., Frost et al. [2000]), among others. The strength of a priming effect can provide insight into how closely linguistic units (i.e., words, sounds) are associated, what kind of information is stored in the mind, and the factors that influence access (Collins and Loftus [1975]). An absence of priming may reveal processing factors that impede access or suggest a weak degree of relatedness between the tested linguistic units (e.g., McNamara [2005]). Priming tasks can also demonstrate competing processes at play. For instance, primes and targets sharing the same sounds can boost activation but may also slow recognition due to lexical competition (Marslen-Wilson [1990], Radeau et al. [1995]). The interaction between different factors, for example at different linguistic levels such as sound and word, can impede access rather than facilitate it. Priming serves as a central tool in psycholinguistics, offering insight into the structure of linguistic representations in the mind and, crucially, how these representations are interconnected and mutually influence each other during processing (Meyer and Schvaneveldt [1971], Hutchison [2003]).

Cross-language priming is a type of priming that utilizes at least two languages: the prime and target are presented in different languages, and facilitation typically arises from shared form or meaning. A cross-language priming effect occurs when activation of a word in one language also activates its corresponding representation in another language. Translation priming is a form of cross-language priming, in which a word in one language facilitates access to the target item in another language based on approximate equivalence in meaning (e.g., Potter et al. [1984], Escudero et al. [2008], Chen and Ng [1989], Jiang and Forster [2001]). For instance, *squid* in English, *ika* in Pidgin, and *mūhe‘e* in Hawaiian all represent the mollusc with an elongated soft body, large eyes, and eight arms, and thus the three of these items can be referred to as translation equivalents.

Previous work has suggested that translation equivalents can elicit a priming effect via the activation of a shared conceptual link or *lemma*. Words that are translation equivalents like *squid*, *ika*, and *mūhe‘e* are representations connected to a shared underlying concept. The prime word from Language A will activate the shared concept, such that when Language B is encountered it is faster to access the item because it has been recently ac-

tivated. A typical translation priming task compares participant response times for translation equivalents with those for unrelated word pairs. For example, the reaction time for identifying the word *mūhe'e* following the word *squid* will be compared to that of the word *mūhe'e* following the word *chair*. An effect found in a cross-language priming task, such as in translation priming, may offer insight into how multiple language systems are interconnected in the multilingual mind. A stronger conceptual link may indicate that lexical items across languages are more closely associated in the multilingual mental lexicon, resulting in greater facilitation in translation priming tasks (e.g., Basnight-Brown and Altarriba [2007], Duyck [2005]).

Order of language acquisition, language proficiency, and language dominance have been demonstrated to be important factors influencing outcomes in cross-language priming. In particular, priming effects are typically stronger in the L1-L2 direction than the L2-L1 direction, a phenomenon referred to as bilingual translation asymmetry (Keatley et al. [1994b], de Groot and Nas [1991], Kroll and Stewart [1994], Williams [1994], Gollan et al. [1997], Jiang [1999], Jiang and Forster [2001], Duñabeitia et al. [2010],

According to the Revised Hierarchical Model (e.g., Kroll and Stewart [1994]), translation asymmetry arises because the L1 has privileged access to conceptual meaning, whereas the L2 requires mapping via the L1 translation equivalent. In this model, translation asymmetry is assumed to persist until proficiency is sufficient to support direct access to meaning in the L2. More recently, however, models such as the Bilingual Interactive Activation (BIA, and its extension BIA+) (e.g., Dijkstra and van Heuven [1998], Dijkstra and van Heuven [2002]) argue that access across-languages is parallel and nonselective; bilingual translation asymmetry is instead due to differences in activation strength. Alternatively, asymmetries may also reflect inhibitory control processes, whereby the dominant L1 requires stronger suppression than the L2, and this residual inhibition interferes with subsequent processing (e.g., Meuter and Allport [1999]).

Empirical evidence strongly supports the view of parallel, nonselective access in multilingual processing (e.g., Marian and Spivey [2003], Spivey and Marian [1999], Weber and Cutler [2004]). Complementing this perspective, the Fuzzy Lexical Representations (FLR) hypothesis suggests that L1 representations themselves may be stronger and less “fuzzy” than those of the L2+, allowing them to dominate lexical competition more quickly (Gor et al. [2021]).

Word frequency has also been shown to influence cross-language priming, such that higher-frequency words produce stronger facilitation effects (e.g., Brysbaert et al. [2017], Brysbaert et al. [2016], Brysbaert et al. [2018],

Diependaele et al. [2013]). Word frequency effects may interact with language proficiency in cross-language priming, as demonstrated by a study (Chaouch-Orozco et al. [2021]) examining the interaction between word frequency, language proficiency and translation pairs, finding that facilitation from L1 translation matched primes for low-frequency L2 targets only among participants with high L2 proficiency.

It has also been suggested that stronger facilitation in priming occurs between lexical items from languages that are socially linked. Szakay et al. [2016] found that for L1 English L2 Māori listeners, Māori English (the variety of English associated with those of Māori descent) facilitated the priming of Māori lexical items more strongly than did Pākehā English (the variety of English associated with those of European descent). Moreover, a priming effect was also observed in the typically weaker L2–L1 direction, but only from Māori to Māori English, not from Māori to Pākehā English. These findings suggest that the indexical and linguistic overlap between Māori and Māori English strengthened the priming relationship, supporting facilitation even in the more sensitive L2–L1 direction. More broadly, this work aligns with the theory that socio-indexical information is accessed across languages, as is well established in monolingual processing (e.g., Babel and Munson [2013b], Abercrombie [1967], Munson et al. [2006], Fogerty [2015]). Further, this work suggested that the pathway between lexical items across-languages can be more strongly tied in priming between a strongly socially connected pair of languages than a weaker one.

In the current study, we hypothesize that short-term cross-language auditory priming will reveal a stronger link between Pidgin and ‘ōlelo Hawai‘i than between Hawai‘i English and ‘ōlelo Hawai‘i. Similar to the Māori context (where Māori is considered to have a stronger link to Māori English than to Pākehā English (e.g., Szakay et al. [2016])), Pidgin and ‘ōlelo Hawai‘i are described as having a stronger association than Hawai‘i English and ‘ōlelo Hawai‘i (e.g., Marlow and Giles [2010a]). We therefore predict that targets preceded by translation-equivalent primes in both languages will be processed faster than those preceded by unrelated control primes, with greater facilitation when the prime is in Pidgin. As in Experiment 1, this link may reflect multiple levels of connection, including linguistic, social, and contextual factors, which are described in detail in Chapter 3. Due to the context of the current study, although both Pidgin and English are L1 languages for many participants, dominance is likely to be context-dependent rather than uniform across domains. An individual may be more English-dominant in academic or institutional settings, while being more Pidgin-dominant in family or community contexts. Because the experimental task was conducted in

a laboratory setting, Hawai'i English may have functioned as the dominant language for many participants in this context, even if Pidgin is dominant in other domains. This context-specific dominance is therefore an important factor to consider when interpreting the present results.

In the study population included in Experiment 2, all participants identified English, Pidgin, or both as their L1, treated here as simultaneous first languages and coded as L1 for analytical purposes, while 'ōlelo Hawai'i was identified as an additional language and labeled L2 for analytical purposes. In the current study, translation priming was examined only in the L1–L2 direction, with all target items in 'ōlelo Hawai'i and primes in either Pidgin or Hawai'i English. We focused on the L1–L2 translation priming direction because it has been previously found to produce stronger effects. Examining this direction also allowed us to lend more focus to our specific research question (concerning the relation between Pidgin to 'ōlelo Hawai'i versus Hawai'i English to 'ōlelo Hawai'i, in this direction) and thus reduce the number of stimuli, minimizing participant fatigue.

Based on these hypotheses and the study design, we expect a cross-language priming effect for 'ōlelo Hawai'i targets with both Hawai'i English and Pidgin primes, with greater facilitation when the prime is in Pidgin. In other words, while targets preceded by translation-equivalent primes in both languages should be processed with faster reaction times than those preceded by unrelated control primes, the facilitation is predicted to be strongest for Pidgin primes.

4.2 Participants

Experiment 2 investigates trilingual speakers of 'ōlelo Hawai'i, Pidgin and Hawai'i English. As in Experiment 1, participant language background was determined via self-report. The participant pool was the same as in Experiment 1; however, the relevant information is repeated here for clarity. Inclusion in the analysis was based on pre-specified accuracy threshold of 80% or higher applied uniformly across all conditions as a data quality control measure. This criterion was defined prior to analysis and was independent of any statistical outcomes.

Participants were recruited through the linguistics department at the University of Hawai'i at Mānoa "Linguistics Beyond the Classroom" program, the Hawaiian Language Studies Department at the Universities of Hawai'i at Mānoa and Hilo campuses, through the use of physical flyers posted at the Universities of Hawai'i at Mānoa and Hilo campuses, by posting elec-

4.2. Participants

tronic flyers on Facebook ‘ōlelo Hawai‘i advocacy groups including “Hoa ‘ōlelo Hawai‘i”, and lastly via advertisements on radio and podcasts (e.g. Hawai‘i Public Radio, the Moana Nui Podcast). Participants were compensated in the amount of \$15 USD in one of three forms: 1) Zippy’s Gift Card (a local restaurant in Hawai‘i), 2) Amazon Gift Card, or 3) Donation to the ‘Aha Pūnana Leo immersion schools.

During the priming task of the current study, overall accuracy was 87%. This falls within the range reported for cross-linguistic priming and lexical decision tasks in the literature (see Chapter 3 for discussion), in which proficient bilingual speakers typically achieve high accuracy rates (approximately 85–95%), with substantially lower performance often interpreted as reflecting reduced task engagement or limited language proficiency (e.g., Szakay et al. [2016] Dijkstra et al. [1999], Brysbaert et al. [1999]). Earlier participants largely met expected thresholds (80%), whereas a growing proportion of later participants fell below them. Consequently, general recruitment was suspended, and subsequent recruitment focused exclusively on Hawaiian language consultants. Hawaiian language consultants were required to be either Hawaiian language teachers or professional Hawaiian language translators. In total, 7 Hawaiian language consultants participated in Experiment 2. Participants in this group were compensated at a rate of \$40 USD per hour in cash. The purpose of this targeted recruitment was to ensure a sample of highly proficient speakers, allowing us to assess whether the observed accuracy patterns persisted even among participants with language knowledge that was externally validated as highly advanced and sufficient for task performance. These participants were recruited by word of mouth and did not include the four Hawaiian language consultants that assisted in the creation of the stimuli in the current study.

A total of 37 participants completed the study. Of them, 7 did not meet the minimum qualifications (e.g., had not completed Hawaiian 102), and 4 appeared to provide nonsensical responses (e.g., entering random numbers or nonwords, or listing nonexistent cities, islands, or high school names) leaving 26 participants. Next, 9 additional participants were excluded based on the pre-specified accuracy threshold, leaving 17 participants for analysis. We acknowledge that this represents a substantial number of exclusions; further discussion of the potential implications is provided in the Discussion section.

Among these included 17 participants, substantial variation was observed across languages, indicating that trilinguals were imbalanced in their self-reported age of acquisition and proficiency across the three languages. These scores were obtained from the Language Background Questionnaire (see Appendix A). For age of acquisition, English was predominantly reported as

4.2. Participants

having been learned from birth (age 0) across participants, with relatively little variation. Pidgin was also most commonly reported as being acquired from birth; however, responses showed greater variability, with some participants indicating later ages of acquisition. In contrast, reported ages of acquisition for ‘ōlelo Hawai‘i were widely distributed, reflecting substantially more variation across participants (see Figure 4.1).

Proficiency ratings, measured on a scale from 1 to 10 (1 = lowest proficiency, 10 = highest proficiency), showed a similar pattern. English was consistently rated as highly proficient across all participants. In contrast, Pidgin and ‘ōlelo Hawai‘i exhibited greater variability. Self-reported proficiency for ‘ōlelo Hawai‘i ranged from as low as 1 to higher values, indicating heterogeneous experience with the language. Pidgin was generally rated as relatively high in proficiency, though typically lower than English, with scores spanning a broad range from 3 to 10 (see Figure 4.2).

Overall, these differences reflect substantial variability in participants’ linguistic experience across the three languages, with English consistently dominant in both age of acquisition and proficiency, Pidgin showing intermediate and more variable patterns, and ‘ōlelo Hawai‘i characterized by later acquisition and lower, more dispersed proficiency ratings. This pattern may give rise to differences in expected performance in language processing across the three languages, with an advantage for English relative to Pidgin, and the weakest performance expected for ‘ōlelo Hawai‘i.

4.2. Participants

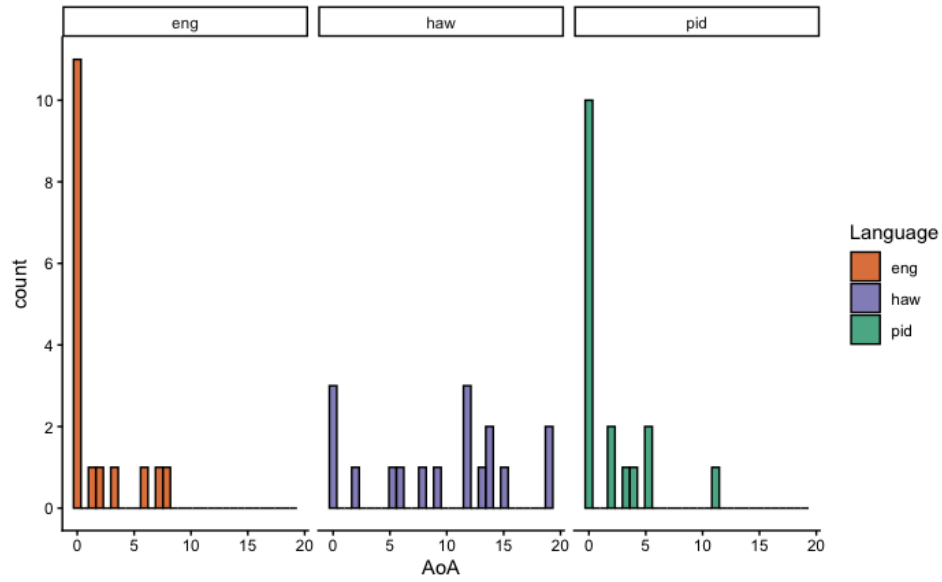


Figure 4.1: Distribution of age of acquisition (AoA) ratings for Pidgin, English, and ‘ōlelo Hawai‘i across participants. Values are plotted separately by language, allowing comparison of AoA distributions across the three language groups.

4.2. Participants

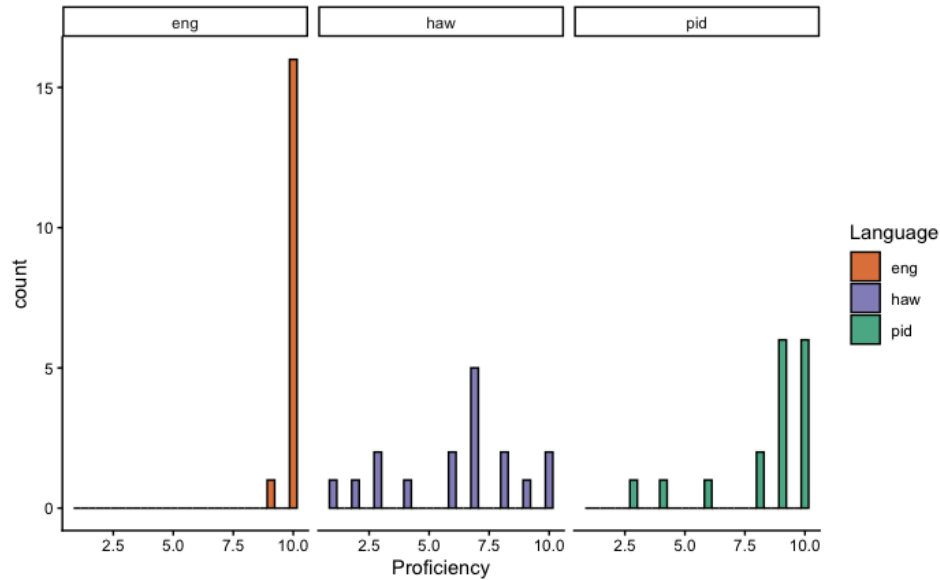


Figure 4.2: Distribution of self-reported proficiency ratings for Pidgin, English, and ‘ōlelo Hawai‘i across participants. Values are plotted separately by language, allowing comparison of proficiency distributions across the three language groups.

4.2.1 Inclusion and exclusion criteria

As with Experiment 1, all materials which advertised the study stated the following requirements: Participants must be 1. Minimally 18 years of age, 2. Speak ‘ōlelo Pa‘i‘ai (Pidgin) (and have at least 10 years of exposure), 3. Speak ‘ōlelo Hawai‘i (completion of Hawaiian 102 level+ or equivalent), and 4. Speak English. No proof of qualification was required by participants, but after many participants began signing up for the study without reading the qualifications, participants were all asked prior to the start of the study to confirm their qualifications, then describe their language background thoroughly in the language background questionnaire. The language background questionnaire frequently helped determine whether participants genuinely met the eligibility criteria and whether they approached the experiment attentively, as reflected in nonsensical results (e.g., random entries, reporting attendance at nonexistent schools in response to the question “Which high school in Hawai‘i did you attend?”, or listing nonexistent cities or islands in

response to “What city and island are you from?”).

4.3 Materials

4.3.1 Wordlist

Participants were presented a prime word in either Pidgin or Hawai‘i English, followed by a target word in ‘ōlelo Hawai‘i. Critical test pairs were cross-language translation equivalents: Pidgin-to-‘ōlelo Hawai‘i and Hawai‘i English-to-‘ōlelo Hawai‘i. Filler items included semantically unrelated real words and nonwords from all three languages. See Table 4.1 for a summary of this design.

Prime Language	Target Language	Prime Type	Target Type	Translation Relation	Trial Type
ENG	‘ōH	real	real	translation equivalent	critical
PID	‘ōH	real	real	translation equivalent	critical
ENG	‘ōH	real	real	non-translation (same target as critical)	control
PID	‘ōH	real	real	non-translation (same target as critical)	control
ENG	‘ōH	real	real	non-translation (unique pair)	filler
PID	‘ōH	real	real	non-translation (unique pair)	filler
ENG	‘ōH	real	nonword	–	filler
PID	‘ōH	real	nonword	–	filler
ENG	‘ōH	nonword	real	–	filler
PID	‘ōH	nonword	real	–	filler
ENG	‘ōH	nonword	nonword	–	filler
PID	‘ōH	nonword	nonword	–	filler

Table 4.1: Cross-lingual stimulus types used in Experiment 2. Within each list, ‘ōlelo Hawai‘i target items were evenly divided between translation-equivalent trials and non-translation control trials, such that each target appeared either with a translation-equivalent prime or with an unrelated control prime. Prime languages were counterbalanced across targets, with half of translation and control trials preceded by Hawai‘i English primes and half by Pidgin primes. Filler trials consisted of unique prime–target pairings.

Each list including critical, control, and filler pairs, contained 320 pairs in total: (i) 24 related real word pairs (critical), (ii) 24 unrelated real word pairs (control), (iii) 32 unrelated real word pairs (filler), (iv) 80 pairs with a nonword prime and real-word target, (v) 80 with a real-word prime and nonword target, and (vi) 80 with nonword primes and nonword targets.

Example pairs for each type are provided as follows: (i) 24 related real-word pairs (critical), e.g., squid–mūhe‘e (ENG–‘ōH) or ika–mūhe‘e (PID–‘ōH); (ii) 24 unrelated real-word pairs (control), e.g., nutmeg–mūhe‘e (ENG–‘ōH) or alalas–mūhe‘e (PID–‘ōH); (iii) 32 unrelated real-word pairs (filler), e.g.,

fruit–ānuenue (ENG–‘ŌH) or popo–lani (PID–‘ŌH); (iv) 80 pairs with a nonword prime and real-word target, e.g., milak–niho (ENG–‘ŌH) or bito–niu (PID–‘ŌH); (v) 80 pairs with a real-word prime and nonword target, e.g., rash–lauhi (ENG–‘ŌH) or bento–okola (PID–‘ŌH); and (vi) 80 pairs with nonword primes and nonword targets, e.g., zothy–hemāe (ENG–‘ŌH) or tanah–kawo (PID–‘ŌH).

The 24 related real word pairs consisted of translation equivalents that were matched as closely as possible in meaning. The control condition included 24 unrelated real word pairs, ensuring that all target items were also presented to some participants in a condition where the Pidgin or Hawai‘i English prime was a real word but was not semantically related to the target. Filler trials manipulated the presence of words and nonwords in the prime and or target position to reduce repetition effects and to prevent participants from developing expectations about wordhood or semantic relatedness. To maximize the likelihood of observing cross-language priming effects, only a small proportion of the total stimulus set consisted of translation-equivalent prime–target pairs. This design choice is supported by prior research showing that priming effects are strongest when translation matches are infrequent and embedded among mostly unrelated or nonword pairings (e.g., de Groot and Nas [1991], Keatley et al. [1994a]). By minimizing the overall proportion of translation-equivalent matched pairs, the design reduces participants’ expectations of semantic relatedness, thereby encouraging automatic rather than strategic processing. For full item lists 1-4, see the Appendix under subsection Experiment 2 Stimuli.

4.3. Materials

Table 4.2: Cross-language translation equivalents

Hawai'i English Prime	Pidgin Prime	‘ōlelo Hawai'i Target	Hawai'i English Prime	Pidgin Prime	‘ōlelo Hawai'i Target
undies	bibadeez	palema'i	later	bumbai	‘auane'i
big	bambucha	nui	woohoo	cheehoo	hūtā
toad	bufo	poloka	awesome	charreh	makepono
soilderfish	menpachi	‘ū‘ū	pouty	habuteru	nuha
idiot	babooze	hūpō	urine	shishi	mimi
pompous	highmakamaka	ho'oki'eki'e	chaotic	hammajang	kapakahi
breasts	chichis	waiū	karma	bachii	pohō
badass	hammah	loea	clique	gangies	hoaloha
shame	halā	hilahila	sore	itai	‘eha
pig	buta	pua'a	trivial	minahs	‘u'uku
unkempt	hash	‘eke	wassup	howzit	pehea
food	kaukau	‘ai	whatchamacalit	dakine	mea
little	schmolkine	iki	worn	bassap	palahē
rags	boroboros	weluwelu	cockroach	b52	‘elelū
sarong	lavalava	kikepa	squid	ika	mūhe'e
mucus	hanabata	hūpē	sloppy	nailz	kāpulu
baldy	bolo	‘ōhule	putrid	kusai	pilo
stupid	bakatare	hepa	ghost	obake	lapu
seasnail	tsubu	pipipi	goodness	aisos	auwē
bath	bocha	‘au'au	disgusting	uji	ho'opailua
phlegm	galas	male	liar	bulai	wahahe'e
sleep	nainai	hiamoe	hot	atsui	wela
toilet	benjo	lua	uptempo	chalangalang	mele
cocky	tantaran	hō'oiio	surely	garans	hiki

To isolate language-based priming effects from item-specific ones, four counterbalanced stimulus lists were created. Key differences were in the critical and control pairings; all lists contained the same filler pairs. Participants were randomly assigned to one of the four lists. Each list included all 48 target ‘ōlelo Hawai'i words, but the prime preceding each target differed by list. For all 48 target ‘ōlelo Hawai'i words, both a Pidgin–‘ōlelo Hawai'i and a Hawai'i English–‘ōlelo Hawai'i pairing existed, such that across lists, each ‘ōlelo Hawai'i target was preceded by either a Pidgin or Hawai'i English prime. This counterbalancing ensured that any observed differences in accuracy or reaction time were not attributable to the inherent familiarity or word frequency of specific ‘ōlelo Hawai'i words, but rather whether the prime matched in meaning, and if the language of the prime made a difference in priming response. In every list, half of the targets were paired with a translation-equivalent prime (half Pidgin, half Hawai'i English); the other half were paired with unrelated control primes (again, split evenly between Pidgin and Hawai'i English). This ensured each target word appeared in all conditions across participants, but only once per list. All 48 translation equivalents included across experimental lists are included in Table 4.2.

The task in total had two bilingual test conditions: PID-‘ōH and ENG-‘ōH (e.g., ika-mūhe'e, squid-mūhe'e, respectively). As our research question is interested in the ability of Pidgin versus English to prime ‘ōlelo Hawai'i lexical items, we did not test prime pairs such as ‘ōlelo Hawai'i-to-Hawai'i English, or

‘ōlelo Hawai‘i-to-Pidgin. To minimize participant fatigue, pairings that were not directly relevant to the research question were excluded from the study. This design also allowed us to first assess whether effects differed between Pidgin and Hawai‘i English in priming ‘ōlelo Hawai‘i while maintaining a more manageable stimulus set.

4.3.2 Creation of nonwords and selection of real words

The nonwords for Pidgin, Hawai‘i English, and ‘ōlelo Hawai‘i were generated using a concatenation tool called Uni-Pseudo (New et al. [2024]). This tool implements a bigram algorithm that recombines segments drawn from real words in a given list to create novel letter strings. As such, the generated pseudo-words approximate the surface patterns of the source languages without replicating existing lexical items. To ensure that our list of nonwords did not inadvertently include real words, we cross-checked the output against attested vocabulary contained in dictionaries for each language (Pidgin: Pidgin to Da Max (Simonson et al. [2005]), Pidgin to da Max Hana Hou (Simonson et al. [1992]); ‘ōlelo Hawai‘i: The Pocket Hawaiian Dictionary (Pukui et al. [1975], The Ulukau (Kuamo‘o [2009]); English: Merriam Webster (Merriam-Webster, Inc. [2025])). All items (real and nonword) were also validated for meaning (as translation equivalents or not, where appropriate) and assessed for familiarity by the researcher and four language consultants. If a real word was unfamiliar to any of the consultants, it was replaced with another word of comparable frequency (if information was available) and syllable count. If a nonword was recognized as a real word, it was replaced with a different nonword.

All target items were matched for syllable count and word frequency (log frequency) across experimental blocks. Prime items were matched for syllable count and word frequency if the information was available across experimental blocks within each language. Because the three languages have different syllable structure and word shape, and frequencies were sourced from different corpora, there was not an attempt to balance differences across language in word frequency, word length, or syllable count; instead, these factors were balanced only within each language across blocks.

Word frequency information for ‘ōlelo Hawai‘i and Hawai‘i English were sourced from two corpora. The corpus used for ‘ōlelo Hawai‘i was based on the Ulukau Electronic Library (Kuamo‘o [2009]), frequency information analyzed by (Doherty [2024]). The corpus used for English was the English Lexicon Project (Balota et al. [2007b]), chosen for its accessibility and comprehensive lexical information, which met the needs of our experimental

design, though other corpora could also have been used.

The Ulukau Electronic Library ⁴ is a publicly accessible collection of texts and oral recordings in ‘ōlelo Hawai‘i, containing both written text and annotated auditory and video in ‘ōlelo Hawai‘i. The collection includes a vast range of sources, including fiction and non-fiction, books, dictionaries, newspapers, historical documents, genealogies, and radio. Most notably, the Ulukau includes the kani‘āina (Kimura [1972]), a digital corpus of ‘ōlelo Hawai‘i speech including transcripts from Ka Leo Hawai‘i, a public radio program, and Mānaleo, a public television program. While the corpus itself does not include frequency values, all printed text was converted to into machine-encoded text using Optical Character Recognition (OCR) (Doherty [2024]).

To generate frequency values, we processed the corpus in Python, listing all unique words and computing their word counts and log frequencies. Extensive filtering processing with regular expressions (Regex) was required to address three major issues from the ‘ōlelo Hawai‘i corpus: (1) the presence of non-‘ōlelo Hawai‘i words (e.g., English, Pidgin, Sāmoan, Māori), likely due to code-switching in speech; (2) OCR errors (e.g., ‘okina (‘) was misread as an apostrophe or the letter l, and the kahakō (̄) misread as a hyphen); and (3) variation in orthographic conventions, including inconsistent use or omission of kahakō and ‘okina. The steps that were taken were as follows: 1) Taking advantage of the relatively small ‘ōlelo Hawai‘i small sound inventory, with standard spelling conventions including only 19 characters (m,n,p,t,k,h,w,l,‘,a,e,i,o,u,ā,ē,ī,ō,ū), all words that contained any other characters were eliminated. 2) ‘ōlelo Hawai‘i also has a limited syllable inventory (C)V(V)(V)(V), thus any words with illicit sound combinations as per syllable structure rules were eliminated (e.g., there should be no consonant clusters in the word; a word should never end with a consonant; the maximum size of a vowel cluster can be 4; etc.). 3) All (-) dash characters that were located rightward to a vowel were assumed to be a long vowel. 4) Finally, it was observed that many words in the corpus exceeding 10 characters were improperly concatenated (e.g., the item nohea‘okeoki contained in the corpus consists of four separate words without spaces separating them). As a result, all words longer than 10 characters were removed from the corpus. The resulting ‘ōlelo Hawai‘i corpus contained approximately 4.4 million word tokens and 50,000 unique word types.

Steps 1-4 addressed the majority of errors in the corpus, but these filtering methods had some limitations. In some cases, filtering was over-applied,

⁴Current URL: <http://www.ulukau.org/>

resulting in the exclusion of true ‘ōlelo Hawai‘i words that contained unusual sounds (e.g., segments such as /s/ or /b/ may occur in ‘ōlelo Hawai‘i phonetically but are not phonemic) or exceeded 10 characters in length. In other cases, filtering was under-applied, resulting in the inclusion of unattested ‘ōlelo Hawai‘i words. For example, words that satisfied the syllable structure and sound inventory requirements were included even though they were not real words. Improper inclusions and exclusions were identified through careful review by language consultants. For the top 3,000 most frequent words, a Hawaiian language consultant manually checked each item, and any errors were corrected. Some errors still may have remained after these processes, however substantial effort was made to ensure that the majority of high-frequency words in the final corpus were accurate. The resulting ‘ōlelo Hawai‘i corpus included the fully reviewed top 3,000 words.

Words that were excluded (either manually or computationally) due to being misspelled were manually reviewed to determine the intended correct form, and their corresponding frequencies were added to the frequency of the correct word. For example, the word *Hawaii*, a misspelling of *Hawai‘i*, had a frequency count of 13483, while the correctly spelled *Hawai‘i* had a separate frequency of 16625. The accurate frequency for the word *Hawai‘i* was thus determined as the sum of these two counts, giving a total of 30108.

Other steps which were taken but were ultimately reversed included elimination of words that were contained in the English dictionary. An extensive amount of true ‘ōlelo Hawai‘i homonyms with English words existed (e.g. *make* (Hawaiian: ‘to die’).

English Lexicon Project includes approximately 40000 unique words. This project includes manually annotated speech data from 816 participants across six universities across the United States. While speakers of Hawai‘i English were not included in this corpus, we believe the frequency information in this corpus is an acceptable proxy of Hawai‘i English. We note that other corpora such as SUBTLEX could also have been used, but ELP provided frequency information that was sufficient and consistent with prior work using similar experimental paradigms (e.g. Kinoshita et al. [2014]).

No corpus of conversational Pidgin with word frequency information currently exists. Consequently, frequency counts for Pidgin items could not be established or used for matching. Instead, syllable count was used as a matching criterion across experimental blocks within each language.

Throughout the word list, cross-linguistic homophones were excluded (e.g. *tako* (Pidgin: ‘octopus’), *taco* in English). Thus, English or Pidgin cognates or homophones shared with ‘ōlelo Hawai‘i (e.g. *titah* (Pidgin, ‘tough woman’) *tita* (Hawaiian: ‘woman, sister’), *ukulele* (Hawaiian: ukulele

instrument lit ‘jumping flea’) and *ukulele* (English, borrowed from Hawaiian) were also excluded. Excluding primes of ‘ōlelo Hawai‘i origin, particularly in the case of Pidgin, also allowed us to exclude word similarity as a potential factor leading to the facilitation of bilingual processing, and focus on the social or contextual relatedness of Pidgin and ‘ōlelo Hawai‘i or the linguistic effects below the word level (e.g., phonological, phonetic, subphonemic). For nonwords, language consultants validated that no nonwords were homonyms of real words or contained a sociolinguistic variation that would potentially allow a participant from Hawai‘i to misperceive the word as real. All words judged as too similar to a real word when spoken aloud were excluded (e.g., nonword, English: ‘bax’ and real word, English: ‘backs’).

After word lists were finalized, we conducted analyses of variance (ANOVAs) on critical and filler items to confirm that they were equal across word frequency and syllable count within-language, across blocks. Specifically, we examined whether ‘ōlelo Hawai‘i targets and Hawai‘i English primes differed in frequency or syllable count across blocks within each language, and whether Pidgin primes differed in syllable count, since frequency information was not available. Critical items were also compared with fillers within each language to confirm that the two sets did not differ statistically.

For ‘ōlelo Hawai‘i critical targets, there were neither differences across blocks in neither word frequency [$F(3, 92) = 0.022, p = 0.996$] nor syllable count, [$F(3, 92) = 0.072, p = 0.975$]. The same was true for Hawai‘i English critical primes, where no block-level differences were found in frequency [$F(3, 36) = 0.328, p = 0.805$], nor syllable count [$F(3, 36) = 0.924, p = 0.439$]. For Pidgin critical primes, no differences were observed across blocks for syllable count [$F(3, 44) = 1.469, p = 0.236$]. Differences in numerator degrees of freedom reflect the number of unique word frequencies or syllable counts present in each language and factor, while differences in denominator degrees of freedom reflect the total number of items analyzed, including any items removed during preprocessing to balance distributions across conditions.

Next, critical and filler items were compared within each language. For Hawaiian, there were no differences between critical word targets and filler words in terms of frequency [$F(1, 542) = 0.044, p = 0.835$], or syllable count [$F(1, 542) = 0.3088, p = 0.573$]. Similarly, Hawai‘i English critical word primes and fillers did not differ in frequency [$F(1, 262) = 1.319, p = 0.524$] or syllable count [$F(1, 262) = 0.319, p = 0.573$]. Finally, Pidgin critical primes and filler words were also comparable, with no difference in syllable count [$F(1, 590) = 0.853, p = 0.356$]. The variation in denominator degrees of freedom across these comparisons reflects differences in the number of items analyzed per language and factor. Together, these analyses confirm

that the materials were well balanced across blocks, and that within each language, critical items were comparable to filler items on the dimensions available for analysis.

4.3.3 Creation of auditory stimuli

All auditory stimuli were created from recordings of three speakers, each representing one of the language varieties in the study. One female Pidgin speaker and one female Hawai'i English speaker were recorded for the prime words, and one male 'ōlelo Hawai'i speaker was recorded for the target words. All three speakers were from O'ahu, Hawai'i, and were in their mid to late 30s at the time of the recording. Speakers were selected based on their interest, their qualifications as language speakers, and their prior experience with narration work.

To avoid priming due to mere voice similarity, prime words were consistently produced by female voices and target words by a male voice. Because Hawai'i English and Pidgin lie on a spectrum of likeness⁵, we further ensured distinctness by recruiting a Hawai'i English speaker identified with what is locally referred to as "Punahou English", a formal register of Hawai'i English associated with elite schooling and considered culturally opposed to Pidgin. A common saying in Hawai'i references this association, "My Pidgin went public sku, but my customer service voice went to Punahou".

All speakers were presented with the word list with the researcher on a laptop screen and read the words out loud at the pace of the speaker. The researcher would instruct the speaker over Zoom, and would be put on mute during the stimuli recording session and the speaker would record using their own computer and upload their sounds to a shared folder. The sound files were recorded directly onto a laptop computer with built-in microphone using Audacity 3.1 Team [2021] in mono, at a sampling frequency of 44100 Hz in .wav format.

At the start of each session, the speaker would practice recording to ensure equipment and recording software were functioning appropriately. The speaker would send the test recording to the researcher and the sound was assessed for quality or ambient noise. Changes were made to meet any needs determined from the test recording session (e.g. speaker sitting too close to the microphone, speaker had a window open, etc.). Next, the speaker was

⁵Often described as a creole continuum, referring to the range of speech varieties that exist between a basilectal creole and an acrolectal standard, with intermediate varieties showing gradient phonological, lexical, and grammatical features rather than forming discrete categories (DeCamp [1971]).

asked to look at the list of words and nonwords and recite them aloud. Some coaching was involved, if the speaker had trouble reciting particularly the nonwords or if there were alternative ways to pronounce the word. In such cases, the researcher would recite the word and the speaker would repeat it. When the speaker was ready, the researcher muted their microphone and the speaker recorded themselves, with the researcher watching. Recording paused if there was loud background noise or if the speaker needed a break. The word lists were broken into parts, where the speakers first read the critical items, followed by the fillers then finally the nonwords. All words were repeated three times. Each speaker was instructed to speak in a natural way, as if they were speaking in conversation. Each recording session took approximately one hour to complete.

4.3.4 Language Background Questionnaire

The same questionnaire was included in Experiments 1, 2, and 3. See section 3.3.2 for more information about the Language Background Questionnaire (see Appendix A). Any participants that took part in more than one experiment only took the Language Background Questionnaire one time. Returning participants were given access to a shortened version of the experiment which did not include the Language Background Questionnaire.

4.4 Experimental Procedure

Experiment 2 used the Gorilla experimental builder (Anwyl-Irvine et al. [2020]), and was accessed by participants online through the use of a personalized link.

Participants first read a consent form, and only proceeded to the next part of the study if they indicated that they accepted the terms of the study.

Next, participants continued to a headphone check (Woods et al. [2017]). This task was designed to detect whether stereo headphones were being used rather than speakers and to exclude participants unable to comply with instructions from continuing further. Participants were instructed to sit somewhere quiet, turn off distracting media, and close unrelated tabs. The headphone check consisted of six trials in which participants were presented three tones per trial of various degrees of loudness and were asked to identify the quietest tone (1st, 2nd, or 3rd). Participants who failed the check were able to retake the test twice more before finally being automatically excluded from the study and advised to troubleshoot and contact the researcher for

4.4. Experimental Procedure

further guidance. If the headphone check was passed, participants continue to the experimental procedure.

Participants were randomly assigned to one of four experimental lists. Next, participants viewed instructions. Participants were instructed that each trial would consist of a two-word auditory sequence. The first item was either a real word or nonword in English or Pidgin, presented in a female voice. Nonwords were defined to reduce ambiguity in the task (e.g. instructions read, “The first word will be either a word or fake *nonword* in English, *or* a word or nonword in Pidgin”). The second item was either a real word or nonword in ‘ōlelo Hawai‘i, spoken by a male voice. Participants were told to listen to both words but focus only on the second word. Their task was to determine whether the second item (the ‘ōlelo Hawai‘i word) was a real word or a nonword. Instructions emphasized that this judgment was the sole task required on each trial. (e.g. “Your task is to determine if the SECOND ITEM ONLY is a real word OR a nonword in ‘ōlelo Hawai‘i. This is your only task.”)

All stimuli were presented auditorily, with no words shown in written form. This decision was made because Pidgin is primarily an oral language without standardized spelling conventions, and because Hawaiian orthography, while officially standardized, carries a history of contested conventions and variation. Presenting words auditorily prevented participant perception from being biased by orthographic differences across languages, or based on one’s affiliation with different orthographic norms.

Participants were instructed that responses were to be made via keyboard: pressing 1 indicated the participant believed the word was real, and pressing 0 indicated they believed it was a nonword. Participants were instructed to respond as quickly and accurately as possible. Each trial was time-limited; if no response was recorded within 10 seconds, the experiment automatically advanced to the next trial. Breaks were offered every 50 trials to reduce participant fatigue.

At the end of these instructions, participants were told to press next to continue onto two practice trials to familiarize themselves with the procedure. The first practice trial played a prime item *chone* [tʃon.e], a nonword in Pidgin, then played a target item *wahine* [va.hi.ne] ‘woman’, a real word in ‘ōlelo Hawai‘i. After a response was entered the participant could continue to the second practice trial which played a prime item *brah* [brɪə] ‘dude’, a real word in Pidgin, then played a target item *pahine* [pa.hi.ne], example of a nonword in ‘ōlelo Hawai‘i). All practice items were recorded by the experimenter herself. To mask the identity of the experimenter, the pitch of the recordings for practice items was manipulated using Pitch-Synchronous

4.4. Experimental Procedure

Overlap and Add (PSOLA) in Praat, a method that splits the waveform into short pitch-synchronous segments and recombines them to change pitch while preserving natural speech quality. Practice prime items were raised in pitch, while practice target items were lowered to fall within the typical range of a male voice.

After the practice trial, a screen displayed the message, “Maika‘i! “wahine” was an example of a real word, and “pahine” was an example of a nonword. You may now proceed to the experiment.” The instructions were reiterated to make it very clear to participants that the task was to decide whether the second word only was a real or nonword in ‘ōlelo Hawai‘i, and to press 1 on the keyboard to provide the response the item was a real word and to press 2 on the keyboard to provide the response the item was a nonword. Participants were instructed to press next to continue to the experiment.

Participants next completed the full list of items to which they were assigned, presented in a randomized order. At the start of each trial a fixation + symbol would appear for 250 ms. The prime item automatically followed. Then an interstimulus interval (ISI) screen displayed for 500 ms, and automatically advanced to the target item. Immediately following the target item a decision screen displayed “Word Press 1, nonword Press 0”. Participants could respond only while this screen was displayed, which remained on for 10 seconds or until a response of 1 or 0 was recorded. If participants did not respond within 5 seconds, a countdown timer appeared in the top-right corner of the screen, displaying the remaining time from 5 to 0 seconds until the trial was automatically terminated and the screen would automatically advance to the next trial. Trials in which participants failed to respond were coded as no response. Advancing from the decision screen would lead to a 1000 ms intertrial interval (ITI) before advancing to the start of the next trial. The fixation symbol indicated to participants the onset of a new trial. The ISI was included between the prime and target stimuli to allow participants sufficient time to process the first item before the onset of the second, and to ensure adequate time for semantic priming to occur. This brief pause reduces auditory overlap and prevents participants from skipping the prime (Buonomano et al. [2009]). Crucially, the magnitude and nature of semantic priming effects depend on the temporal interval between stimuli, with shorter stimulus onset asynchronies (SOAs) primarily reflecting automatic spreading activation and longer SOAs allowing the engagement of controlled, expectancy-based processes (Neely [1977]).

The ITI was used to provide consistent pauses between trials, giving participants time to respond and prepare for the next trial. This interval helps to reduce fatigue, maintain engagement, and avoid rapid carryover

effects that could confound response times or decision accuracy (Compton et al. [2017]).

As stated in the instructions, every 50 trials participants were allowed to take a break at a screen that displayed that they were able to take a break and press a “next” button to resume the experiment. Participants could take a break for as long as they desired. During the course of the study, the top edge of the screen displayed a progress bar that progressively filled as fewer trials remained toward the completion of the stimuli list. When participants reached the end of the study, a screen displayed a message indicating that they had reached the end of the experimental portion of the study. If the participant did not complete any other experiment, they were required to complete the Language Background Questionnaire and Language Attitudes Survey, the same questions used in Experiment 1 and Experiment 3 (see Appendix A). If they had completed another experiment prior to this study, they would skip the questionnaire and survey. Following this step, all participants would proceed to a screen indicating that the study has concluded in its entirety.

4.5 Analysis and Results

Among 17 included participants performing at 80% or above, Table 4.3 presents overall accuracy by trial type and Table 4.4 presents overall accuracy by prime language (Hawai‘i English or Pidgin) and target type (nonword or word).

As shown in Table 4.3, overall accuracy was high across trial types, with differences observed depending primarily on whether the target item was a word or a nonword. Accuracy for cross-language translation (word–word) equivalents was relatively high at 0.91, alongside other conditions in which the target item was a word. These included cross-language control (word–word) trials, which showed similarly high accuracy at 0.90, nonword–word trials with an accuracy of 0.91, and word–word filler trials, which yielded the highest accuracy overall at 0.94. In contrast, conditions in which the target item was a nonword showed lower accuracy. Nonword–nonword filler trials had an accuracy of 0.84, while word–nonword filler trials showed the lowest accuracy at 0.80. Overall, rejecting nonword items was more difficult than accepting words as real. Accuracy did not appear to be influenced by whether pairs were matched in terms of word–nonword identity or meaning.

4.5. Analysis and Results

Trial type	Mean Accuracy
cross-language translation (word-word) equivalents	0.91
cross-language control (word-word) trials	0.90
nonword-nonword filler trials	0.84
nonword-word filler trials	0.91
word-nonword filler trials	0.80
word-word filler trials	0.94

Table 4.3: Overall accuracy across trial type after exclusion of participants performing under 80%.

Observing accuracy by prime language and target type, as shown in Table 4.4, accuracy again appears to be influenced by target type (word vs. nonword). Accuracy for word targets was above 0.90 across both prime languages, whereas accuracy for nonword targets ranged between 0.80 and 0.90. Across prime languages, small differences were observed. Hawai'i English primes resulted in slightly higher overall accuracy compared to Pidgin primes, particularly for word targets (0.92 vs. 0.91). For nonword targets, performance was lower in both language conditions, with accuracy of 0.83 following Hawai'i English primes and 0.81 following Pidgin primes.

Prime Language	Target Type	Mean Accuracy
english	nonword	0.83
english	word	0.92
pidgin	nonword	0.81
pidgin	word	0.91

Table 4.4: Overall accuracy across prime language and target type after exclusion of participants performing under 80%.

Among those scoring 80% or higher, outlier participants were then identified and removed by inspecting individual RT distributions. This revealed that one participant (ID: 11741447; see Figure 4.3) responded substantially more slowly than the rest and was therefore excluded from further analysis (overall mean RT = 661.24 ms, SD = 475.38 ms; participant 11741447 mean RT = 2948.84 ms, SD = 1013.32 ms). After this, at the trial level, all responses with RTs greater than 3000 ms were removed. In total, the dataset initially contained 740 trials. After removing 54 trials with RTs greater than

4.5. Analysis and Results

3000 ms, 686 trials remained. Importantly, these RT-based exclusions only affected the reaction time analyses; no trial accuracy values were modified. Trials exceeding the RT cutoff were simply removed from the RT dataset, without any recoding of correct or incorrect responses. The resulting dataset was used for all subsequent RT analyses.

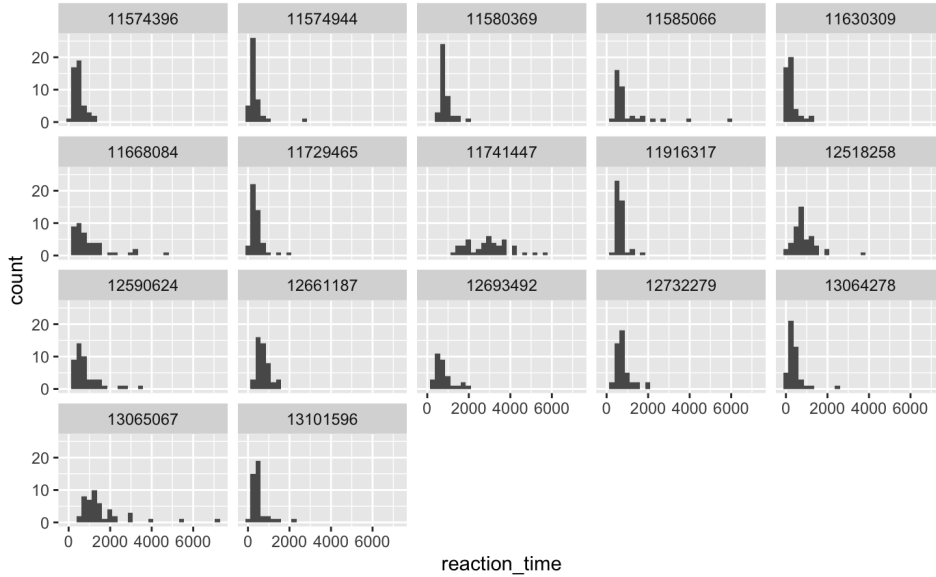


Figure 4.3: Histograms with RT distributions for participants with greater than 80% accuracy.

Next, we examined the cross-language priming effect as a function of prime language and trial type with correct responses only. For this analysis, a priming effect was defined as faster RTs to targets preceded by translation-equivalent primes compared to targets preceded by unrelated control primes. If the priming effect is stronger for Pidgin than for Hawai'i English, the difference in RTs between translation-equivalent and unrelated control primes will be larger for Pidgin than for Hawai'i English, crucially in the direction where RTs for translation-equivalent primes are faster than for unrelated control primes.

We begin our examination of the cross-language priming effect in Figure 4.4, viewing reaction times (RTs in ms) for cross-language translation equivalent trials compared to control trials by prime language after all removals. As depicted, it appears there is no clear priming effect for Pidgin trials, with

4.5. Analysis and Results

distributions appearing similar for prime–target pairs matched in meaning and for control pairs. In contrast, a possible priming effect is observed for Hawai‘i English trials, where reaction times are slightly faster for translation equivalents than for controls.

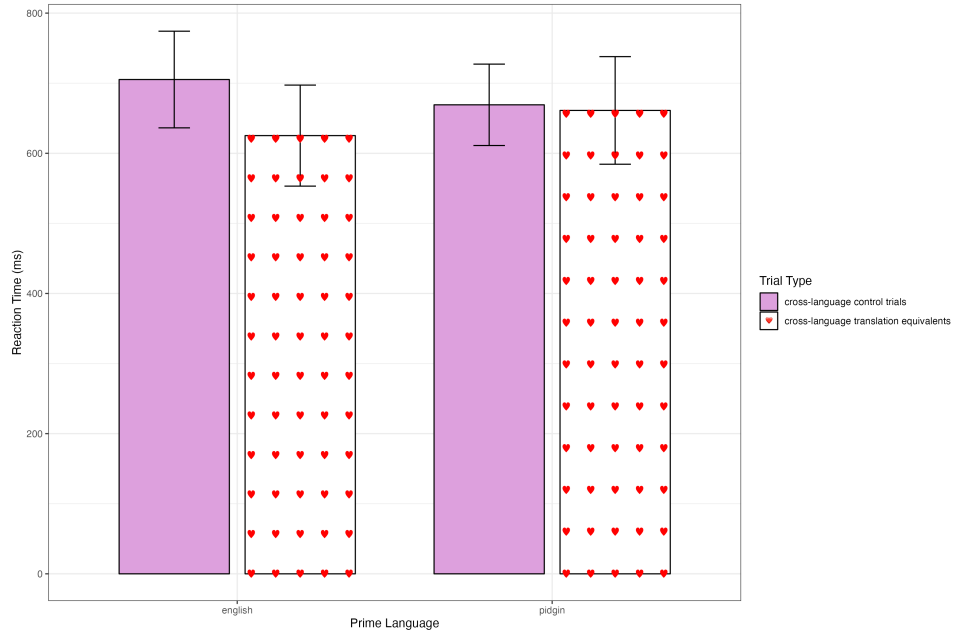


Figure 4.4: RT differences by trial type and prime language for correct responses only.

The validity of these apparent effects in Figure 4.4 was evaluated using a Bayesian model mixed-effects model using the brms package (Bürkner [2017]). In this model, reaction times (RTs, in ms) served as the dependent variable. Fixed effects of prime language (Pidgin vs. Hawai‘i English), trial type (translation equivalent vs. control), and their interaction. In this analysis, Hawai‘i English served as the reference level for prime language, and control trials served as the reference level for trial type. Coefficients represent the differences from these baseline conditions: how Pidgin primes differ from Hawai‘i English primes, how translation equivalent trials differ from control trials, and how the effect of trial type varies depending on prime language. The model included random intercepts for target items and by-participant random slopes for prime language, trial type, and their interaction. Because

RT data are strictly positive and typically right-skewed, we specified a log-normal likelihood (family = lognormal()), which models the log-transformed RTs as approximately normal. This ensures that estimates respect the lower bound of zero and appropriately capture the skewed distribution of RTs. The following formula was used in this model:

$$\begin{aligned} \text{reaction_time} \sim & \text{prime_language} \times \text{trial_type} \\ & + (1 + \text{prime_language} \times \text{trial_type} \mid \text{participant_id}) \\ & + (1 + \text{prime_language} \mid \text{target}) \end{aligned}$$

Priors were specified as weakly informative. This is common in Bayesian modeling (following e.g., McElreath [2020]) because it helps with model convergence and prevents extreme parameter estimates without strongly biasing results. The intercept was assigned a normal distribution with $\mu = 6$, and $\sigma = 1$. The use of this prior sets the expectation of the log(RT) to be normally distributed about the mean of 6 log ms, with a standard deviation of 1 log ms. A mean of 6 on the log-transformed scale corresponds to an expected reaction time of approximately 403 ms, as $e^6 \approx 403$, reflecting studies have reported mean RTs in lexical decision tasks ranging from 400 ms to 600 ms (e.g., Ratcliff et al. [2004]). A standard deviation of 1 on the log-transformed scale representing moderate uncertainty around the mean ($e^1 \approx 2.72$).

Priors for fixed effects (class = b) were assigned a normal distribution with $\mu = 0$, and $\sigma = 0.25$, again reflecting the default assumption of no effect for the predictor while allowing moderate variability. Random effect standard deviations and residual error (class = sd, sigma, respectively) were given a cauchy distribution with $\mu = 0$, and $\sigma = 0.25$, to constrain implausibly large variability while also remaining relatively uninformative. Model fitting was performed using four chains, each with 6000 iterations (1000 warm-up, thinning factor of 2), and a fixed seed for reproducibility.

The model's population-level effects (fixed effects), as seen in Table 4.5, indicated an intercept of 6.28 log ms (Estimate = 6.28, 95% CI: [5.98, 6.57]), representing the predicted log-transformed reaction time for trials in which the prime language was Hawai'i English and the target was preceded by an unrelated control prime. In real RT (back-transforming the log-transformed intercept to milliseconds using the exponential function), corresponding to a mean baseline reaction time of $e^{6.28} \approx 536$ ms, and a range between $e^{5.98}$ - $e^{6.57} \approx 397$ -714 ms across participants.

Relative to Hawai'i English primes, the effect of Pidgin primes was small and uncertain (Estimate = -0.04, 95% CI [-0.17, 0.10]), providing no evidence

4.5. Analysis and Results

of a systematic difference in reaction times between Pidgin and Hawai'i English primes.

The effect of cross-language translation equivalent trials was also small (near zero) and its confidence interval crossed zero (Estimate: -0.07 , 95% CI: $[-0.20, 0.07]$), providing no credible evidence for difference in RTs between trials where primes matched targets in meaning and trials where they did not. This suggests there was no priming effect in the baseline condition, when the prime language was Hawai'i English. Thus, the potential priming effect suggested in Figure 4.4 was not found to be credible.

Finally, the interaction between prime language and trial type was also small (near zero) and crossed zero (Estimate = 0.07 , 95% CI: $[-0.11, 0.25]$), indicating no credible evidence of an interaction. This means that, when targets were preceded by Pidgin primes, the priming effect (i.e., the difference between translation-equivalent and control trials) was approximately 0.07 log ms slower than when preceded by Hawai'i English primes, a difference that is not meaningfully large and lies within a wide credible interval and is thus inconclusive. When relating this to Figure 4.4, this indicates that the apparent visual difference in priming effects between Hawai'i English and Pidgin primes is not statistically meaningful and should not be interpreted as evidence of a reliable effect.

Predictor	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept	6.28	0.15	5.98, 6.57	1.00	3983	5885
Prime Language (Pidgin vs. Hawai'i English)	-0.04	0.07	-0.17, 0.10	1.00	9454	9238
Trial Type (translation equivalent vs. control)	-0.07	0.07	-0.20, 0.07	1.00	9286	9029
Prime Language $\times$ Trial Type	0.07	0.09	-0.11, 0.25	1.00	9657	8628

Table 4.5: Population-level effects from the Bayesian mixed-effects model estimating log reaction times (log RTs), including fixed effects for Prime Language (Pidgin vs. Hawai'i English), Trial Type (translation equivalent vs. control), and their interaction.

Overall, the credible intervals for the effect of prime language, and even for trial type (i.e., translation-equivalent versus control), each included zero. Consequently, the results provided no evidence that Pidgin primes facilitated processing of 'ōlelo Hawai'i targets more effectively than Hawai'i English primes. Furthermore, there was no indication that either language exerted a measurable priming effect on 'ōlelo Hawai'i within the parameters of the present study.

This study found no evidence of priming for either prime language, in contrast to prior research reporting cross-language facilitation effects (e.g., Szakay et al. [2016]). In the discussion that follows, we examine potential

sources of complexity that may have contributed to the lack of a priming effect, situating the findings in the broader literature, highlighting their theoretical significance, and identifying directions for future research.

4.6 Discussion

No evidence was found to support the hypothesis that Pidgin primes facilitate processing of ‘ōlelo Hawai‘i targets more effectively than Hawai‘i English primes. The model also did not demonstrate evidence of a priming effect for ‘ōlelo Hawai‘i words overall, with neither Hawai‘i English nor Pidgin primes preceding ‘ōlelo Hawai‘i target items. The absence of a priming effect crucially points to several potential challenges in the study design, primarily including features of the word and nonword stimuli and the task paradigm that may have been less than ideal for the current language context.

The absence of a priming effect may in part reflect lexical frequency and proficiency effects documented in the L2 processing literature, where lower-frequency L2 words often elicit weaker or delayed priming effects (e.g., Chaouch-Orozco et al. [2021]). In the present study, these effects are most likely to arise at the level of the target language, as ‘ōlelo Hawai‘i was the sole target language in the task. As a result, participants’ proficiency and lexical familiarity in ‘ōlelo Hawai‘i are expected to exert a particularly strong influence on performance. Additional factors related to the priming languages (Pidgin and Hawai‘i English) are considered later in the discussion. Indeed, the target words in the current study were lower in frequency than ideal, potentially limiting participants’ familiarity, though they were known to the language consultants. Lowered frequency is expected to slow response times. In retrospect, some items may have been unknown to certain participants, increasing rate of participant exclusions, while with other participants, the words may have been known but more seldom used, contributing to longer response times. The high exclusion rate in Experiment 2 due to low task performance further highlights this issue. Part of the challenge in selecting higher-frequency target items stemmed from the need to avoid cognates between ‘ōlelo Hawai‘i targets and Pidgin primes, while also ensuring clear translation equivalents across all three languages and preventing overlap between Pidgin and Hawai‘i English items. Avoiding cognates between Pidgin and ‘ōlelo Hawai‘i was essential, as cognates could artificially inflate priming effects between Pidgin and ‘ōlelo Hawai‘i, confounding the measurement of true cross-language activation. Similarly, avoiding cognates between Pidgin and Hawai‘i English was important, since such items would not have been

sufficiently distinct from one another. Using the same corpora as in Experiment 2, we found that among the 1000 most frequent ‘ōlelo Hawai‘i words, approximately 30% overlap with Pidgin, underscoring the difficulty of selecting distinct stimuli. These constraints unfortunately eliminated many high-frequency words and substantially limited the available stimuli set. Although considerable effort was made to select translation pairings that were equivalent across ‘ōlelo Hawai‘i, Pidgin, and Hawai‘i English while choosing the highest-frequency items possible given the study constraints, it is likely that the words were still not frequent enough to produce a robust priming effect. Improving the stimulus set, for example by excluding low-frequency words, may help increase accuracy in future tasks. Additionally, incorporating a familiarity assessment, as in prior work (e.g., Soo and Monahan [2023], Lee-Ellis [2012]), before the start of participation could provide additional experimental control. Overall, this issue highlights the importance of selecting high-frequency target words for paradigms of this kind, particularly when the target language is an L2.

Feedback from participants in Experiment 2 supported this concern, with several noting informally to the researcher that the task was more difficult than that in Experiments 1 and 3. One Hawaiian language teacher who completed the study remarked that the task “seemed too hard” for students that could have met the criteria. Taken together, this suggests that the ‘ōlelo Hawai‘i targets may have been overly challenging overall for the given participant sample, thereby reducing the likelihood of detecting robust priming effects from either language. Although participants achieved high overall accuracy, the task may have been sufficiently difficult to produce reaction times that were longer than ideal, potentially reducing sensitivity to subtle experimental effects.

Relatedly, the lexical decision task, which entailed asking participants to decide whether a target ‘ōlelo Hawai‘i word was real or not, was identified as particularly challenging in an L2 context. One participant noted in informal feedback that although they were confident in their spoken Hawaiian, they were also aware that they did not know all Hawaiian words, which made judging whether a given word was real or not especially difficult. “I am aware there are many words I do not know,” the participant said, “I began to just accept more words than not as probably real”. Indeed, in studies such as Schilling et al. [1998], performance in lexical decision tasks was found to be more strongly influenced by word frequency than in naming tasks, suggesting that lexical decision may place greater demands on lexical access and familiarity. With Hawaiian being a minoritized language, the difficulty of the task may have been especially exacerbated, as even highly proficient

Hawaiian speakers may feel less confident in their ability to discern whether a word is a nonword or simply a word that is unfamiliar to them. This issue is further compounded by several additional factors, including questions of authenticity (what counts as a “real” word), dialectal or regional variation (disagreements over which words are considered “real”), limited exposure (even fluent speakers may have limited exposure to less common words), and so-called *linguistic insecurity* (proficient speakers may doubt their knowledge). Overall, the demands of the experimental paradigm may have made the task too challenging for many participants, thereby reducing the chances of observing a priming effect.

Another point raised pertained to the mode of which the Hawaiian items were presented in the study, specifically concerning the decision for all items to appear in spoken form only rather than written form. Several participants informally reported that in Experiment 2, it was difficult to differentiate long and short vowels or to detect the presence or absence of the glottal stop with isolated words in spoken Hawaiian. Certainly, these distinctions are more subtle for many English L1 speakers, and it is possible that these contrasts may not have been fully acquired by some lower-proficiency participants, whose exposure was more limited. Exposure has been linked to stronger language access and increased processing efficiency (e.g., Reichle and Perfetti [2003]). Even more proficient participants, who have likely well acquired these distinctions, may have been more inclined to accept a nonword differing in the presence or absence of a glottal stop or in vowel length as a potential variant pronunciation. Future work could address this issue by presenting targets in both written and auditory forms, while presenting primes only auditorily to avoid ambiguities related to Pidgin’s lack of a standardized orthography. Prior research supports this methodological approach; Wang and Tragant [2022], for instance, found that providing written text alongside spoken language can enhance listening comprehension by reducing ambiguity and supporting more accurate processing of phonemic contrasts.

Although no priming effect was found for either language, Pidgin may have been particularly vulnerable to reduced priming effects due to participants’ limited familiarity with certain primes, potentially stemming from dialectal variation. While Pidgin dialectal variation, to our knowledge, has not yet been empirically studied, the dictionaries used in this study as well as the researcher’s own expertise reflect a dialect of Pidgin rooted in Japanese ancestry, which may represent an ethnolect of Pidgin that includes a higher proportion of Japanese-derived words. As a result, some items may have been less familiar to participants of non-Japanese ancestry (e.g., the Pidgin word *buta* ‘pig’ is widely recognized among Pidgin speakers of Japanese ancestry,

whereas those of Hawaiian ancestry may be more likely to use the Hawaiian word *pua'a*). In addition, the experimental design required the selection of Pidgin words with Hawaiian cognates, which necessitated the exclusion of words originating in Hawaiian. Given that Hawaiian is the primary substrate language of Pidgin, this exclusion likely also removed many of the higher-frequency Pidgin words. While no frequency norms were available at the time of the study, for Pidgin, this exclusion plausibly reduced the overall familiarity of the remaining items for the participant sample recruited. Together, dialectal variation and restrictions on lexical selection may have reduced familiarity with some Pidgin primes, thereby weakening associations between primes and their translation equivalents compared to controls for some participants and increasing the variability between-participants. Indeed, prior research shows that more proficient speakers develop stronger word-concept associations than less proficient speakers (Kroll and Stewart [1994], Francis et al. [2014]), suggesting that limited access to certain primes due to unfamiliarity to a particular dialect of Pidgin could have further diminished priming effects.

A key limitation of this study was the lack of access to a Pidgin corpus and the challenges associated with the Hawaiian corpus, both of which complicated access to reliable word frequency information. The absence of a Pidgin corpus prevented systematic control for word familiarity in item selection, potentially contributing to variability in participant responses. Although the 'ōlelo Hawai'i corpus was extensive, it contained errors and inconsistencies that required substantial cleaning, which may have affected the true frequency values of words. Such limitations highlight the need for more reliable Pidgin and 'ōlelo Hawai'i corpora in future research to improve stimulus selection and frequency matching.

Participant variability may have also contributed to the diminished priming effects. Notably, as described in the procedures section, in this study critical and control trials for a given item were presented to different participants rather than within the same participant. This is necessary to avoid repetition of target items in the task. As a result, individual differences in familiarity, proficiency, or dialectal background could have introduced additional variability, potentially masking true priming effects. In other words, the absence of a significant effect for some items may reflect listener variation rather than a genuine lack of priming. This issue may have been further impacted by the relatively small sample size, which would increase the influence of participant-specific differences.

Based on the participant language background questionnaire, variability in reported proficiency differed across the three languages (see Figure 4.2).

English proficiency was consistently rated at the highest level, with most participants reporting maximum proficiency scores (a score of 10 out of 10) in all but one participant. Pidgin proficiency was also generally rated as high, though with greater variability across participants, some participants rating themselves much lower in proficiency. In contrast, ‘ōlelo Hawai‘i proficiency was much more widely distributed, with most participants reporting substantially lower proficiency than for either English or Pidgin, with a score of around 6 out of 10. Overall, these self-reported proficiency patterns may contribute to differences in relative processing across the three languages.

Considering that language priming effects may have been less likely to emerge due to overall task difficulty, particularly among participants with lower proficiency, we conducted a follow-up analysis focusing on a subset of participants identified as language consultants. Based on their status as language consultants, as well as their mean proficiency ratings, these participants were expected to have higher proficiency in ‘ōlelo Hawai‘i than the rest of the participant sample and therefore to be less affected by the challenges associated with low-frequency ‘ōlelo Hawai‘i words. These participants were filtered from the larger dataset, and the same Bayesian lognormal model (including in the Analysis and Results section) was rerun on this participant subset using the same reference levels, priors, chains, and iterations as in the primary analysis.

The results indicated that, even within this highly proficient subgroup ($n=7$), there was still no evidence of a priming effect for either Pidgin nor Hawai‘i English. The population-level estimates for the simple effects of prime language (Estimate = -0.05, 95% CI [-0.24, 0.14]) and trial type (Estimate = -0.09, 95% CI [-0.29, 0.10]), as well as their interaction (Estimate = 0.08, 95% CI [-0.17, 0.33]), all had credible intervals that included zero. See Table 4.6 for a full summary.

To assess whether any participants showed evidence of priming, we further examined individual participant posterior estimates for the effect of Pidgin priming. Across the seven participants all 95% credible intervals overlapped zero. This indicates that there was no strong evidence for priming at the individual level in this dataset. The lack of a priming effect among the highly proficient participants indicates that design limitations, including the choice of stimuli, likely played a significant role in the null findings, rather than participants’ proficiency in Hawaiian.

To further explore potential factors limiting priming, we examined accuracy as a function of prime language and trial type (translation primes versus controls). Including accuracy measures is consistent with prior work on cross-language priming, which often investigates accuracy more in-depth

4.6. Discussion

Predictor	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept	6.33	0.22	5.88, 6.77	1.00	4977	6358
Prime Language (Pidgin)	-0.05	0.10	-0.24, 0.14	1.00	9033	8973
Trial Type (Cross-language Translation Equiv)	-0.09	0.10	-0.29, 0.10	1.00	8539	8745
Prime Language $\times$ Trial Type	0.08	0.13	-0.17, 0.33	1.00	8989	9314

Table 4.6: Population-level effects from the Bayesian lognormal model for language consultant participants.

alongside RT analyses in studies where performance is not at ceiling (e.g., Smith et al. [2019], Kiran and Lebel [2007], Dimitropoulou et al. [2011]). This analysis allowed us to investigate whether participants correctly recognized or processed the items, providing additional insight beyond the reaction time results. To assess whether reaction time effects were accompanied by systematic differences in accuracy, we conducted a Bayesian mixed-effects logistic regression on trial-level accuracy, with prime language, trial type, and their interaction as predictors, and random intercepts for participants and items. This analysis evaluates potential speed–accuracy tradeoffs and contextualizes the absence of priming effects in the reaction time data. For this model, accuracy was coded as a binary variable (correct = 1, incorrect = 0). These data included only participants that scored above 80% accuracy (including both Hawaiian language consultants and non-consultant participants). We fit the model using the `brms` package (Bürkner [2017]). We modeled accuracy using a Bernoulli distribution with a logit link, due to accuracy being a binary variable. Weakly informative priors were specified for the fixed effects (Normal(0,1)), random effect standard deviations (Cauchy(0,1)), and the intercept (Normal(0,1)), following recommendations for binary outcome mixed-effects modeling in Bayesian frameworks (McElreath [2020]). Model fitting was performed using four chains with 6000 iterations each, 1000 warm-up iterations, and a thinning factor of 2. The reference levels were Hawai‘i English (for prime language) and cross-language control trials (for trial type).

Population-level effects from the Bayesian binomial regression model are summarized in Table 4.7. The intercept (Estimate = 3.15, 95% CI: [2.52, 3.80]) represents the baseline log-odds of a correct response for the reference trial type (cross-language control trials) and prime language (Hawai‘i English). Across prime languages, there was no evidence of an accuracy difference when primes were in Pidgin compared to Hawai‘i English (Estimate = -0.26, 95% CI: [-0.82, 0.31]). Comparing cross-language translation equivalent trials with control trials revealed no evidence of a meaningful difference in accuracy, with translation-equivalent prime-target pairs performing similarly to control trials that were not equivalent in terms of meaning (Estimate

= 0.06, 95% CI: [-0.62, 0.78]). Similarly, the interaction between prime language and the cross-language translation equivalent trials was not credibly different from zero (Estimate = 0.30, 95% CI: [-0.60 to 1.24]), suggesting no reliable effect of accuracy by prime language in this trial type.

Interactions between prime language and trial types also did not show credible deviations from zero across all trial types (Translation Equivalents: Estimate = 0.30, 95% CI: [-0.60, 1.24]; nonword-nonword fillers: Estimate = 0.24, 95% CI: [-0.65, 1.11]; nonword-word fillers: Estimate = 0.26, 95% CI: [-0.69, 1.20]; Word-nonword fillers: Estimate = -0.10, 95% CI: [-0.98, 0.75]); and Word-word fillers: Estimate = -0.10, 95% CI: [-1.31, 1.12]). This pattern suggests that prime language did not differentially influence accuracy across the various trial types.

Examining the various filler trial types, several effects were observed depending on whether the ‘ōlelo Hawai‘i target was a real word or a nonword. nonword–nonword and word–nonword trials were associated with significantly lower odds of correct responses relative to control real words (Estimate = -0.87, 95% CI: [-1.65, -0.09]; Estimate = -1.04, 95% CI: [-1.83, -0.26], respectively), whereas word–word and nonword–word filler trials showed estimates overlapping zero, indicating no reliable accuracy differences. These findings support the pattern observed in the initial accuracy results for participants scoring above 80% (Tables 4.3 and 4.4), where nonword targets in ‘ōlelo Hawai‘i were generally more difficult and elicited lower accuracy than real word targets. In other words, participants appear to exhibit a strong bias toward real words. This pattern is consistent with extensive lexical decision research showing that nonwords are generally harder to classify than words, particularly when they resemble real lexical items or require stronger inhibitory control during decision making (e.g., Samuel and Larraza [2015]; Sumner and Samuel [2007]; Lupker and Pexman [2010]).

This pattern also aligns with a language consultant’s comment during the study: “I am aware there are many words I do not know,” highlighting the challenge participants face in judging word validity. These findings highlight the importance of careful stimulus selection for both words and nonwords, and draw attention to the potential bias that arises when nonwords are misclassified as words in lexical decision tasks. Such challenges may be particularly pronounced in minoritized language contexts, as in the present study. Alternative experimental paradigms may therefore provide a more reliable means of eliciting responses while reducing participant uncertainty.

In summary, trial type, prime language, and their interaction showed no credible effects for accuracy. Instead, accuracy outcomes reveal that

4.6. Discussion

Predictor	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept	3.15	0.33	2.52, 3.80	1.00	5019	7121
Prime Language (Pidgin)	-0.26	0.29	-0.82, 0.31	1.00	6662	8746
Trial Type (translation equivalents)	0.06	0.35	-0.62, 0.78	1.00	7918	8214
Trial Type (nonword-nonword fillers)	-0.87	0.40	-1.65, -0.09	1.00	4737	7246
Trial Type (nonword-word fillers)	0.21	0.41	-0.58, 1.01	1.00	5186	7457
Trial Type (word-nonword fillers)	-1.04	0.40	-1.83, -0.26	1.00	4816	6865
Trial Type (word-word fillers)	0.78	0.52	-0.21, 1.80	1.00	6176	7887
Prime Language (Pidgin) x Trial Type (translation equivalents)	0.30	0.47	-0.60, 1.24	1.00	7150	8848
Prime Language (Pidgin) x Trial Type (nonword-nonword fillers)	0.24	0.44	-0.65, 1.11	1.00	4628	6549
Prime Language (Pidgin) x Trial Type (nonword-word fillers)	0.26	0.48	-0.69, 1.20	1.00	5355	7796
Prime Language (Pidgin) x Trial Type (word-nonword fillers)	-0.10	0.44	-0.98, 0.75	1.00	4253	7065
Prime Language (Pidgin) x Trial Type (word-word fillers)	-0.10	0.61	-1.31, 1.12	1.00	6680	8315

Table 4.7: Population-level effects from the Bayesian binomial regression model of accuracy with participants that scores about 85% or more.

error rates were primarily driven by the status of the target item: nonwords consistently elicited lower accuracy than real words. In the following section, I examine nonword accuracy in greater detail.

Closer examination of the relationship between nonword stimuli and accuracy suggests that, according to prior research, highly word-like nonwords can slow response times and reduce accuracy in word-nonword discrimination tasks (e.g., Lupker and Pexman [2010], Yap et al. [2014]) and may amplify word frequency effects for real words (Stone and Van Orden [1993]). Additionally, research has shown that nonwords are more difficult for participants to reject when they are composed of existing morphemes compared to when they are not (Taft and Forster [1975]), suggesting that plausible nonwords can interfere with lexical processing and potentially reduce priming effects. Due to the methodology used to create nonwords, some nonwords may have inadvertently resembled real Hawaiian words. This was particularly likely given use of the Uni-Pseudo concatenation tool, which broke apart real words and attached them to other real words. It is thus probable that the use of word-like nonwords also contributed to variability in response times and overall accuracy, potentially attenuating observed priming effects.

Returning to the issue of high exclusion rates, among participants who met the formal qualifications but were excluded for scoring below 80% accuracy, all had either: been currently enrolled in Hawaiian 201 or 202, completed the equivalent of Hawaiian 102 in high school (which may have been too long ago or not equivalent to the current 102 university standard), or reported being largely self-taught, indicating that although they met the qualifications on paper, they had not actively used Hawaiian for an extended period. We do not believe dialectal variation played a significant role in lower accuracy, as none of the participants self-reported speaking the Ni‘ihau dialect of Hawaiian. Nonetheless, the fact that accuracy did not cluster near

90% or higher among qualifying participants that met the accuracy threshold (scoring 80% or higher) still warrants caution, underscoring complicating factors. Future work could reduce exclusion rates by enforcing stricter qualifications and encouraging meaningful responses, such as by running the study in person rather than online.

Overall, the limitations observed in this study appear to arise primarily from challenges in stimulus selection for words and nonwords, and in the task paradigm. While there are numerous ways future research could refine the investigation of whether Pidgin facilitates priming of Hawaiian more than Hawai'i English, the present study was constrained particularly by limited resources for selecting appropriate stimuli. Ensuring that primes and targets across the three languages were matched in meaning, consisted only of high-frequency words, and avoided cross-language cognates or borrowings between prime and target for all three languages imposed substantial constraints on the items that could be used. Together, these factors likely contributed more to the absence of robust priming effects, highlighting the practical challenges of conducting psycholinguistic research in minoritized language contexts.

4.7 Conclusion

In sum, Experiment 2 did not provide statistical evidence of a priming advantage of Pidgin over Hawai'i English for 'ōlelo Hawai'i targets, nor was any priming effect observed overall, regardless of language. This pattern was observed both in the broader population of participants who performed above 80% and in the smaller subgroup of Hawaiian language consultants. While participant variability and the limited sample size may have further obscured potential effects, the main limitations arose largely from challenges in stimuli selection (for both lower frequency words and nonwords). At the time of this study, the constraints of creating suitable stimuli were particularly prohibitive given the language context (i.e., the relation between Pidgin and Hawaiian). Ensuring that prime–target pairs were meaning-matched, high frequency, and lacking cross-language cognates imposed considerable restrictions on item selection. These constraints were particularly onerous given the limited pool of eligible words that met all criteria and the lack of resources providing sufficient information, such as a large, error-free written corpus containing frequency norm data for either Pidgin or Hawaiian. These findings highlight the unique methodological difficulties of conducting experimental psycholinguistic research in minoritized language contexts. Future work will benefit from improved stimulus design that reduce participant un-

certainty and better capture the dynamics of cross-language processing in ‘ōlelo Hawai‘i.

4.8 Mo‘olelo

Experiment 2 tells a broader story about what it means to conduct research within the scope of community. The challenges encountered in developing stimuli reflected the broader context in which language is used, understood, and studied in Hawai‘i. One challenge was the limited availability of accessible and reliable language resources. Unlike widely studied languages with extensive corpora, frequency databases, and standardized resources, ‘ōlelo Hawai‘i and Pidgin have been shaped by histories of unequal documentation and limited institutional support for their inclusion in linguistic research. The limited availability of community-accessible Pidgin resources reflects broader patterns in which linguistic knowledge was often produced and maintained within academic spaces rather than returned to the community.

Creating open-access language resources can advance linguistic research while also serving the needs of the community. These resources can allow community members to explore their own questions about language and support educational efforts, whether in ‘ōlelo Hawai‘i classrooms or in creative writing contexts where documented examples of Pidgin can be used as teaching materials.

The findings from Experiment 2 also call attention to the importance of prioritizing how language is experienced and used in the community. The results from this study suggest that, for these participants, the Pidgin–‘ōlelo Hawai‘i and English–‘ōlelo Hawai‘i word pairs did not show strong connections in the way measured by this task. Observing which words are frequently used within the community, the contexts in which they occur, and the speakers who use them could strengthen future studies. While certain decisions were prioritized to meet experimental constraints, these choices may have resulted in stimuli that were less representative of how these words are encountered and understood in everyday community use. Thus, future work should place greater importance on how words are experienced within the community to better understand what words and language-pairings most accurately reflect community language use. Returning to the process of creating appropriate stimuli, research must continue to consider not only how words are represented, but also how they are used and understood by speakers themselves. Developing resources and approaches that more accurately

reflect community language use can strengthen both the empirical contributions of this work and its relevance to the communities whose languages are being studied. Paying greater attention to this could also help create research experiences that better reflect and respect participants’ own language knowledge, making participation itself an experience that affirms their linguistic knowledge and leaves participants with a greater sense of confidence in their abilities.

In the next chapter, I will explore the connections between Pidgin and ‘ōlelo Hawai‘i using a completely different approach, by assessing the potential advantage Pidgin speakers may have in accessing the Hawaiian protolexicon compared to non-Pidgin Hawai‘i English speakers.

Chapter 5

Experiment 3: Implicit knowledge of ‘ōlelo Hawai‘i by Pidgin speakers

5.1 Introduction

Through the American colonial gaze and the myth of monolingualism, English is positioned as the dominant and defining language of modern Hawai‘i, reinforcing its association with the United States (Asensio [2010]). In contrast, the linguistic reality of Hawai‘i is far more complex, where Pidgin is spoken as the majority language (Sakoda and Siegel [2003]), co-existing alongside numerous immigrant languages primarily from Asia and the Pacific, and at the heart of the linguistic landscape lies ‘ōlelo Hawai‘i (Saft [2022]), lauded as one of the most successful revitalization efforts in the world (Wilson and Kamanā [2011]). A key marker of the Hawaiian language revitalization movement’s success lies in its ambient presence in Hawai‘i. While many kama‘āina are not (yet) speakers of ‘ōlelo Hawai‘i, the language permeates daily life through public signage, media, politics, education, and community spaces and all residents of Hawai‘i are exposed to the language to varying degrees (Saft et al. [2018]). Previous studies have shown that ambient exposure enables adults in Aotearoa who do not speak Māori to develop a *proto-lexicon* of the language (e.g., Oh et al. [2020]), extending work on ambient learning originally developed in the context of child language acquisition. The proto-lexicon is defined as implicit form-based knowledge, including word structure, morphemes, and sound sequences, that are learned in the absence of associated semantic representations, often through exposure to statistical regularities (Johnson [2016], Martin et al. [2013], Ngon et al. [2013]). While first proposed to account for early stages of first language acquisition (Pierrehumbert [2003b], Curtin and Archer [2015]), the construct has since been extended to capture proto-lexical knowledge that emerges in adults in second or additional language acquisition under conditions of

sustained ambient and community-level exposure. Crucially, unlike the acquisition of the lexicon, the proto-lexicon emerges from speech input and top-down lexical information, without requiring explicit form-meaning mappings (Panther et al. [2023b], Martin et al. [2013]). Experiment 3 examines the ambiently learned linguistic knowledge of ‘ōlelo Hawai‘i for adults who do not identify themselves as ‘ōlelo Hawai‘i speakers and do not report having taken any formal Hawaiian language classes, comparing performance on two experimental tasks across individuals with and without Pidgin knowledge, both from Hawai‘i and from outside Hawai‘i, relative to fluent ‘ōlelo Hawai‘i speakers. A comparison of these participant groups raises the research question: Is Pidgin associated with greater implicit knowledge of ‘ōlelo Hawai‘i compared to English? If Pidgin indeed is associated with greater implicit knowledge of ‘ōlelo Hawai‘i, participants with Pidgin experience would be expected to outperform non-Pidgin speakers of *English*⁶.

Research on language acquisition highlights the foundational role of the proto-lexicon in both first and additional languages, providing a basis for the subsequent learning of meaning. The majority of research on the proto-lexicon has centered on first language acquisition. Infants acquire the fundamentals of the language(s) to which they are exposed rapidly and without explicit instruction. Many foundational linguistic skills develop before a child is able to speak (Curtin and Archer [2015]), reflecting sensitivity to structured speech input prior to the acquisition of explicit lexical knowledge (Pierrehumbert [2003a]). In fact, exposure to native language(s) begins prior to birth, with newborns already able to differentiate between prenatally experienced languages and unfamiliar ones (Mehler et al. [1988], Nazzi et al. [1998], Byers-Heinlein et al. [2010]).

During the first year of life, children begin developing a range of perceptual and segmentation abilities that support early proto-lexical development. They can perceive native-language speech sounds and distinguish them from sounds outside their native phonological inventory (e.g., Kuhl et al. [2008]; Werker and Tees [1999], Johnson [2016], Kaushanskaya et al. [2013]); extract words from continuous speech (Jusczyk and Aslin [1995]); and identify sub-word units including morphemes, sounds, sound sequences, and syllables (Johnson et al. [2014]; Jusczyk et al. [1993]; Saffran et al. [1996b]). Many of these learned skills are argued to rely on the exploitation of statistical cues.

⁶In Experiment 3, because no auditory stimuli were present, we refer more generally to the language as *English*, and distinguish speakers as either English speakers from Hawai‘i or English speakers from Canada. These broad categories may include different varieties of English but are defined here according to participant criteria, specifically time spent living in Hawai‘i.

For instance, knowledge of syllable transitional probabilities and common word shapes supports identification of syllable and word boundaries (Saffran et al. [1996b]); distributional learning of sounds aid in early recognition of allophones (Jusczyk et al. [1999]); and sensitivity to word frequencies helps infants identify lexical categories (Marino et al. [2020] as well as recognize familiar high-frequency words and help discriminate likely real words from nonwords (Bortfeld et al. [2005], Jusczyk and Aslin [1995], Bergelson and Swingley [2012]).

Probabilistic phonotactics is the frequency with which phonological segments and segmental sequences occur within syllables and words in a given language (Vitevitch and Luce [2005]). It is well established that listeners are sensitive to probabilistic differences among phonotactic patterns in their first language (e.g., Treiman et al. [1996], Vitevitch et al. [1997]). Speakers of a language are able to evaluate whether word forms are more or less probable by generalizing statistical patterns over known words (Hay et al. [2004]). Many studies have demonstrated sensitivity to sound patterns is exhibited in infancy (e.g., Jusczyk et al. [1994], Mattys et al. [1999], Mattys and Jusczyk [2001]). By nine months of age, infants show a preference for frequent over infrequent phonotactic patterns (Jusczyk et al. [1994]), and those with larger vocabularies exhibit greater sensitivity to phonotactic likelihood (Fenson et al. [2000]). This suggests that probabilistic phonotactic awareness arises from statistical generalizations over known words (Edwards et al. [2004], Frisch et al. [2001]) and is closely tied to vocabulary growth. Probabilistic phonotactics provide a useful means of probing the proto-lexicon, as they provide insight into knowledge of word forms independent of knowledge of meaning. Sensitivity to these patterns reflects continuously developing skills that support proto-lexical processing, including tracking sub-word and word boundaries and inferring possible word forms (Archer and Curtin [2016], Mattys and Jusczyk [2001], Zamuner [2009]).

It has also been demonstrated across several studies that knowledge of probabilistic phonotactics, or *phonotactic awareness*, in the ambient language contributes to early word learning in first language acquisition. In Storkel and Rogers [2000], phonotactic probability was found to affect the rate of word learning such that high-frequency sound sequences were learned more rapidly and required less exposure than low-frequency sound sequences. Phonotactic probabilities also were found to influence the speed and strength of the associations between form and meaning (Storkel [2001]). Sensitivity to phonotactic probability has been argued to be a necessary step in mapping word form to meaning, with findings showing that prior segmentation experience, rather than mere exposure frequency, facilitates word learning

(Hay et al. [2011], Graf Estes et al. [2007]). This suggests that statistical phonotactic learning of word forms supports subsequent word learning and provides a foundation for vocabulary development (Graf Estes et al. [2016]).

Evidence also suggests that adults exhibit statistical knowledge of word form and sound sequences in the acquisition of second or additional languages. Adult participants who were exposed to an artificial language for only ten hours were still able to recognize high-frequency words three years later (Frank et al. [2013]). Across multiple studies, adult participants exposed to an artificial language demonstrated the use of statistical segmentation strategies to identify word boundaries in continuous speech (Pena et al. [2002], Onnis et al. [2005], Newport and Aslin [2000]). These findings are consistent with the existence of a proto-lexicon in adults and suggest that proto-lexical knowledge precedes the establishment of stable form-meaning mappings. In this view, learners maintain an implicit repository of distributional regularities over word forms and sub-lexical patterns, which supports early structural learning prior to the acquisition of lexical meaning in adult second language learning. This suggests that adults learning an additional language rely on statistical cues to support foundational perceptual and segmentation processes that precede, and potentially scaffold, mappings between form and meaning.

The adult proto-lexicon has also been observed in natural language contexts. For example, Kittleson et al. [2010] found that individuals that did not speak Norwegian were able to discriminate between words and nonwords at an accuracy well above chance after a brief familiarization period with Norwegian continuous speech, followed by a spoken list of isolated target words. According to Kittleson et al. [2010], these results demonstrate that adult listeners can use statistical regularities generalized from implicit exposure to a language they do not speak to infer word boundaries and segment words from continuous speech, and can also use stress differences to distinguish likely words from nonwords. These findings are consistent with the development of a proto-lexicon in adults through passive exposure, in which statistical knowledge of sound patterns supports early word segmentation and enables learners to distinguish words from nonwords even without full knowledge of the language.

In Saffran et al. [1996c], adults were found to utilize statistical regularities in speech, including syllable transitional probabilities, to support links between word form, sublexical structure, and later lexical learning, in a manner similar to that observed in children in the earlier study (Saffran et al. [1996a]). These distributional cues were used to segment novel words in the artificial language. As in infants, Mirman et al. [2008] found a direct

link between statistical segmentation ability, the storage of such sound sequence knowledge, and word learning in adults, such that novel object–label pairings were learned more quickly when labels contained higher-probability syllable transitions. Similarly, in Sulzberger [2007], non–Russian speaking adults were more likely to recognize Russian words containing phonotactically probable sound patterns than those with less probable ones after exposure to Russian speech in the laboratory.

It has also been found that phonotactic awareness is reflected in explicit judgments of nonword *well formedness*, or the extent to which a nonword conforms to the phonological patterns of real words (Frisch et al. [2000]). Speakers of a given language are able to distinguish possible sound patterns from impossible ones with a high degree of accuracy and demonstrate precise knowledge of the gradient well formedness of nonwords in the languages they speak, such that their ratings of how word-like a nonword is are highly correlated with corpus-based measures of phonotactic probabilities (Brown and Hildum [1956], Coleman and Pierrehumbert [1997], Frisch et al. [2000], Hay et al. [2004], Kirby and Yu [2007], Vitevitch and Luce [1999]). These sensitivity patterns are often interpreted as reflecting an implicit statistical representation of permissible sound sequences, suggesting that listeners encode distributional regularities of their language in a way that abstracts beyond individual lexical items. From the perspective of the proto-lexicon, such representations can be understood as an early form of structured word knowledge, in which learners track probabilistic phonotactic constraints across exposure and thereby build pre-lexical representations of word-like forms that are not yet mapped to meaning but are nonetheless organized according to language-specific phonological regularities.

Word–nonword identification tasks can be used to probe the proto-lexicon of language speakers. Speakers with larger lexicons tend to be better at discriminating words in their native language (Yap et al. [2012]). However, word–nonword tasks do not require knowledge of meaning, as these tasks allow researchers to examine the presence of word forms and the organizational relationships among them, rather than their detailed semantic representations (Whaley [1978]). Word–nonword tasks reveal the existence of word form representations, and more gradient confidence ratings allow the potential to capture knowledge of weak or “fuzzy” forms and form-to-meaning mappings (Cook and Gor [2015]; Cook et al. [2016]). The Fuzzy Lexical Representation Hypothesis predicts that lexical representations may contain underspecified or imprecise semantic information despite relatively stable word-form representations (Cook and Gor [2015]; Jiang [2000]; Llompарт [2021]). Consequently, speakers who are more accurate at discriminating

words from nonwords are likely to have a larger proto-lexicon (Mattingley et al. [2024b]).

Word–nonword identification tasks have also revealed factors that bias the selection of a word as real or not. These includes word length (Forster and Chambers [1973], Whaley [1978], Balota et al. [2004]), word frequency (Stanners et al. [1975], Perea et al. [2005], Duchek and Neely [1989], Ziegler et al. [2001])), and neighborhood frequency (Andrews [1989], Andrews [1992], Balota et al. [2004]), suggesting that such factors interact with competing lexical representations. This highlights the need to control these factors (e.g., Yap et al. [2014]).

Beyond languages that have been learned as a first or additional language, or through familiarization in laboratory settings, Bice and Kroll [2019] found that self-identified “monolingual” adults living in linguistically diverse environments were more sensitive to novel phonological patterns than those in homogeneous environments. Specifically, participants in diverse contexts more successfully learned a Finnish vowel harmony rule, suggesting that passive exposure to multiple languages enhances phonological learning ability even without active bilingual experience.

Similarly, Oh et al. [2020] found that participants who merely lived in a region demonstrated greater proto-lexical knowledge than those who did not. In their study, participants residing in Aotearoa who did not speak Māori (NMS-A) nonetheless showed substantial implicit knowledge of Māori in two proto-lexical tasks. In a well formedness rating task, the NMS-A participants performed nearly as well as Māori speaking participants (MS), and significantly outperformed non-Māori speaking participants from the United States (NMS-U). In a word identification task, NMS-A participants were able to distinguish real Māori words from highly Māori-like nonwords, whereas NMS-U participants were not. This suggests that NMS-A participants had implicit lexical knowledge that could be accessed to differentiate real Māori words from Māori-like nonwords. Moreover, they appeared able to generalize over this proto-lexicon to generate phonotactic knowledge and assess how well-formed a given nonword was in Māori, allowing them to perform nearly as well as the MS speakers.

Since Oh et al. [2020], several studies have investigated the implicit knowledge held by adults due to ambient exposure. While Oh et al. [2020] examined the two proto-lexical tasks with separate participant groups, Panther et al. [2023b] directly linked them by administering both tasks to the same participants. Those with larger proto-lexicons, as indicated by their performance in the word identification task, showed greater sensitivity to phonotactic patterns in the well formedness rating task. This suggests a

connection between the two measures, linking form and statistical sensitivity.

The impact of the proto-lexicon accumulated from one’s ambient environment has been found to lead to significant benefits in language acquisition. In artificial experimental contexts, implicit exposure to an L2 has been shown to facilitate faster subsequent explicit vocabulary learning in adult learners (e.g., Bisson et al. [2013], Gullberg et al. [2010]). Prior research has likewise demonstrated that phonotactic and proto-lexical knowledge, built up through lifetime exposure, can influence explicit word learning in real-world contexts. In Mattingley et al. [2024b], implicit linguistic knowledge represented in the proto-lexicon was found to provide an advantage for explicit language learning in a classroom environment. Participants with larger Māori proto-lexicons were shown to learn more words in the classroom, demonstrating a link between proto-lexical knowledge and subsequent vocabulary acquisition. This was argued in Mattingley et al. [2024b] as evidence that adults are able to unconsciously extract statistical regularities from repeated exposure to running speech in a language they do not speak and store these patterns in long-term memory, such that these representations can later be reactivated to support learners in mapping meanings onto word forms.

While Mattingley et al. [2024a] do not specify whether their participants spoke Māori English or Pākehā English, two dialects associated with distinct ethnic groups, related literature has reported that the Māori English dialect may be associated with a processing advantage, in addition to effects of exposure to *te reo Māori* (the Māori language). Szakay et al. [2016] found that Māori English primes Māori more effectively than Pākehā English. However, it is not explicitly clear whether this advantage arises from greater exposure to Māori among speakers of Māori English, or from the fact that Māori English is typologically closer to Māori than Pākehā English.

The current study, Experiment 3, utilizes two tasks designed to probe linguistic knowledge: a well formedness rating task and a word-nonword identification task. These tasks were incorporated to assess participants’ ‘ōlelo Hawai‘i linguistic knowledge as a whole. While these tasks may either tap into lexical or proto-lexical representations, neither task requires access to meaning, relying instead on knowledge of word form. As such, performance on these two tasks reflects generalizations of linguistic knowledge, including lexical and proto-lexical representations, such that a larger overall linguistic repertoire is associated with responses that more closely reflect the phonotactic reality of ‘ōlelo Hawai‘i, as well as with more confident, accurate performance on word identification. In line with prior work in this area (e.g., Oh et al. [2020], Panther et al. [2023b], Mattingley et al. [2024a]), we adopt

the general term proto-lexicon to describe the knowledge minimally required by non-‘ōlelo Hawai‘i speakers to engage in the well formedness rating task and word–nonword identification task. We note, however, that performance may also reflect lexical knowledge acquired through ambient learning. To provide a concrete example, it is common to encounter ‘ōlelo Hawai‘i word forms in use without knowledge of their meaning (e.g., “maika‘i” in “Maika‘i Rewards,” a marketing term used by a local grocery store chain, without understanding the meaning of maika‘i). However, in other cases, ambient exposure includes an understanding of the associated meaning (e.g., aloha). Both lexical and proto-lexical knowledge play a role in more Hawaiian-like well formedness ratings and in distinguishing words from nonwords; however, only proto-lexical knowledge is required for successful performance on these tasks.

Performance on the two tasks will be compared for individuals with and without Pidgin knowledge to ‘ōlelo Hawai‘i speakers. A comparison of these participant groups raises the research question: Is Pidgin associated with better phonotactic knowledge of ‘ōlelo Hawai‘i than English? Unlike previous work, which examines proto-lexical knowledge solely via environmental exposure, our study compares Pidgin and English speakers from Hawai‘i to identify differences in their implicit knowledge of ‘ōlelo Hawai‘i. In addition to this, I also compare the behavior of Pidgin and English speakers from Hawai‘i to speakers of English without exposure to ‘ōlelo Hawai‘i. The hypothesis is that there may be a gradient effect, where Pidgin speakers’ proto-lexicons are most similar to Hawaiian speakers, English speakers from Hawai‘i are second most similar, and English speakers without exposure are the least similar to ‘ōlelo Hawai‘i speakers in phonotactic tasks. If Pidgin knowledge generalizes to better implicit knowledge of ‘ōlelo Hawai‘i, we expect participants with Pidgin knowledge to be closer to Hawaiian speakers in regard to their phonotactic awareness than non-Pidgin speakers of English.

Pidgin knowledge may support the development of an ‘ōlelo Hawai‘i proto-lexicon through several interrelated dimensions, including linguistic similarity, frequent co-occurrence, and social connection. First, the significant structural overlap between Pidgin and ‘ōlelo Hawai‘i, including phonological patterns, phonotactic constraints, and lexical features (e.g., Grama [2015], Sakoda and Siegel [2008]), may enable Pidgin speakers to draw on typologically similar representations when recognizing Hawaiian forms. While a Pidgin speaker may not be a speaker of ‘ōlelo Hawai‘i, they may nonetheless possess partial knowledge of the language due to the structural overlap between the languages, including phonological, phonotactic, and lexical representations. This partial knowledge of ‘ōlelo Hawai‘i may therefore sup-

port stronger generalization than that observed for speakers of English alone, which shows less structural overlap with ‘ōlelo Hawai‘i and thus affords fewer relevant representations than does Pidgin.

Second, there is evidence to suggest that ‘ōlelo Hawai‘i and Pidgin are socially co-associated, in opposition to English. This may result in Pidgin speakers encountering ‘ōlelo Hawai‘i more frequently in their everyday environments, thereby increasing ambient exposure to the language. Relative to English speakers, this increased exposure may provide more opportunities for Pidgin speakers to acquire lexical and sublexical representations of ‘ōlelo Hawai‘i, thereby expanding both lexical and proto-lexical knowledge. Qualitative work describes the use of ‘ōlelo Hawai‘i and Pidgin as a linguistic means to simultaneously align with “local” Hawaiian values, establishing belongingness to the in-group (Sato [1991b], Marlow and Giles [2008], Marlow and Giles [2010b]), and both are used to position oneself in shared opposition to threats of American assimilation (Reinecke [1938]). Thus, speakers may encounter and use both languages within the same domains. Further, because both Pidgin and ‘ōlelo Hawai‘i are closely tied to “local” and Hawaiian identities, Pidgin speakers may be more likely than English speakers, who may be more strongly aligned with U.S.-dominant cultural ideologies and values (e.g., Sato [1991b]), to seek out ‘ōlelo Hawai‘i media, participate in activities involving ‘ōlelo Hawai‘i such as hula or mele, or attend events where ‘ōlelo Hawai‘i is spoken. A Pidgin speaker may also be more likely to be a person who places greater value on ‘ōlelo Hawai‘i. These factors may increase the opportunities Pidgin speakers may have to develop their ambient knowledge of ‘ōlelo Hawai‘i as well as increase their alignment toward the language. Stronger valuation and more positive attitudes toward a language within a society, alongside exposure, have previously been shown to be empirically linked to a stronger proto-lexicon (Todd et al. [2023a]).

Additionally, because Pidgin speakers are more likely to come from Kānaka Maoli and/or intergenerational local families, they may have closer ties with ‘ōlelo Hawai‘i speakers through family members, friends, and broader community networks and thus have greater exposure to ‘ōlelo Hawai‘i. Pidgin speakers may additionally be more likely to reside in neighborhoods where ‘ōlelo Hawai‘i is commonly heard in public spaces, compared to speakers from more English-dominant backgrounds. This connection between Pidgin and ‘ōlelo Hawai‘i identity may make it more likely that a Pidgin speaker has been heavily exposed to ‘ōlelo Hawai‘i, thereby increasing their opportunities to develop their ambient knowledge of ‘ōlelo Hawai‘i.

It is important to note that the current experimental design does not allow these factors to be disentangled. From this study alone, we cannot

determine the source of any observed advantage for Pidgin speakers, should one emerge. If a proto-lexical advantage is observed, it could be the result of linguistic overlap, co-association and increased exposure, social identity factors, or a combination of these determinants.

Unlike Experiments 1 and 2, Experiment 3 focuses on proto-lexical knowledge in Pidgin-speaking adults who are not yet fluent in ‘ōlelo Hawai‘i, using tasks designed to assess implicit linguistic knowledge rather than overt perception in self-identified ‘ōlelo Hawai‘i speakers. This allows us to understand the benefits of Pidgin in a new way, highlighting how prior language experience with Pidgin may support the implicit protolexical learning of ‘ōlelo Hawai‘i beyond formal instruction.

In the current study, relative ‘ōlelo Hawai‘i knowledge across speaker populations is probed using orthographic representations of the language rather than in spoken form. In prior Māori proto-lexical studies (e.g., Oh et al. [2020]; Panther et al. [2023b]; Mattingley et al. [2024b]), stimuli were similarly presented in written rather than auditory form. As argued by Panther et al. [2023b], it is unlikely that participants’ performance in these tasks primarily reflected exposure to written Māori-only as opposed to Māori more broadly, given that non-speakers in New Zealand are far more frequently exposed to spoken or sung Māori than to written text. Furthermore, written Māori has a highly transparent orthography, in which graphemes map closely onto phonemes. Previous work on the proto-lexicon has demonstrated that reading engages phonological and lexical awareness comparably to spoken language, particularly in transparent orthographies. (Ziegler and Goswami [2005]). This allows participants to access sound-structure information in a consistent and controlled manner, without variability introduced by differences in dialectal differences, pronunciation, or recording quality. Collectively, these considerations motivate the use of orthographic stimuli in the present study as a proxy for access to phonological structure.

The current research question is motivated by two primary community-driven goals. First, because Pidgin is the predominant L1 in Hawai‘i and for many Kānaka Maoli (Sakoda and Siegel [2003]), the typical L2 learner of ‘ōlelo Hawai‘i is a Pidgin speaker, followed by English speakers from Hawai‘i. Understanding how Pidgin and English speakers engage with ‘ōlelo Hawai‘i has broad community implications, including strengthening revitalization efforts and supporting equitable access to ‘ōlelo Hawai‘i education for Pidgin speakers. Investigating speaker backgrounds from Pidgin and English speakers from Hawai‘i has the practical benefit of addressing the largest language communities in the islands, as well as the groups most likely to desire to learn ‘ōlelo Hawai‘i.

Second, Pidgin is often stigmatized as a language that hinders the acquisition of “proper English” or ‘ōlelo Hawai‘i. Yet, Hawaiian language teachers have observed the opposite in classroom contexts (e.g., Martin [2018]; Rozet [2017a]; Rozet [2018]; Nordstrom [2018]; Solomon [2018]). Recognizing the linguistic and cognitive benefits of Pidgin knowledge challenges long-held misconceptions and affirms the value of Pidgin as a bridge to learning ‘ōlelo Hawai‘i. Ultimately, this work contributes to understanding how to better support Pidgin speakers in taking the next step toward ‘ōlelo Hawai‘i learning and revitalization.

5.2 Participants

Unlike experiments 1 and 2, experiment 3 did not require the ability to speak Hawaiian. Five groups were recruited based on language background and regional background: 1) speakers of ‘ōlelo Hawai‘i, 2) Pidgin speakers that do not speak ‘ōlelo Hawai‘i from Hawai‘i (lived in Hawai‘i for at least 1 year), 3) Pidgin speakers that do not speak ‘ōlelo Hawai‘i from the diaspora (that have not lived in Hawai‘i for at least 1 year), 4) English speakers that do not speak ‘ōlelo Hawai‘i that have lived in Hawai‘i for at least 1 year, and 5) English speakers that do not speak ‘ōlelo Hawai‘i who have not lived in Hawai‘i for at least 1 year. All qualifying information was gleaned from participants LBQ (see chapter 2, experiment 1 for more information on the LBQ, see Appendix A for the full questionnaire). Pidgin speakers from both Hawai‘i and the diaspora were recruited for two reasons: first, to disentangle the effects of language experience from regional exposure; and second, to acknowledge the growing population of Kānaka Maoli living in the diaspora who are also invested in (re)learning ‘ōlelo Hawai‘i.

Participants in the ‘ōlelo Hawai‘i speaking group were not required to speak Pidgin or English, although the majority identified themselves as trilingual speakers of Pidgin, English, and ‘ōlelo Hawai‘i. To qualify for this group, participants needed to indicate adequate experience in their language background questionnaire. Adequate experience included 10 years of home exposure of ‘ōlelo Hawai‘i, or the completion of at least 2 years of explicit ‘ōlelo Hawai‘i instruction. To recruit the most proficient speakers of ‘ōlelo Hawai‘i, a special subset of the ‘ōlelo Hawai‘i-speaking participants, referred to as “Hawaiian language consultants,” were invited to participate. To qualify as Hawaiian language consultants, participants were required to be either a Hawaiian language teacher or Hawaiian translator. Regardless of whether they were Hawaiian language consultants, all of these participants

5.2. Participants

were classified in the same ‘ōlelo Hawai‘i-speaking group.

Participants were excluded from the study if they had too little experience with Hawaiian to be categorized as ‘ōlelo Hawai‘i speakers (for example, those who had completed less than Hawaiian 102 or its equivalent), but too much knowledge to be placed in the group of participants who did not know ‘ōlelo Hawai‘i. For instance, participants who reported completing only Hawaiian 101 were excluded from the study, as they did not fit into any language group.

Speakers of Pidgin were recruited from Hawai‘i and diaspora Kānaka Maoli populations. Speakers were required to have had at least 10 years of exposure to Pidgin in the home and consider themselves a speaker of Pidgin. Pidgin speakers from Hawai‘i were recruited from Hawai‘i (the University of Hawai‘i Linguistics Beyond the Classroom program, social media, and physical flyers posted in public domains) and Pidgin speakers from diaspora were recruited through social media (on diaspora accounts such as “Kānaka living in Washington”, “Kānaka living in Vegas”, and “Frozen ‘Ohana” accounts). Participants in this category also were required to not consider themselves speakers of Hawaiian, and must not have had taken any formal Hawaiian language classes. Speakers of Pidgin from the diaspora were defined as individuals without regular exposure to Hawaiian through residence in Hawai‘i. These participants could report knowing “some Hawaiian words” but must not have lived in Hawai‘i for longer than one year or completed more than Hawaiian 102 (or its equivalent). Many individuals recruited from the diaspora did not meet the minimal requirements for the diaspora Pidgin group and were instead categorized as either ‘ōlelo Hawai‘i speakers (if they spoke too much Hawaiian) or Pidgin speakers from Hawai‘i (if they had lived in Hawai‘i for an extended period of time). Speakers of English were recruited from Hawai‘i (through the University of Hawai‘i Linguistics Beyond the Classroom program, social media, and physical flyers posted in public spaces) for the English from Hawai‘i group. Most participants in this category report to have been born and raised in Hawai‘i, although this was not a requirement.

Participants in the English from Hawai‘i and Pidgin groups frequently reported exposure to ‘ōlelo Hawai‘i, but critically did not self-identify as speakers of the language. This exposure primarily occurred through everyday contexts, including songs, greetings, public ceremonies, airport announcements, formal event openings, place names (e.g., streets, neighborhoods, schools, or businesses), and in media such as television, internet, and radio. Acceptable additional forms of ambient exposure to ‘ōlelo Hawai‘i included hula, mele, and other Hawaiian cultural classes. For instance, all students from

Hawai'i take *Hawaiiana* classes (Hawaiian cultural classes) from kindergarten through grade 6. In these classes, students typically are exposed to minimal vocabulary related to colors, animals, numbers, place names, names of significant Hawaiian figures and deities, and other foundational Hawaiian concepts, either explicitly or in the context of songs. In line with the body of literature on the proto-lexicon of te reo Māori among residents of Aotearoa, where school children have comparable classroom exposure (e.g., Oh et al. [2020]; Panther et al. [2023b]), this level of Hawaiian knowledge was accordingly considered acceptable for participants in non-ʻōlelo Hawai'i groups so long as they also do not consider themselves as Hawaiian speakers. Such exposure is best understood as ambient contact with the language rather than evidence of functional proficiency. While participation in either cultural or school-based activities such as hula or *Hawaiiana* classes may result in limited lexical knowledge of ʻōlelo Hawai'i, the amount of vocabulary acquired in these contexts is minimal and is therefore treated as part of general ambient exposure.

English speakers who were not from Hawai'i could not have lived in Hawai'i for more than one year. These participants were recruited through the University of British Columbia Linguistics Outside the Classroom program. Participants were not excluded if they had visited Hawai'i or lived there for less than one year. They may have had limited or incidental exposure to Hawaiian through popular media, a vacation to Hawai'i, or linguistics courses that discussed Hawaiian as a case study; however, they could not report having taken a Hawaiian language class, identified as Pidgin speakers, or resided in Hawai'i for over a year.

In total, there were 33 participants in the ʻōlelo Hawai'i category, 50 in the Pidgin from Hawai'i category, 7 in the Pidgin from diaspora category, 32 in the English from Hawai'i category, and 87 in the English without Hawaiian-exposure category. Due to concerns about whether 7 participants constituted an adequately representative sample, the Pidgin from diaspora group was excluded from analysis.

5.3 Experimental Procedure

5.3.1 Experimental Tasks

To test for proto-lexical knowledge of ʻōlelo Hawai'i, we conducted well formedness rating and word-nonword identification tasks following the designs of Oh et al. [2020] and Panther et al. [2023b].

5.3.2 Well formedness Rating Task

The well formedness rating task was the first task completed by all participants. This task examined participants' ability to judge the phonological well formedness of items explicitly identified as nonwords, providing a measure of their proto-lexical knowledge of the phonotactic regularities of 'ōlelo Hawai'i, that is, their sensitivity to the phonotactic regularities that govern the structure of real 'ōlelo Hawai'i words. Prior to beginning the task, participants were given written instructions and practice trials in English to ensure that they understood the procedure and the rating scale. Instructions on page one read, "On each trial, you will see a written nonword (a fake word, not attested in the dictionary) which was constructed to sound similar to a real word in Hawaiian. Your task is to rate how much each word sounds like Hawaiian to you, to the best of your abilities. To rate words, you will use your mouse to select buttons 1-5 on your screen: rate on a scale of 1-5 how Hawaiian-like you believe these words to be. 1: Not at all Hawaiian-like nonword, 2: Somewhat not Hawaiian-like nonword, 3: Neutral Hawaiian-like nonword, 4: Somewhat Hawaiian-like nonword, 5: Very Hawaiian-like nonword." To prevent participants from looking up words during the task, a 10-second response limit was imposed for each trial, as stated in the instructions. Two practice trials preceded the main task. Participants were told: "To rate words, you will use your mouse to select buttons 1–5 on your screen: rate on a scale of 1–5 how English-like you believe these words to be." The practice items were "dzik" and "bick." Following these trials, participants saw a message stating: "*dzik* was an example of a non-English-like nonword and *bick* is an example of a very English-like nonword."

After the practice trials, participants were instructed: "Now, in the same manner, you will be asked to determine how Hawaiian-like the following nonwords sound to you. You are not required to know any Hawaiian. We are interested in differences between groups in their answers. Try your best and trust your instincts. Press next to continue onto the experiment." The well formedness rating task then began.

Participants' responses to the English practice items were examined and were generally consistent with the expected pattern (i.e., *dzik* rated as less English-like and *bick* as more English-like). However, performance on these practice trials was not used to include or exclude participants and was not considered in subsequent analyses.

During the well formedness rating task, participants would see a fixation cross for 250 milliseconds, then see a written Hawaiian nonword with written instructions on each slide to rate each nonword from 1-5. Buttons 1-5 were

displayed on the screen. After a button press was given or after 10 seconds were up (whichever came first), a inter-trial interval (ITI) pause for 1000 milliseconds followed. Every 40 trials, participants were able to take a break if they wanted one, where they would land on a break screen they would need to press “next” to continue from. In total each participant had 212 trials of unique nonword items, drawn from a pool of 580 items. Trials were randomly ordered for participants. 212 items were balanced on 10 lists participants were randomly assigned to, matching the lists with comparable averages and range of phonotactic probabilities.

5.3.3 Word Identification Task

After the well formedness rating task, the word identification task followed. The Word Identification Task tests a different aspect of linguistic knowledge by asking a different question, “Your task is to rate how confident you are that each item you see is a Hawaiian word.” Stimuli consisted of written words, exactly half of words were real Hawaiian words and half were non-words. Participants from each language background group were compared to the Hawaiian language-speaking group in terms of their accuracy in judging whether items were real words, as well as the relationship between their confidence ratings and the calculated phonotactic probabilities of the stimuli.

At the start of the task, participants were instructed, “Using your mouse, selecting buttons 1-5 on your screen: rate on a scale of 1-5 how confident you are on whether a word is a word or not in Hawaiian. 1: Confident word is NOT a Hawaiian word, 2: Somewhat confident word is NOT a Hawaiian word, 3: Not at all confident this word is a Hawaiian word or not, 4: Somewhat confident word is a Hawaiian word, 5: Confident word is a Hawaiian word. Again, after instructions were given, participants next were given two practice trials in English. “To practice, let’s try two items in English. Rate on a scale of 1-5 how confident you are that each item you see is an English word. The first word was “car” and the second word was “mlar”. Participants were told after these English examples that “car” was an example of a real word in English and “mlar” was an example of a nonword in English, and that their confidence of these words being real or not will be recorded. Next instructions read, “Now let us move onto the task. Again, now you will be determining your confidence in whether each word is a real word or not in Hawaiian.” Instructions were repeated with which numerical values were associated with which responses. Instructions also included the maximum time of 10 seconds each trial, as in the well formedness rating task. “Please respond as quickly and as accurately as possible. (You will

have a maximum of 10 seconds per word to decide).” This generous response window, which is explicitly cited in previous written word recognition work (e.g., Balota et al. [2007a]), was selected to balance the instruction to respond as quickly and as accurately as possible with methodological considerations aimed at reducing time pressure effects and increasing participant comfort. This consideration is particularly relevant given the metalinguistic demands of identifying words that may be fuzzily represented and the inclusion of multilingual speakers. The 10 second limit served as a ceiling rather than an expected response time, and most responses occurred well within this window.

During the Word Identification Task, participants would see a fixation cross for 250 milliseconds, then see a written Hawaiian word or nonword with written instructions on each slide to rate each nonword from 1-5. Buttons 1-5 were displayed on the screen. After a button press was given or after 10 seconds were up (whichever came first), a inter-trial interval (ITI) pause for 1000 milliseconds followed. Every 40 trials, participants were able to take a break if they wanted one, where they would land on a break screen they would need to press “next” to continue from. In total all participants had 280 trials of unique items, 140 of which were nonwords and 140 of which were real words. All participants saw the same list of 280 words, in random order for each participant.

5.3.4 Language background questionnaire and language attitudes survey

At the end of each study, participants took part in a language background questionnaire and language attitudes survey. For more details on these tasks see Experiment 1. The same questionnaire and survey were included in Experiments 1-3.

5.4 Stimuli and Materials

5.4.1 Overall stimuli

Following Oh et al. [2020] and Panther et al. [2023b], participants viewed all stimuli in written form rather than hearing them auditorily. Like Māori, ‘ōlelo Hawai‘i has a highly transparent orthography, meaning that written forms closely reflect pronunciation. As reasoned in prior work (e.g. Oh et al. [2020], Panther et al. [2023b]) using written stimuli therefore provides

a consistent representation of word form and minimizes the variability that spoken Hawaiian speakers might typically encounter in natural speech.

Oh et al. [2020] reported that letters marked with a macron were perceived as more Māori-like, leading them to exclude long vowels from their stimuli. In the present study, long vowels were retained to minimize restrictions on the already limited Hawaiian stimulus set. Assuming that similar orthographic cues may influence perceptions of Hawaiian-likeness in ‘ōlelo Hawai‘i, however, we also controlled for the potential effects of orthographic symbols absent in English, the ‘okina and the kahakō, by proportionally balancing these features such that half of all stimuli contained at least one ‘ōlelo Hawai‘i-specific orthographic feature and the other half contained none.

5.4.2 Corpus

To construct the experimental materials, it was necessary to identify attested lexical items in ‘ōlelo Hawai‘i, generate phonotactically plausible nonwords, and compute phonotactic probability estimates for both item types. These requirements necessitated a probabilistic model of Hawaiian phonotactics trained on a large written corpus that reflects commonly occurring forms in contemporary written Hawaiian, corresponding to the varieties of Hawaiian most widely encountered in public and educational contexts in Hawai‘i.

‘O Wikipikia (Wikipedia contributors [2025]), a corpus compiled from Wikipedia articles translated into ‘ōlelo Hawai‘i, served as the training data for the language model and was used to identify real words, construct nonwords, and calculate phonotactic probabilities. This corpus was selected for several reasons. First, an ideal corpus for this study needed to include frequency information in order to control for the potential influence of word frequency on phonotactic ratings. This could not be obtained from a dictionary corpus, for example. Second, for convenience, a written corpus would be in a text-readable format, without requiring additional Optical Character Recognition (OCR) processing to convert PDF documents into machine-readable text. Third, the ideal corpus would be as large as possible and composed of language that most closely resembles the input a typical Hawaiian speaker would encounter. Thus, the Hawaiian Bible (Hawaiian Bible Project [2025]), for instance, was considered not ideal as the words that would be designated as highly frequent and the sounds that are often used (transliterations based on English) adopt somewhat atypical phonotactics (e.g., the use of traditionally “non-Hawaiian” phonemes such as /s/ in “sabeta” (“sabbeth”), which is a highly frequent word in the Hawaiian Bible corpus.

5.4.3 Stimuli Limitations

The ‘O Wikipikia corpus (Wikipedia contributors [2025]), as an open-source resource, was not entirely error-free and did not consistently adhere to standardized conventions. To minimize editing and maintain feasibility, the only modifications applied involved the removal of words containing non-standard Hawaiian characters. Specifically, all items including characters outside the standard Hawaiian sound inventory (m, n, p, t, k, h, w, l, ‘, a, e, i, o, u, ā, ē, ī, ō, ū) were excluded from the corpus used to train the language model. Despite these steps, some errors may have remained, which could have influenced the calculated phonotactic probabilities.

The standard⁷ Hawaiian orthographic conventions were used (i.e. inclusion of kahakō and ‘ōkina), as all participants reported exposure to this dialect of ‘ōlelo Hawai‘i. Some degree of dialectal variation was represented in the corpus as translators came from different regional backgrounds. It is notable to state, however, the /t/ sounds were very rare in this corpus based on phonotactic probabilities, consistent with the standard Hawaiian (/k/-variety) dialect.

Using standard orthography consistently with the stimuli was intentional, as it reflects the “standard” form of ‘ōlelo Hawai‘i most commonly encountered within public and educational contexts. However, variation across dialectal and colloquial uses that do not adhere to these conventions may have introduced variability in judgments of real words and nonwords. Furthermore, defining word “realness” based on whether the word is included in the dictionary and whether it was attested by a single language consultant may have offered a limited view of how participants would categorize words as real or nonword in ‘ōlelo Hawai‘i, given existing language variation. Although efforts were made to exclude nonwords that might be confusable with real words based on known patterns of variation, some forms reflecting variation beyond the scope of available knowledge may not have been captured.

⁷It is important to note that the term Standard Hawaiian is problematic, as it implies the existence of a single, prescriptive norm of ‘ōlelo Hawai‘i. In reality, variation exists across speakers, regions, and teaching contexts, and what is often referred to as Standard Hawaiian typically reflects the variety institutionalized through revitalization and formal education efforts. In this study, the term is used only to refer to the variety commonly described as Standard Hawaiian in both colloquial and linguistic contexts. This dialect was selected because the features represented are those most familiar to the participant population, based on their responses reported in their language background questionnaires. The term standard is not intended as a statement of linguistic accuracy or authenticity.

5.4.4 Calculating phonotactic probabilities

Phonotactic probabilities were calculated using KenLM (Heafield [2011]), a language model. KenLM is not provided phonological rules or constraints. Instead, it learns statistical patterns of phoneme combinations from real language data (i.e. the corpus). The output log-probabilities “phonotactic probability scores” approximate how statistically likely a sequence is, based purely on how often similar sequences appear in the training set.

KenLM was used in three stages: (1) during model training to capture the phonotactic patterns of ‘ōlelo Hawai‘i words, (2) to calculate phonotactic probabilities for real words, and (3) to score constructed nonwords. The specific procedures and KenLM implementation used for each stage are described below.

Preprocessing the corpus for the language model

To prepare the corpus for training and subsequent phonotactic probability calculation, the first step was to convert all words in the corpus into phoneme sequences. For a written corpus, as used in this study, all words will be transcribed into their pronunciation equivalents. While a phonemically transcribed corpus was not available for Hawaiian at the time this study was conducted, because the Hawaiian language is quite transparent, letters were converted to their phoneme equivalents using a basic substitution script in Python. Pronunciation equivalents were primarily drawn from the literature on the Hawaiian phonological system (Schütz [1994], Elbert and Pukui [1979], ‘Ōiwi Parker Jones [2006]). We converted the lexicon to ARPA-format to improve text-readability.

substitutions =

```
[('āe', 'AY:1-'), ('āi', 'AI:1-'), ('āo', 'AO:1-'), ('āu', 'AW:1-'),
 ('ēi', 'EY:1-'), ('ōu', 'OW:1-'), ('ai', 'AI1-'), ('ao', 'AO1-'),
 ('au', 'AW1-'), ('ae', 'AY1-'), ('ei', 'EY1-'), ('eu', 'EW1-'),
 ('oi', 'OY1-'), ('iu', 'IW1-'), ('ou', 'OW1-'), ('h', 'H-'),
 ('k', 'K-'), ('l', 'L-'), ('m', 'M-'), ('n', 'N-'), ('p', 'P-'),
 ('w', 'V-'), ('', 'GS-'), ('t', 'T-'), ('a', 'AA1-'), ('e', 'EH1-'),
 ('i', 'IY1-'), ('o', 'OO1-'), ('u', 'UW1-'), ('ā', 'AA:1-'),
 ('ē', 'EH:1-'), ('ī', 'IY:1-'), ('ō', 'OO:1-'), ('ū', 'UW:1-')]
```

Practical considerations, however, were also included as outlined below. Diphthongs were designated to their corresponding pronunciations, rather than a sum of their monophthong parts. This meant these vowels needed

to be converted to their phoneme equivalents first. Additionally, all long vowels were analyzed as individual vowel sounds instead of sequences of two vowels (e.g. where “AA” represented /a/, and “AA:” is /a:/, rather than “AA-AA” for /a:/). This decision was made based on previous work (e.g., Daland et al. [2011], Hayes and Wilson [2008], Jäger [2007]) which discuss treating long vowels (e.g., /a:/) as distinct phonemes, not sequences of repeated short vowels. However, when calculating phonotactic probabilities in both formats (e.g. “AA:” is /a:/ and “AA AA” for /a:/, calculations were extremely similar (differed only by a fraction of a phonotactic probability score). This convention also applied to long diphthongs.

Some variation was oversimplified. All “w” letters were converted to the “V” (/v/) phoneme as the /v/ $\sim$ /w/ variation with the “w” letter is not completely predictable by environment (e.g., Elbert and Pukui [1979], Schütz [1994]).⁸ All “t” letters were converted to the /t/ phoneme and “k” letters were converted to the /k/ phoneme although there is some spoken variation when it comes to this (free variation). Some spoken variation is thought to occur with some vowels, but not in a predictable manner.

Stress was also generalized in this pronunciation guide. Future iterations of this study could incorporate stress into the analysis; however, in the current analysis, all sounds were treated as stress-neutral, with each indicated by “1.” Hawaiian stress is generally predictable and falls on the penultimate mora (Pukui et al. [1975]), for words up to four syllables long (but not longer) (Schütz [1994]). Long vowels and diphthongs function as heavy syllables and are stressed, with primary stress assigned to the final heavy syllable and all other heavy syllables bearing secondary stress (Schütz [1994]).

All pronunciation sequences were originally separated by spaces rather than dashes, so these separators were replaced with dashes to prepare the corpus for modeling.

Training the language model

A tri-gram (n=3) model was trained using KenLM to capture the sequential phonotactic structure of ‘ōlelo Hawai‘i. The input to the model consisted of phoneme sequences derived from the prepared Hawaiian corpus, where each word was represented as a sequence of phoneme separated by spaces. During training, KenLM estimated the conditional probabilities of phoneme

⁸Regarding the predictability of /v/ $\sim$ /w/ variation, some accounts propose that /w/ occurs before /i/ or /e/, while /v/ occurs before /a/, /o/, or /u/. However, the presence of numerous exceptions has led other researchers to argue that this variation cannot be straightforwardly analyzed as allophonic (Parker Jones [2018]).

sequences, such that the probability of a given phoneme was modeled as dependent on the two preceding phonemes. This allowed the model to estimate the relative likelihood of specific phoneme sequences based on their frequency of occurrence and to capture the phonotactic constraints that define well-formed sound combinations in ‘ōlelo Hawai‘i.

Although KenLM can incorporate frequency information to weight the training process, the present study focused on structural phonotactic probabilities rather than lexical frequency effects. To this end, all word frequencies in the corpus were normalized to a numerically neutral value of 1. This approach follows Oh et al. [2020], who found that type-based frequency representations more effectively capture probabilistic patterns of form than token-based frequency counts. By doing so, the model’s estimates reflect the likelihood of phoneme co-occurrence based on form alone, rather than being influenced by lexical frequency distributions.

Calculation of Phonotactic Probability

Once the language model was trained, real words from the ‘O Wikipikia corpus were processed through the model to obtain phonotactic probability estimates. KenLM implements a statistical n -gram language model in which the probability of a sequence is computed as the product of conditional probabilities of each unit given the preceding $n - 1$ units (Heafield [2011]).

For the given trigram model ($n = 3$), this corresponds to estimating the probability of each phoneme given the two preceding phonemes, with word boundaries represented by start and end of sequence symbols. The log-probability of a word is calculated as the sum of the log conditional probabilities across the sequence (Heafield [2011]). For example with the ‘ōlelo Hawai‘i word “pake”,

$$\log P(\text{PAKE}) = \log P(\text{P}) + \log P(\text{AK} \mid \text{P}) + \log P(\text{E} \mid \text{PAK})$$

When KenLM is queried with a list of phoneme sequences, a log-probability is returned per word based on how frequently those n -grams (phonemes) appear in training data. The returned score is a logarithm (base 10) of the total probability of the sequence under the n -gram model.

KenLM was trained using modified Kneser–Ney smoothing, the default and recommended smoothing method in KenLM, to handle rare or unseen sequences (Heafield [2011]). This ensures all sequences get a non-zero probability.

Next, normalization is applied internally within KenLM as part of modified Kneser–Ney smoothing, ensuring that conditional probabilities are prop-

erly normalized across contexts (Heafield [2011]). The application of sequence-normalization helps control for the fact that longer sequences naturally have lower (more negative) log-probabilities, the final phonotactic score yielding a negative number. A higher log-probability (i.e., less negative) means the sequence occurs more frequently in the training data, is more “phonotactically probable”.

5.4.5 Construction and Selection of nonword Stimuli

Following the procedures outlined in Panther et al. [2023b], I constructed nonwords by first identifying real ‘ōlelo Hawai‘i words from the ‘O Wikipikia corpus that consisted of two to three syllables. These words were then modified by altering one to four phonemes. Any items that resembled English words or that exhibited reduplication or other complex morphological patterns were excluded from consideration.

To generate nonwords systematically matched to existing real words, I wrote a Python script that produced phonotactically matched nonwords through controlled character substitutions. All consonants could be replaced with any other consonant, and all vowels could be replaced with any other vowel, with the only restriction that long vowels were replaced only by other long vowels (e.g., “ā” with “ē,” “ī,” “ō,” or “ū”) and short vowels only by short vowels (e.g., “a” with “e,” “i,” “o,” or “u”). This ensured that consonant–vowel positions were preserved while allowing broad variability in the resulting nonwords. Long vowels were treated specially because words containing macrons were found to be associated with Māori-likeness ratings in Oh et al. [2020]. A list of real input words from the Hawaiian corpus were imported. For each input word, the script randomly selected between one and four positions to substitute. For every position available, one to four character substitutions were selected at random and used to replace the original one to four characters. This process was repeated until ten unique nonwords were generated per original input word. All generated nonwords were then calculated for their phonotactic probabilities.

This process occasionally produced real words. Therefore, each constructed nonword was checked individually to verify that it was not a real word. A nonword was defined as an item that (1) was not identified as a real word by a Hawaiian language consultant who assisted with stimulus creation, and (2) did not appear in the Ulukau Hawaiian language dictionary database (Kuamo‘o [2009]).

Next, all nonwords were processed through the language model using the same calculation procedure described in the previous section. Prior to scor-

ing, the nonwords were preprocessed and then input into the KenLM model to calculate their phonotactic probabilities. Crucially, the phonotactic probabilities of the nonwords were calculated using the language model trained on real ‘ōlelo Hawai‘i words.

To maintain as much phonotactic integrity as possible in our list of nonwords, a Python script was used to systematically apply the following matching criteria between real words and their corresponding nonwords: (i) The word and nonword had the same number of syllables; (ii) The word and nonword had the same number of phonemes; (iii) The word and nonword had consonants and vowels in the same positions; (iv) The word and nonword had similar phonotactic scores (defined here as within ± 2 points)⁹, following the general matching procedures described in Panther et al. (2023), who paired words and nonwords with comparable phonotactic probabilities. The number of phonemes ranges from three to six phonemes.

Mid-frequency real words only were selected as real words to include as stimuli, and as templates to create nonwords. This decision was made to control for the potential influence of frequency on ratings of Hawaiian-likeness, as word frequency is another factor known to affect perceived word-likeness (Bailey and Hahn [2001]).

For the well formedness rating task, only nonwords were required. An approximately normal distribution of words across a range of phonotactic probabilities were selected. To draw from the same pool of stimuli for task 1, 10 lists were made, drawing 212 words from a pool of 580 items. 212 items were balanced on 10 lists participants were randomly assigned to, matching the lists with comparable averages and range of phonotactic probabilities. This semi-random selection with relatively equal matching for features in terms of phonotactics was conducted using a Python script.

For the word-nonword identification task, an equal count of words (140) and nonwords (140) was included. Each selected word and its matching nonword were included in the stimulus set. This ensured that the words and nonwords were similar across multiple factors, including syllable length, phonotactic probability, number of phonemes, and consonant–vowel positioning. The phonotactic probabilities of the full set of stimulus words were approximately normally distributed. Words and nonwords exhibited similar averages and variability in their phonotactic probability values. The complete wordlist, including phonotactic scores for both nonwords and their corresponding real word matches, is provided in Appendix D.

⁹We selected ± 2 as our matching tolerance based on the distribution of phonotactic scores in our corpus; the ± 2 threshold was not explicitly defined in Panther et al. [2023b]

5.5 Data inclusion criteria and participant exclusion

A total of 215 participants completed the study. Eight participants in the ‘ōlelo Hawai‘i group were excluded from statistical analyses because they did not qualify for any participant group; they did not meet the Hawaiian 102 threshold but had completed too much Hawaiian coursework (Hawaiian 101) to be placed in another category. Their data were retained for future analysis. An additional five participants were excluded after being identified as likely bots or as not taking the study seriously, as indicated by nonsensical responses in their language background questionnaires (e.g., listing nonexistent high schools or entering random letters or symbols in response fields). The seven participants in the diaspora group were also excluded from analysis because the group size was too small to provide a representative sample. Their data were retained for future analysis. In total, after all exclusions, 195 participants remained for analysis. 33 were Hawaiian speakers, 49 were Pidgin speakers, 31 were English speaking from Hawai‘i and 82 were English speaking from Canada.

5.6 Well Formedness Task: Statistical Analyses and Results

For the Well Formedness Task, we fit a Bayesian cumulative probit regression model using the *brms* package in R (Bürkner [2017]) to predict ordinal participant responses (1-5) on how “Hawaiian-like” each nonword was. The model included phonotactic probabilities and language background as fixed effects, along with their interaction. Phonotactic probability was modeled as a continuous predictor, operationalized as the non-frequency weighted phonotactic probability computed for each nonword item. Values were centered and scaled (z-scored) using the *scale()* function in base R (R Core Team [2021]), producing predictors with mean 0 and standard deviation 1 prior to inclusion in the model. This standardization improves interpretability and model convergence. Language background was encoded as a categorical variable with four levels, Hawaiian speakers (the reference level), Pidgin speakers, English speakers from Hawai‘i, and English speakers from Canada. The model included random intercepts by item and by participant to account for item-level and participant-level variability in responses. In this analysis, the reference level for the categorical predictor Language Background was Hawaiian speakers.

The model was specified as: $\text{Response} \sim 1 + \text{phonotactic-probability} * \text{language-background} + (1 \mid \text{item}) + (1 + \text{phonotactic-probability} \mid \text{participant})$.

The model was fit with four Markov Chain Monte Carlo (MCMC) chains, each run for 4,000 iterations with a warmup of 1,000 and thinning was set to 1 ($\text{thin} = 1$), resulting in 12,000 post-warmup draws in total. All parameters showed excellent convergence ($\text{Rhat} = 1.00$) and high effective sample sizes for both Bulk ESS and Tail ESS, well above commonly recommended minimum thresholds ($\approx 1,000$). Accordingly, no additional thinning was required, and the model was retained with $\text{thinning} = 1$.

Priors on the fixed effects and on the intercepts were specified as weakly informative normal distributions with $(\mu = 0, \sigma = 5)$. These priors regularize estimates toward zero while still allowing the model to detect effects of meaningful magnitude, and are appropriate for ordinal response models on the probit scale (e.g., McElreath [2020]). The outcome variable was modeled using a cumulative family with a probit link function for the threshold model and an identity link for the discrimination parameter (disc), which was fixed at 1.00.

The primary effect of interest for the current research question is the interaction between phonotactic probability and language background. An interaction indicates that the relationship between participant response and phonotactic probability differs across language groups. With the reference level set to Hawaiian speakers, a 95% credible interval (CI) that excludes zero suggests that this relationship differs meaningfully from that of the ‘ōlelo Hawai‘i speaker group. Positive and negative estimate directions describe the relationship between participant response and phonotactic probability relative to the ‘ōlelo Hawai‘i speakers, who serve as the baseline (0). Estimates whose absolute values are closer to zero indicate behavior that more closely matches that of the baseline group. Positive estimates indicate a stronger correspondence between participant responses and corpus-derived phonotactic probabilities, whereas negative estimates indicate a weaker correspondence compared to the ‘ōlelo Hawai‘i speakers. Effects with 95% CIs excluding zero are interpreted as meaningfully different from baseline, while intervals including zero suggest uncertainty or limited evidence of meaningful difference (e.g., Bürkner [2017]; McElreath [2020]). Threshold parameters (Intercepts 1 through 4) mark the cumulative cutpoints that separate the five ordered response categories on the probit scale.

According to our hypothesis, all non-Hawaiian speaking groups will have responses that deviate further away from the corpus-derived phonotactic probabilities (negative estimates) relative to Hawaiian speakers. However,

consistent with Oh et al. [2020], participants with more substantial exposure to Hawaiian, including Pidgin and English speakers from Hawai‘i, may pattern more similarly to Hawaiian speakers. Specifically, we predict a small negative interaction estimate for Pidgin speakers from Hawai‘i; a moderately negative estimate for English speakers from Hawai‘i; and the largest negative estimate for English speakers from Canada.

A summary of the model output is presented in Table 5.1. There was a positive effect of phonotactic probability (Estimate = 0.14, 95% CI = [0.04, 0.24]) indicating that ‘ōlelo Hawai‘i listeners’ ratings of word-likeness increased as phonotactic probability increased. This relationship is illustrated in Figure 5.1. Responses from all other language groups differed from Hawaiian speakers both as simple effects and in interactions with phonotactic probability.

The simple effect of phonotactic probability reflects the baseline relationship between response and phonotactic probability for Hawaiian speakers, who served as the reference group in the model. We expect this effect to be positive, indicating that words with higher phonotactic probability elicit higher participant response values, with a 95% CI excluding zero to suggest that phonotactic knowledge is credibly reflected in the Hawaiian speakers’ responses. The baseline intercepts represent the overall response distribution for the Well Formedness task, providing the reference point against which all other language groups are compared. The simple effects for each language group describe how their overall response tendencies differ from this baseline. These simple effects thus estimate each group’s general response bias or tendency, independent of how phonotactic probability shapes their responses as indicated by the interaction terms.

5.6. Well Formedness Task: Statistical Analyses and Results

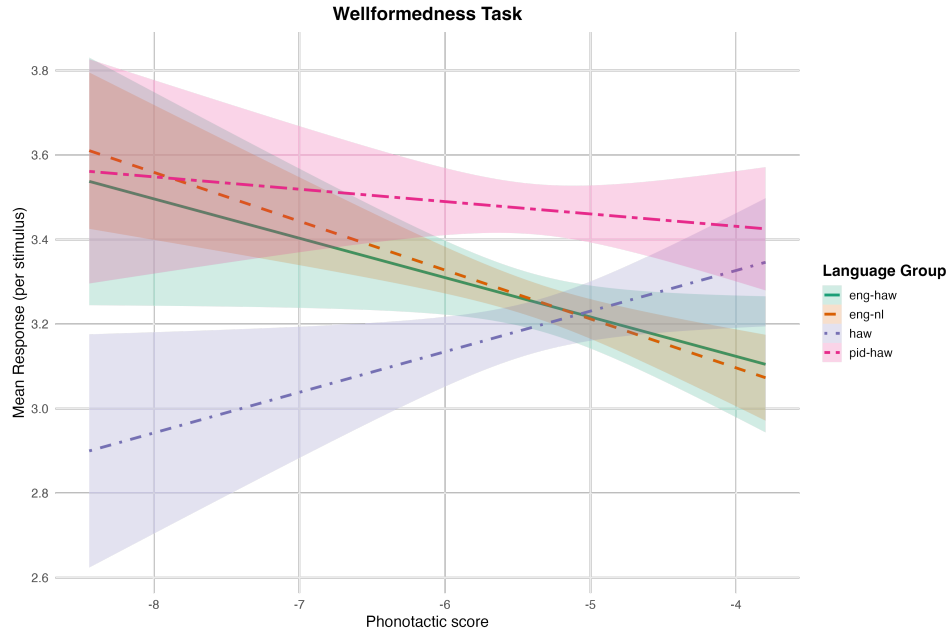


Figure 5.1: Well formedness ratings as a function of phonotactic probability and language background. The figure displays empirical participant responses, with phonotactic probability on the x-axis and well formedness ratings on the y-axis, shown for each stimulus and stratified by language group (green = English participants from Hawai‘i, orange = English speakers from Canada, purple = Hawaiian speakers, pink = Pidgin speakers from Hawai‘i). Lines show group-specific linear fits to the observed responses, with shaded regions indicating standard error. Higher phonotactic probability values correspond to items that are more phonotactically attested in the corpus, and higher response values reflect greater perceived well formedness. Phonotactic scores appear as negative values because they are expressed on a log probability scale: more negative values indicate lower phonotactic probabilities, whereas values closer to zero indicate higher phonotactic probabilities.

5.6. Well Formedness Task: Statistical Analyses and Results

Term	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept[1]	-2.22	0.28	-2.75, -1.67	1.00	2129	4382
Intercept[2]	-1.29	0.28	-1.83, -0.75	1.00	2130	4406
Intercept[3]	-0.61	0.28	-1.15, -0.07	1.00	2133	4490
Intercept[4]	0.32	0.28	-0.22, 0.87	1.00	2133	4438
Phonotactic Probability	0.14	0.05	0.04, 0.24	1.00	1978	4347
Pidgin speakers from Hawai'i	-0.60	0.29	-1.18, -0.04	1.00	3366	5903
English speakers from Hawai'i	-1.28	0.32	-1.91, -0.66	1.00	3376	5705
English speakers from Canada	-1.38	0.26	-1.91, -0.86	1.00	3085	5242
Phonotactic Probability: Pidgin speakers from Hawai'i	-0.17	0.05	-0.27, -0.07	1.00	2839	5347
Phonotactic Probability: English speakers from Hawai'i	-0.26	0.05	-0.36, -0.15	1.00	2881	5615
Phonotactic Probability: English speakers from Canada	-0.27	0.04	-0.36, -0.18	1.00	2644	4836

Table 5.1: Fixed effects from a Bayesian-mixed effects cumulative probit model predicting participants' ordinal well formedness ratings of Hawaiian-like nonwords. The model estimates how phonotactic probability and language background influence these judgments, as well as their interactions, with Hawaiian speakers serve as the reference group for all comparisons. Estimates with lower absolute magnitudes reflect response behavior that is less differentiated from the Hawaiian speaker baseline, indicating greater similarity in well formedness judgments. Positive estimates indicate higher well formedness ratings, while negative estimates indicate lower ratings relative to the Hawaiian speaker baseline. 95% CIs excluding zero are interpreted as meaningfully different from baseline, while intervals including zero suggest uncertainty or limited evidence of meaningful difference.

Model results (see Table 5.1) indicated a positive simple effect of phonotactic probability for the reference Hawaiian speaker group, such that nonwords with higher phonotactic probability scores received meaningfully higher well formedness ratings overall (Estimate = 0.14, 95% CI = [0.04, 0.24]; see Table 5.1), with the 95% CI excluded zero. This suggests that Hawaiian speakers' responses were more strongly influenced by the underlying phonotactic structure of the stimuli compared to any other group, as their judgments associated more positively with the corpus-derived phonotactic probability measures and the other groups showed comparatively more negative associations.

Comparing the interaction effects between phonotactic probability, language background, and language group, all non-Hawaiian speaking groups show negative interaction estimates that were meaningfully different from the Hawaiian speakers (95% CI excluded zero). This indicates that the link between phonotactic probability and well formedness ratings was not only weaker for these groups compared to the Hawaiian speaker baseline, it was notably in the opposite direction. Pidgin speakers from Hawai'i show a smaller negative effect (Estimate = -0.17, 95% CI = [-0.27, -0.07]), English speakers from Hawai'i show a moderately larger effect (Estimate =

-0.26, 95% CI = [-0.36, -0.15]), and English speakers from Canada show the largest deviation from the Hawaiian baseline (Estimate = -0.27, 95% CI = [-0.36, -0.18]). The 95% credible intervals for all three groups do not include zero, thus we interpret the relationship between participant response and phonotactic probabilities as meaningfully different for each language group compared to that of Hawaiian speakers. While all non-Hawaiian speaker groups demonstrate a negative relationship between well formedness ratings and phonotactic probability, the gradient across groups followed the expected order. Pidgin speakers showed the smallest attenuation of the phonotactic probability effect on participant responses relative to the Hawaiian baseline, English speakers from Hawai'i showed a larger attenuation, and English speakers from Canada showed the greatest attenuation.

To further investigate why the relationship between well formedness ratings and phonotactic probability was negative for all non-Hawaiian speakers, we considered the possibility that uniquely Hawaiian orthographic features, namely the 'okina and kahakō, might bias some participants to rate these items as "more Hawaiian-like." Previous literature (e.g., Oh et al. [2020]) excluded uniquely Māori orthographic markers, as it was found that these items were disproportionately rated as "more Māori-like," independent of whether the phonotactic structure was indeed more probable. After excluding responses associated with items containing 'okina and kahakō, we re-ran the same model to determine whether these orthographic features accounted for the observed negative relationship. The same model specification and fitting procedure described above were used for this follow-up analysis. Convergence diagnostics again indicated excellent model performance (Rhat = 1.00), with high Bulk ESS and Tail ESS values and no need for additional thinning.

A summary of the follow-up analysis is presented in Table 5.2. There remained a positive effect of phonotactic probability (Estimate = 0.18, 95% CI [0.08, 0.29]) indicating that among items that did not contain 'okina and kahakō 'ōlelo Hawai'i speakers still displayed a positive relationship between well formedness ratings and phonotactic probability (Estimate = 0.18, 95% CI [0.08, 0.29]). Pidgin speakers indicated a positive relationship between their word-likeness ratings and phonotactic probability (Estimate = -0.13, 95% CI [-0.25, -0.02]); combined with the Hawaiian speaker reference-level baseline of 0.18, this yields an overall slope of 0.05 ($0.18 - 0.13 = 0.05$), a reliably positive relationship. This indicated that after excluding uniquely Hawaiian orthographic characters from the stimuli, Pidgin speakers' ratings of word-likeness increased as phonotactic probability increased, similarly to Hawaiian speakers. This relationship is illustrated in Figure 5.2, where

Hawaiian speakers and Pidgin speakers are positive aligned, such that when phonotactic probability of the word items increased, their well formedness ratings increased, accordingly.

For English speakers from Hawai'i, the effect of phonotactic probability on well formedness ratings is negative (Estimate = -0.28, 95% CI [-0.41, -0.16]). The interaction estimate of -0.28 reduces the slope from the reference level of 0.18 to -0.10 ($0.18 - 0.28 = -0.10$), with the 95% CI not including zero, indicating a confidently negative effect. English speakers from Canada show the largest deviation from Hawaiian speakers, with an interaction of -0.29, resulting in an overall slope of -0.11 (Estimate = -0.29, 95% CI [-0.39, -0.19]), also confidently negative. This relationship is illustrated in Figure 5.2, where English speakers from Hawai'i and English speakers from Canada are negative in slope, such that when phonotactic probability of the word items increased, their well formedness ratings decreased.

5.6. Well Formedness Task: Statistical Analyses and Results

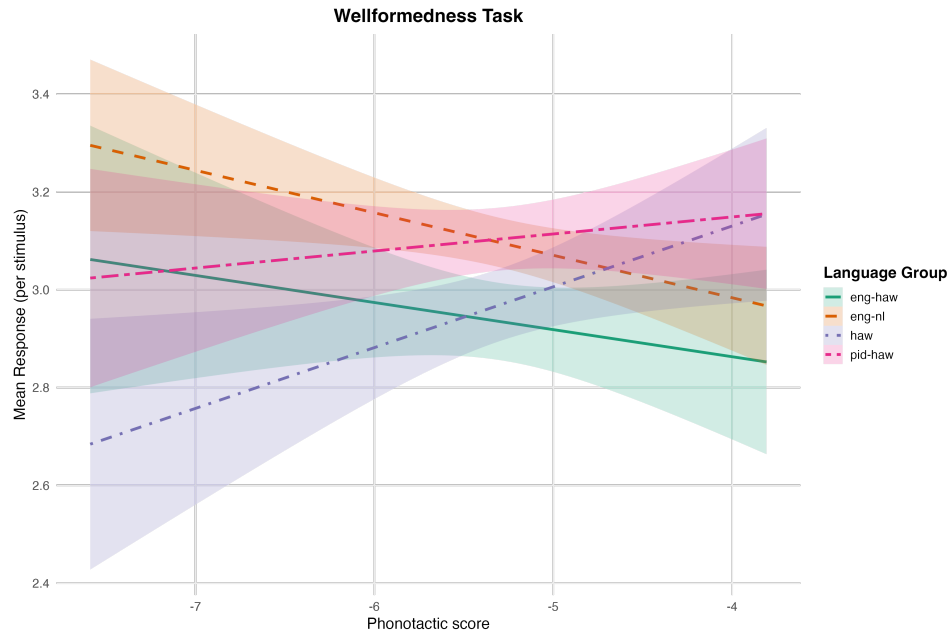


Figure 5.2: Well formedness ratings for each stimulus as a function of phonotactic probability and language background, shown for items without uniquely Hawaiian orthography (kahakō or ‘okina). The figure displays empirical participant responses, with phonotactic probability on the x-axis and well formedness ratings on the y-axis, shown for each stimulus and stratified by language group (green = English participants from Hawai‘i, orange = English speakers from Canada, purple = Hawaiian speakers, pink = Pidgin speakers from Hawai‘i).

5.7. Well formedness Task: Discussion

Term	Estimate	Est. Error	95% CI	Rhat	Bulk ESS	Tail ESS
Intercept[1]	-2.28	0.32	-2.92, -1.67	1.00	820	1551
Intercept[2]	-1.27	0.32	-1.91, -0.66	1.00	821	1588
Intercept[3]	-0.54	0.32	-1.18, 0.07	1.00	819	1570
Intercept[4]	0.39	0.32	-0.24, 1.01	1.00	818	1569
Phonotactic Probability	0.18	0.06	0.08, 0.29	1.00	736	1450
Pidgin speakers from Hawai'i	-0.55	0.34	-1.22, 0.12	1.00	1081	1609
English speakers from Hawai'i	-1.48	0.38	-2.25, -0.73	1.00	966	1622
English speakers from Canada	-1.39	0.30	-1.99, -0.81	1.00	881	1554
Phonotactic Probability: Pidgin speakers from Hawai'i	-0.13	0.06	-0.25, -0.02	1.00	903	1350
Phonotactic Probability: English speakers from Hawai'i	-0.28	0.06	-0.41, -0.16	1.00	840	1302
Phonotactic Probability: English speakers from Canada	-0.29	0.05	-0.39, -0.19	1.01	801	1358

Table 5.2: Population-level effects from the Bayesian cumulative probit model predicting participants' ordinal well formedness ratings of Hawaiian-like nonwords excluding responses from stimuli with uniquely Hawaiian orthographic features. Estimates show the effect of phonotactic probability and language group, as well as their interactions. Positive estimates indicate higher ratings relative to the reference group (Hawaiian speakers), while negative estimates indicate lower ratings. 95% CIs that exclude zero indicate a reliable difference from baseline.

5.7 Well formedness Task: Discussion

The results of the Well formedness Task provide clear evidence that sensitivity to Hawaiian phonotactic structure varies systematically across language backgrounds. Hawaiian speakers were the only group that showed a meaningful alignment between phonotactic probability and well formedness ratings, across all stimuli types that are possible in 'ōlelo Hawai'i, which suggests that their judgments in relation to the nonword stimuli reflect phonotactic knowledge grounded in actual patterns of Hawaiian. Hawaiian speakers showing the closest relationship between phonotactic probability and well formedness validates the phonotactic probability measures used in this study, indicating that these measures indeed capture meaningful distributional information present in the mental representations of Hawaiian speakers. Even though the majority of Hawaiian speakers in the sample did not report Hawaiian as their L1, they nonetheless demonstrated a clear alignment between phonotactic probability and well formedness ratings, consistent with previous work demonstrating speakers' sensitivity to phonotactic probabilities (e.g., Treiman et al. [1996], Vitevitch et al. [1997]).

In contrast, all non-Hawaiian listener groups exhibited a negative relationship between phonotactic probability and well formedness ratings when Hawaiian-unique orthography is included in the stimuli. This negative slope pattern was shared across all non-Hawaiian listener groups and indicates

that, despite differences in degree, none of these groups employed Hawaiian phonotactic probabilities in the same way as Hawaiian speakers across the full Hawaiian phonotactic inventory (see Figure 5.1).

After the exclusion of responses from stimuli items containing either ‘okina or kahakō, however, Pidgin speakers displayed a positive relationship between phonotactic probability and well formedness ratings. This suggests that stimuli containing Hawaiian-unique characters previously biased well formedness ratings, such that items with low phonotactic probability were rated as more well-formed than those with high phonotactic probability. Excluding these items, Pidgin speakers demonstrate evidence of phonotactic awareness that is not present in the case of the two English speaking groups.

Similar orthographic biases were reported in pilot studies for English speakers from Aotearoa (Oh et al. [2020]), which led the authors to exclude uniquely Māori orthographic characters in their design. The follow-up analysis in the current study provides evidence that Pidgin speakers possess knowledge of ‘ōlelo Hawai‘i phonotactics, as their well formedness ratings align positively with phonotactic probability when stimuli exclude uniquely Hawaiian orthographic characters; however, when such orthographic characters are included, their ratings show a negative bias, indicating a limitation in the robustness of their phonotactic knowledge which was not present in Hawaiian speakers.

In our main design, we incorporated all possible sounds, including uniquely Hawaiian orthographic characters, to investigate proto-lexical behavior across the complete Hawaiian phonotactic system. This decision was crucial for maintaining ecological validity, as excluding these characters would have artificially constrained the phonotactic space and limited our ability to capture how speakers across the language groups evaluate well formedness in a system that fully reflects ‘ōlelo Hawai‘i. This was important because proto-lexical judgments are thought to emerge from exposure to the full distributional structure of a language. Additionally, including kahakō and ‘okina allowed us to test whether listener’s judgments were driven by phonotactic structure itself or influenced by familiarity with Hawaiian-specific orthographic cues. Conducting the analysis in two parts, by first including, then next, excluding Hawaiian-specific orthography, allowed us to tease apart the differences in the respective phonotactic awareness between Hawaiian speakers and Pidgin speakers. Even when uniquely Hawaiian orthographic characters were included, Hawaiian speakers maintained a positive relationship between phonotactic probability and well formedness ratings, indicating that their judgments were not influenced by these more salient orthographic cues. In contrast, Pidgin speakers were influenced by these characters, judging them

as more “Hawaiian-like” although this was in fact in the opposite direction of phonotactic probability scores. However, with the exclusion of these characters, Pidgin speakers’ well formedness judgments aligned with phonotactic probability. This demonstrated that, aside from their understanding of items which including ‘okina and kahakō, Pidgin speakers possess sensitivity to phonotactic probability which aligned them closest to Hawaiian speakers, and also distinguished them from other non-English speakers.

Furthermore, this follow-up analysis demonstrated that both English-speaking groups remained negatively associated between phonotactic probability and well formedness ratings after filtering out “Hawaiian-like” orthographic characters. This indicates that the negative relationship between phonotactic probability and well formedness judgments was independent of the presence of uniquely Hawaiian orthographic cues, for both English speakers from Hawai‘i as well as English speakers from Canada. In other words, even when ‘okina and kahakō were absent in the stimuli, English-speaking participants rated phonotactically less probable items as more Hawaiian-like. One possible explanation for this negative relationship between well formedness ratings and phonotactic probability is that participants evaluated items based on how typologically “unlike” other known languages, such as English, they appeared, effectively “exoticizing” the perceived structure of the stimuli in regard to its Hawaiian-likeness. This pattern suggests that English speaking listeners rely more heavily on perceived typological distinctiveness than on phonotactic probability when evaluating well formedness. This behavior indicates weaker proto-lexical knowledge of Hawaiian phonotactic patterns for the two English-speaking groups compared to Pidgin speakers.

This further teases apart the differences between Pidgin speakers and English speakers from Hawai‘i. Inclusive of the full Hawaiian inventory, Pidgin speakers showed the smallest reduction in phonotactic sensitivity relative to Hawaiian speakers. When uniquely Hawaiian orthographic characters were excluded, Pidgin speakers were the only non-Hawaiian speaking group whose ratings remained positively aligned with phonotactic probability, whereas English speakers from Hawai‘i continued to show negative alignment. This pattern is consistent with the hypothesis that Pidgin speakers would exhibit an advantage over other English speakers, reflecting a proto-lexicon that deviates least from the Hawaiian speaker baseline and is generally positively aligned, albeit influenced by Hawaiian-specific orthography.

In this task, English speakers from Canada showed no evidence of sensitivity to Hawaiian phonotactic patterns, as their well formedness ratings did not align with phonotactic probability, with low probability items rated as more well-formed than high probability items, regardless of the inclusion

of Hawaiian-specific orthography. This pattern suggests that in the absence of meaningful exposure to Hawaiian, participants may draw on strategies unrelated to Hawaiian specific structure. Their responses provide weak evidence of attunement to Hawaiian phonotactic probability, which aligns with previous research showing that individuals without exposure do not develop sensitivity to the statistical structure of a language (e.g., Oh et al. [2020], Panther et al. [2023b]).

The simple effects also provide insight into participants' overall response strategies. All non-Hawaiian speaking groups gave lower well formedness ratings on average, and the gradient of similarity to Hawaiian speakers followed the expected pattern with Pidgin speakers rating all words higher on average compared to English speakers from Hawai'i, and English speakers from Hawai'i rating words higher on average compared to English speakers from Canada. In other words, the task required some familiarity with what Hawaiian words look and sound like, and this broader familiarity appears to vary across groups. This indicates that the stimuli were perceived as less Hawaiian-like by participants who lack fluent knowledge of Hawaiian, even before phonotactic probability is considered. This result is consistent with expectations given the nature of the stimuli, which did not contain any illegal structures or sounds absent in the Hawaiian inventory. These baseline effects suggest that well formedness ratings reflect not only sensitivity to phonotactic probabilities but also broader knowledge of the Hawaiian phonological system and its constraints. Similar baseline effects have been reported in previous work on Māori, where Māori speakers had higher ratings on average compared to non-Māori speaking groups (e.g., Oh et al. [2020], Panther et al. [2023b]).

With respect to the high baseline observed for Hawaiian speakers, participant feedback indicated that some items classified as “nonwords” were identified as possible real words by certain Hawaiian-speaking participants, while others judged the same items to be nonwords. This variability likely introduced additional noise into the word identification task and may have contributed to elevated baseline ratings in the Well Formedness task for this group.

Overall, these reductions indicate a gradient in phonotactic sensitivity across the language groups, consistent with the predictions of our hypothesis. Hawaiian speakers were the only group that were positive aligned regardless of inclusion of 'okina and kahakō, meaning that they did not automatically judge these orthographic features as inherently more phonotactically probable. Pidgin speakers appeared to be biased in rating items with 'okina and kahakō as more “Hawaiian-like”, however independent of these items Pid-

gin speakers displayed positive alignment in their well formedness ratings and phonotactic probability scores. English speakers were both negative in slope both between well formedness ratings and phonotactic probability, regardless of the inclusion of ‘okina and kahakō. Although reductions in phonotactic sensitivity differed meaningfully across the two English speaking groups, these differences should not be interpreted as weaker but aligned use of Hawaiian phonotactic knowledge. Instead, they reflect varying magnitudes of deviation from the Hawaiian baseline.

To situate the reduced Hawaiian effects by English speakers in a broader context, Todd et al. [2023a], show that non-Spanish speaking Californians and Texans likewise exhibit weaker effects of Spanish proto-lexical knowledge compared to the much stronger Māori proto-lexicon effects reported in Oh et al. [2020] and Panther et al. [2023a]. Although the authors examined individual differences in Spanish exposure, they found no statistically reliable effects and note that the reduced participant pool after screening limited statistical power, making the null results inconclusive. Their study also lacked fluent Spanish-speaking controls and English-speaking controls without Spanish exposure. The authors suggest that differences in the strength of the effects may arise because Spanish is structurally more complex and includes a greater number of forms and structures to incorporate into the proto-lexicon compared to Māori. While the authors do not suggest that their weaker effects stem from differences in Māori and Spanish exposure contexts, the current Hawaiian study, which involves a language structurally similar to Māori, demonstrates relatively weaker effects among non-Hawaiian speakers similar to Todd et al. [2023a] than those observed for non-Māori speakers in Aotearoa (Oh et al. [2020] and Panther et al. [2023a]). This pattern suggests that gradient exposure differences may better explain cross-study differences than structural ones.

Overall, the observed relationship between well formedness judgments and phonotactic probabilities supports the predicted gradient: Pidgin speaker ratings aligned the strongest with Hawaiian speakers, followed by English speakers from Hawai‘i, and finally English speakers from Canada showing the weakest alignment. This gradient pattern underscores the continuous relationship between relative degrees of exposure to a language and the development of statistical language knowledge.

5.8 Word Identification Task: Statistical Analyses and Results

5.8.1 Bayesian Model and Model Interpretation

For the Word Identification Task, we fit a Bayesian cumulative probit regression model using the `brms` package in R (Bürkner [2017]). The outcome variable was participants' ordinal response (1-5, how "confident" participants were the word was real or not), and predictors included the three-way interaction between: phonotactic probabilities (scaled with z-scoring, calculated for both real and nonword items), item type (real vs. nonword), and language background. The model included random intercepts for items and random intercepts and slopes for phonotactic probability by participant to account for individual variability in responses. In this model, the Hawaiian speaker group served as the reference level for language background, and real-word items served as the reference for item type. Thus, all simple effects and interactions are interpreted relative to the responses of Hawaiian speakers rating real words. The following formula was fit for this analysis:

Formula: $\text{Response} \sim 1 + \text{phonotactic-probability} * \text{language-background} * \text{type} + (1 \mid \text{item}) + (1 + \text{phonotactic-probability} \mid \text{participant})$

The model was fit with four Markov Chain Monte Carlo (MCMC) chains, each run for 6,000 iterations (1,000 warm-up). Sampling diagnostics indicated successful convergence (all Rhat values were approximately 1.00, indicating convergence, and all effective sample sizes (Bulk ESS, Tail ESS) were well above commonly recommended minimum thresholds ($\approx 1,000$)). Priors on the fixed effects and on the intercepts were specified as weakly informative normal distributions with ($\mu = 0, \sigma = 5$), to shrink estimates toward zero enough to prevent overfitting but remain broad enough to permit non-trivial parameter values.

For the word identification task, results most directly tied to our research question come from the interaction effects rather than the simple effects. The interaction effects in this model indicate how confidence responses to different item type and/or phonotactic probability vary across language backgrounds relative to Hawaiian speakers' responses to real words. Recall that participants rated each item on a 5-point scale indicating their confidence that the word was a Hawaiian word where responses were coded as follows: 1 = Confident the word is not a Hawaiian word, 2 = Somewhat confident the word is not a Hawaiian word, 3 = Not confident whether the word is or is not a Hawaiian word, 4 = Somewhat confident the word is a Hawaiian word, and 5 = Confident the word is a Hawaiian word. All model results were ex-

pressed on the probit scale, which means the model estimates reflect changes in an underlying latent response variable rather than the original confidence ratings. In this context, the latent response variable represents an internal continuous sense of confidence drawn from a lexicon or proto-lexicon that is assumed to guide participants' choices among the ordered response categories. The model therefore estimates how predictors shift participants along this internal confidence continuum rather than predicting the numeric ratings themselves. For all Bayesian model estimates, confident differences are indicated by 95% CIs that exclude zero. The magnitude of each estimate is then used to assess the direction of the effect and its distance from the reference level.

We first consider the two-way interaction between phonotactic probability and language background. A negative interaction between these factors would indicate that a given non-Hawaiian language group is less influenced by phonotactic probability than Hawaiian speakers when providing confidence ratings for real words. A positive two-way interaction between phonotactic probability and language background, in contrast, would indicate that a given non-Hawaiian language group was more influenced by phonotactic probability than Hawaiian speakers when rating real words. We hypothesize that all non-Hawaiian groups will show negative interaction values, reflecting reduced influence of phonotactic probability overall in the confidence ratings of real-nonword items. Among non-Hawaiian speaking groups, Pidgin speakers are expected to be closest to zero (closer to the reference level Hawaiian speakers in their performance), followed by English speakers from Hawai'i who are predicted to show negative values of a larger magnitude than the Pidgin speakers (further away from zero), and lastly English speakers from Canada are expected to show the largest negative interaction values, reflecting the greatest divergence from Hawaiian speakers in their influence of phonotactic probability in their confidence ratings.

Next, we examine the two-way interaction between language background and word type. A positive two-way interaction between language background and word type would indicate that the difference in confidence ratings for real and nonword items is smaller for that language group compared to Hawaiian speakers. In other words, nonword items are rated more similarly to real items. Large positive values indicate that the group is much less sensitive to the real versus nonword distinction in their confidence ratings, reflecting a blurred boundary between real and nonword item types. Values close to zero indicate a pattern similar to the confidence ratings of Hawaiian speakers. Large negative values indicate that the group is more sensitive to this distinction than Hawaiian speakers, showing strong differences between real

and nonword items. According to our hypothesis, all non-Hawaiian groups should show positive values for this interaction. Among these groups, Pidgin speakers are expected to be closest to zero, English speakers from Hawai'i are expected to show larger positive magnitudes, and English speakers from Canada are expected to show the greatest positive magnitudes.

A three-way interaction between phonotactic probability, language background, and item type captures whether the influence of phonotactic probability on confidence ratings varies simultaneously by both language group and whether the item is real or nonword. Given that the reference levels are Hawaiian speakers and real items, the three-way interaction reflects how much additional change occurs in the phonotactic probability slope for a non-Hawaiian language group when the item type shifts from real to nonword. This interaction allows us to determine not only whether different language groups differ in how strongly phonotactic probability shapes their confidence ratings, but also whether that difference depends on the lexical status of the item. Negative values for this interaction indicate that a given non-Hawaiian language group shows a weaker relationship between phonotactic probability and confidence ratings specifically for nonword items, relative to Hawaiian speakers rating those same nonword items. In other words, when evaluating nonword items, this would indicate non-Hawaiian speaker groups rely less on phonotactic probability than Hawaiian speakers do. Larger negative magnitudes would reflect a greater reduction in sensitivity to phonotactic probability for nonword items, while values closer to zero would suggest a pattern more similar to that of Hawaiian speakers. Positive values would indicate the opposite pattern in which a non-Hawaiian group is more influenced by phonotactic probability than Hawaiian speakers when rating nonword items.

Simple effects are also interpreted throughout the analysis to contextualize the comparisons between language background groups, since they represent the baseline relationships for the reference group. The simple effect of phonotactic probability reflects how participant word–nonword confidence ratings changed with increasing phonotactic probability among Hawaiian speakers when evaluating real words. A positive estimate indicates that items with higher phonotactic probability received higher confidence ratings. The simple effect of language background captures overall differences in rating tendencies across groups, with negative estimates indicating that non-Hawaiian speaking groups gave lower confidence ratings for real words than did Hawaiian speakers. Finally, the simple effect of item type shows how Hawaiian speakers rated nonword items compared to real ones, with negative estimates indicating that they gave nonword items lower confidence ratings consistent with strong differentiation between words and nonwords.

A three-way interaction between phonotactic probability, language background, and item type tests whether the effect of phonotactic probability on confidence ratings changes simultaneously as a function of both the speaker group and whether the item is real or nonword. Because the model is parameterized with Hawaiian speakers and real items as the reference levels, the three-way interaction coefficients quantify how much additional change in the phonotactic probability slope occurs for a non-Hawaiian group specifically when the item shifts from real to nonword.

In this parameterization, a negative three-way interaction value indicates that, for nonword items, the phonotactic probability slope for a given non-Hawaiian group becomes more reduced (i.e., less positive or more negative) compared to the corresponding slope for Hawaiian speakers evaluating nonword items. Values closer to zero indicate that the group resembles Hawaiian speakers more closely in how phonotactic probability affects confidence for nonword items. Importantly, these coefficients do not tell us the full slope for any group or any item type on their own, and instead indicate how the slope changes relative to the reference group and condition.

For this reason, the three-way interaction should be viewed as a structural indicator confirming that slope differences exist across both language groups and item types, but it does not provide interpretable effect sizes on its own. To tease apart how the three-way interaction presents across language groups and word types, and to evaluate whether the hypothesized gradient of similarity to Hawaiian speakers emerges for real and nonword items across phonotactic probability differences, we proceed with follow-up comparisons using estimated marginal means and estimated slope differences.

5.8.2 Estimated Marginal Means and Trend Analyses

We examined differences and slopes across combinations of language background and item type for confidence ratings using an estimated marginal means (also referred to as least squares means) and estimated slopes analyses with the *emmeans* package in R (Lenth [2025]). To assess overall differences in predicted confidence ratings between language groups and item type, we extracted estimated marginal means using the `emmeans()` function. To assess how sensitivity to phonotactic probability varied across language groups and item type, we extracted estimated slopes using the `emtrends()` function. These additional analyses allow for direct comparison of predicted responses and slopes across language groups and item types, providing an interpretable summary of the model’s complex interactions.

To interpret these results, we rely on the standard approach of exam-

ining both estimated marginal means and estimated slopes. Differences in estimated marginal means show whether certain combinations of language group and item type tend to produce higher or lower overall confidence ratings than others, independent of the effect of phonotactic probability. In contrast, differences in estimated slopes (emtrends) show whether language group and item type combinations differ in how their confidence responses change as a function of phonotactic probability. Here, sensitivity to phonotactic probability is reflected in the steepness of the slope relating confidence ratings to phonotactic probability, with steeper slopes indicating that higher-probability items elicited higher confidence ratings. Overall, these comparisons allow us to disambiguate general rating tendencies from phonotactic sensitivity, providing a clearer picture of both overall response patterns and the extent to which different language backgrounds and item types modulate how strongly phonotactic probability shapes confidence ratings.

From the estimated marginal means analysis, the estimate represents the predicted difference in confidence ratings between each combination of language group and word type. This allows us to examine baseline differences between all relevant combinations without considering phonotactic probability. For example, we can compare baseline confidence ratings for real versus nonwords within a single language group, such as real items compared to nonword items for Hawaiian speakers. We can also compare baseline differences in confidence ratings between all combinations of language groups for real and nonword items. For instance, we can compare the confidence ratings of real words for Hawaiian speakers to the ratings of real words for each non-Hawaiian speaker group, as well as compare the confidence ratings of real words among the non-Hawaiian groups themselves. We can make the same set of comparisons for nonwords by comparing the confidence ratings of nonword items across each language group.

When comparing marginal means, positive values indicate that the first group has a higher predicted confidence rating than the second, whereas negative values indicate that the first group has a lower predicted confidence rating than the second. Values near zero indicate that the groups differ very little in their predicted confidence ratings. The 95% credible interval for this difference is given by the highest posterior density (HPD) interval. 95% HPD intervals that exclude zero indicate that the difference between the marginal means is credible, while intervals that include zero suggest that the difference could plausibly be zero.

From the estimated slopes analysis, the model provides the estimated effect of phonotactic probability for each combination of language group and item type. These slopes represent how strongly confidence ratings change

in response to phonotactic probability within each combination of language group and word type. Examining these slopes makes it possible to evaluate how the influence of phonotactic probability on confidence ratings differs across groups and word types.

When comparing estimated slopes, a positive contrast value indicates that the first group in the contrast shows greater sensitivity to phonotactic probability than the second group. In other words, the first group has a steeper increase in predicted confidence as phonotactic probability increases. A negative contrast value indicates that the first group is less sensitive than the second group, meaning the slope is shallower. Larger absolute contrast values indicate larger differences in sensitivity, and values that are closer to zero indicate that the groups have similar slopes. Again, 95% HPD intervals that exclude zero indicate that the difference between the marginal means is credible, while intervals that include zero suggest that the difference could plausibly be zero.

We hypothesize that sensitivity to item type will follow a gradient reflecting participants' familiarity with Hawaiian. From the estimated marginal means analysis we expect the largest difference between real and nonword items among Hawaiian speakers, with progressively smaller real–nonword differences among Pidgin speakers, English speakers from Hawai'i, and English speakers from Canada. We further expect that pairwise contrasts in estimated marginal means will likewise follow this gradient. Specifically, differences between Hawaiian and Pidgin speakers should be the smallest, followed by differences between Hawaiian and English speakers from Hawai'i, and finally the largest differences between Hawaiian and English speakers from Canada.

From the estimated slopes analysis we expect the steepest positive slopes for Hawaiian speakers for both real and nonword items, indicating the strongest use of probabilistic phonotactic knowledge. Pidgin speakers are expected to show shallower slopes than Hawaiian speakers, followed by even shallower slopes for English speakers from Hawai'i, suggesting decreasing influence of knowledge on phonotactic structure across these groups. Finally, English speakers from Canada are expected to show near zero slopes, consistent with little or no sensitivity to Hawaiian phonotactic probability.

5.8.3 Results: Posterior Estimates from the Bayesian Mixed-Effects Model

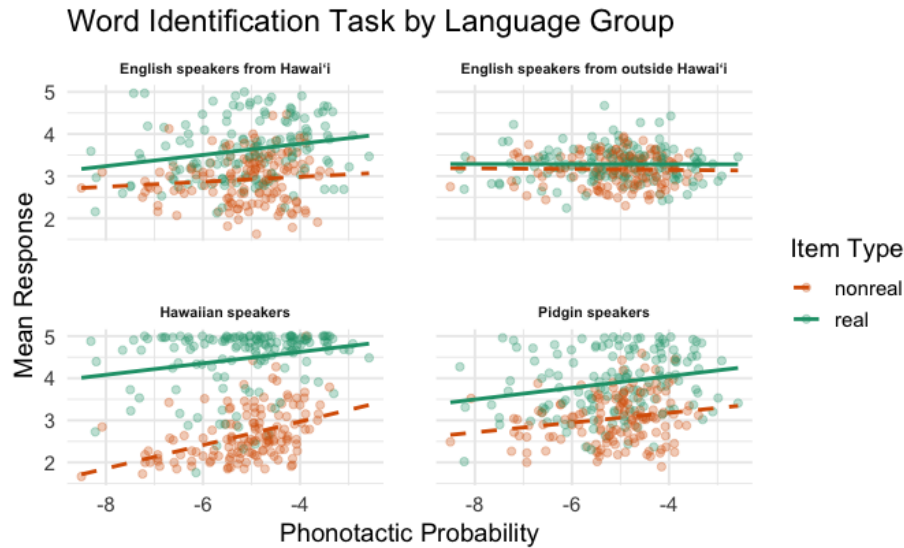


Figure 5.3: Empirical relationship between phonotactic probability and mean confidence rating responses in the word identification task across four language groups. Each panel displays one language group: Hawaiian speakers, Pidgin speakers, English speakers from Hawai'i, and English speakers from outside Hawai'i. The x-axis represents the phonotactic probability of each item (the likelihood of its sound sequence occurring in Hawaiian), while the y-axis shows the mean response across participants for that item. Data points represent individual items, with separate colors for real Hawaiian words (green) and nonword items (orange).

5.8. Word Identification Task: Statistical Analyses and Results

Predictor	Estimate	Est. Error	95% CI
Phonotactic Probability	0.20	0.05	0.11, 0.29
Pidgin speakers from Hawai'i	-0.85	0.11	-1.07, -0.64
English speakers from Hawai'i	-1.23	0.12	-1.46, -0.98
English speakers from Canada	-1.59	0.10	-1.80, -1.39
Word type (nonword)	-2.13	0.06	-2.25, -2.01
Phonotactic Probability $\times$ Pidgin speakers from Hawai'i	-0.07	0.03	-0.14, -0.00
Phonotactic Probability $\times$ English speakers from Hawai'i	-0.10	0.04	-0.18, -0.03
Phonotactic Probability $\times$ English speakers from Canada	-0.25	0.03	-0.31, -0.19
Phonotactic Probability $\times$ Word type (nonword)	0.21	0.06	0.08, 0.32
Pidgin speakers from Hawai'i $\times$ Word type (nonword)	1.28	0.03	1.22, 1.35
English speakers from Hawai'i $\times$ Word type (nonword)	1.52	0.03	1.45, 1.59
English speakers from Canada $\times$ Word type (nonword)	2.11	0.03	2.05, 2.17
Phonotactic Probability $\times$ Pidgin speakers from Hawai'i $\times$ Word type (nonword)	-0.13	0.03	-0.19, -0.06
Phonotactic Probability $\times$ English speakers from Hawai'i $\times$ Word type (nonword)	-0.17	0.03	-0.24, -0.10
Phonotactic Probability $\times$ English speakers from Canada $\times$ Word type (nonword)	-0.09	0.03	-0.14, -0.03

Table 5.3: Population-level effects from the Bayesian model predicting confidence ratings. The model includes fixed effects for phonotactic probability, language background, and item type, as well as all interactions among these predictors. Estimates represent posterior means with 95% credible intervals (CIs). Effects whose CIs exclude zero are interpreted as credible differences. Positive coefficients indicate higher confidence ratings (i.e., stronger confidence that an item is a real word) as a function of the corresponding predictor. Interaction terms capture how these relationships vary across language backgrounds and item types.

Model results are summarized in Table 5.3. Given the focus on interaction effects, I first interpret the two-way and three-way interactions, and then address the simple effects in the model output.

Relative to the reference group of Hawaiian speakers, the two-way interaction between phonotactic probability and language background indicated group differences in the slope relating phonotactic probability to confidence ratings. For Pidgin speakers, the posterior interval for the interaction included zero (Estimate = -0.07, 95% CI = [-0.14, -0.00]). For English speakers from Hawai'i, the interaction term indicated a reduction of 0.10 in the phonotactic probability slope (Estimate = -0.10, 95% CI = [-0.18, -0.03]). For English speakers from Canada, the reduction was 0.27 (Estimate = -0.27, 95% CI = [-0.31, -0.19]). The simple effect of phonotactic probability for the reference group of Hawaiian speakers was positive (Estimate = 0.20, 95% CI = [0.11, 0.29]).

These results indicate that the positive relationship between phonotactic probability and confidence ratings observed for Hawaiian speakers was progressively reduced for the other groups. For Hawaiian speakers, higher phonotactic probability was associated with higher confidence ratings across real and nonword types, and the credible interval for this effect did not in-

clude zero, indicating a clearly positive relationship with a slope of 0.20. The interaction term for Pidgin speakers was smallest compared to all non-Hawaiian speaking groups and included zero, indicating no evidence of a meaningful difference in the slope relating phonotactic probability to confidence ratings across word types compared to Hawaiian speakers. When combined with the baseline slope, this interaction corresponds to an estimated slope of approximately 0.13 for Pidgin speakers ($0.20 - 0.07 = 0.13$). English speakers from Hawai'i demonstrated a larger reduction in slope, with the credible interval indicating a meaningful difference from Hawaiian speakers. This suggests that the relationship between phonotactic probability and confidence ratings was weaker for English speakers from Hawai'i than for Hawaiian speakers. Combining the interaction with the baseline slope yields an estimated slope of approximately 0.10 ($0.20 - 0.10 = 0.10$), suggesting that although the relationship remained positive across word types, English speakers from Hawai'i showed reduced sensitivity to phonotactic probability compared to Hawaiian speakers and Pidgin speakers. Lastly, English speakers from Canada showed the largest reduction in slope. Combining this interaction with the baseline slope yields an estimated slope of approximately -0.07 ($0.20 - 0.27 = -0.07$), suggesting no positive relationship between phonotactic probability and confidence ratings for English speakers from Canada. Overall, this pattern indicates a gradient in sensitivity to phonotactic probability, with Hawaiian speakers showing the strongest positive relationship, followed by Pidgin speakers, English speakers from Hawai'i, and English speakers from Canada.

The two-way interaction between word type and language background also revealed differences relative to Hawaiian speakers. Compared to Hawaiian speakers, Pidgin speakers showed an increase of 1.28 in confidence ratings for nonwords (Estimate = 1.28, 95% CI = [1.22, 1.35]). English speakers from Hawai'i showed an increase of 1.52 (Estimate = 1.52, 95% CI = [1.45, 1.59]), and English speakers from Canada showed an increase of 2.11 (Estimate = 2.11, 95% CI = [2.05, 2.17]). For Hawaiian speakers, the simple effect of word type indicated lower confidence ratings for nonwords relative to real words (Estimate = -2.13, 95% CI = [-2.25, -2.01]).

This means that for Hawaiian speakers, the interaction between word type and language background indicated lower confidence ratings for nonwords compared to real words, reflecting the strongest discrimination between real words and nonwords. The interaction term for Pidgin speakers indicated a smaller reduction in this difference, yielding an estimated word type difference of 0.85 ($2.13 + 1.28$), suggesting that Pidgin speakers still rated nonwords lower than real words but were less confident in their dis-

inction than Hawaiian speakers. English speakers from Hawai'i showed a larger reduction in the difference, producing an estimated slope of 0.61 ($2.13 + 1.52$), indicating a weaker distinction between nonwords and real words relative to Hawaiian and Pidgin speakers. English speakers from Canada showed the largest reduction, corresponding to an estimated difference of 0.02 ($2.13 + 2.11$), suggesting that nonwords and real words were rated almost equivalently, reflecting minimal differentiation on the confidence scale.

The model also included a three-way interaction between phonotactic probability, word type, and language background. Relative to Hawaiian speakers, Pidgin speakers showed an adjustment of -0.13 in the phonotactic probability slope for nonwords (Estimate = -0.13, 95% CI = [-0.19, -0.06]). English speakers from Hawai'i showed an adjustment of -0.17 (Estimate = -0.17, 95% CI = [-0.14, -0.03]). English speakers from Canada showed an adjustment of -0.09 (Estimate = -0.09, 95% CI = [-0.14, -0.03]). For Hawaiian speakers, the simple effect associated with this interaction was positive (Estimate = 0.21, 95% CI = [0.08, 0.32]).

For Hawaiian speakers, the three-way interaction between phonotactic probability, word type, and language background revealed that phonotactic probability had a positive effect for nonwords. This indicates that higher phonotactic probability slightly increased confidence ratings for nonwords. This suggests that, when Hawaiian speakers encountered a word they did not know, they appeared to rely on their knowledge of phonotactic patterns to guide their responses, a strategy that was stronger than that observed in the non-Hawaiian speaking groups. Relative to Hawaiian speakers, all non-Hawaiian speakers showed a reduction in this slope.

It is important to reiterate that these three-way interaction effects apply specifically to the rating of nonwords compared to the reference level of the association between language background, word type and phonotactic probability and therefore do not, on their own, represent the complete pattern of how phonotactic probability shapes confidence ratings across groups. The interpretation of this gradient can be disentangled in the follow-up emmeans and emmtrends analyses, which clarify the structure of the interaction by decomposing it into interpretable group contrasts and differences in phonotactic probability slopes. The next section presents the estimated marginal means and trend analyses. Posterior estimates will be interpreted further after these analyses are discussed.

5.8.4 Results: Estimated Marginal Means and Trend Analyses

Contrast	Estimate	95% HPD
Hawaiian speaker real - Hawaiian speaker nonword	2.1307	2.0109, 2.2463
Pidgin speaker real - Pidgin speaker nonword	0.8484	0.7365, 0.9572
English speaker from Hawai'i real - English speaker from Hawai'i nonword	0.6091	0.4933, 0.7243
English speaker from Canada real - English speaker from Canada nonword	0.0228	-0.0859, 0.1308

Table 5.4: Posterior contrasts of estimated marginal means (emmeans) analysis for confidence ratings presented on the probit scale. Each row shows the median posterior contrast comparing the first group listed to the second, along with the 95% highest posterior density interval. These contrasts summarize within-group differences between real and nonword item confidence ratings for Hawaiian speakers, Pidgin speakers, English speakers from Hawai'i, and English speakers from Canada.

To better understand these interactions and the gradient across groups, we conducted follow-up analyses using estimated marginal means (emmeans) and estimated trends (emmtrends). These analyses decompose the complex interaction between phonotactic probability, word type (word vs. nonword), and confidence ratings into interpretable contrasts between all pairs of speaker groups. This approach allows for a clearer examination of each group's ability to differentiate words from nonwords and their reliance on phonotactic probability to guide responses for each word type, while also enabling direct comparisons between group pairs. The following section presents the emmeans and emmtrends results, which provide the basis for a more detailed interpretation of the posterior estimates.

We first report each group's ability to distinguish real words from nonword items, as reflected in posterior contrasts from the emmeans analysis of confidence ratings. This contrast provides a simple and interpretable summary of overall task performance. Examining within-group separation between item types establishes whether participants reliably differentiated words from nonwords before turning to more complex interaction effects. As such, this result offers a meaningful point of reference for interpreting subsequent effects in the model.

When comparing within groups on their ability to distinguish words from nonwords (Table 5.4), all groups showed clear separation between the two item types, with confident average differences in ratings, except for the English speakers from Canada (Estimate = 0.0228, 95% HPD = [-0.0859, 0.1308]). When it comes to the largest magnitude of separation between real

5.8. Word Identification Task: Statistical Analyses and Results

words and nonwords, this follows the expected order with Hawaiian speakers having the largest separation (Estimate = 2.1307, 95% HPD: [2.0109, 2.2463]), followed by Pidgin speakers (Estimate = 0.8484, 95% HPD: [0.7365, 0.9572]), followed next by English speakers from Hawai'i (Estimate = 0.6091, 95% HPD: [0.4933, 0.7243]), and lastly with English speakers from Canada with no differences between real and nonword items.

Contrast	Estimate	95% HPD
Real words		
Hawaiian speaker – Pidgin speaker	0.8464	0.6418, 1.0738
Hawaiian speaker – English speaker from Hawai'i	1.2294	0.9804, 1.4644
Hawaiian speaker – English speaker from Canada	1.5886	1.3931, 1.8002
Pidgin speaker – English speaker from Hawai'i	0.3823	0.1585, 0.5971
Pidgin speaker – English speaker from Canada	0.7422	0.5724, 0.9159
English speaker from Hawai'i – English speaker from Canada	0.3617	0.1693, 0.5722
Nonwords		
Hawaiian speaker – Pidgin speaker	-0.4357	-0.6461, -0.2130
Hawaiian speaker – English speaker from Hawai'i	-0.2910	-0.5429, -0.0623
Hawaiian speaker – English speaker from Canada	-0.5189	-0.7108, -0.3110
Pidgin speaker – English speaker from Hawai'i	0.1416	-0.0866, 0.3473
Pidgin speaker – English speaker from Canada	-0.0837	-0.2535, 0.0892
English speaker from Hawai'i – English speaker from Canada	-0.2247	-0.4181, -0.0113

Table 5.5: Posterior contrasts of estimated marginal means (emmeans) for confidence ratings on the probit scale. Rows are grouped by word type (real vs nonword). Each contrast reports the median posterior estimate comparing the first group to the second, along with the 95% highest posterior density interval.

The estimated marginal means analysis (see Table 5.5) separates mean estimates by language group for real and nonword item types. When using contrast coding, the comparison between two groups is expressed as the difference between group A and group B (group A is subtracted by group B). Thus, a positive value indicates that group A has a larger estimated confidence rating response on the probit scale than group B, while a negative value indicates that group B is larger than group A. For real word items, a positive value indicates a tighter alignment with real word confidence ratings. The first pattern that emerges is that Hawaiian speakers provided the highest confidence ratings for real items. In all comparisons where Hawaiian speakers are group A and non-Hawaiian speakers are group B, the resulting estimates are positive. This includes the contrast between Hawaiian speaker real (group A) and Pidgin speaker real (group B) (Estimate = 0.8464, 95% HPD [0.6418, 1.0738]), Hawaiian speaker real (group A) and English speaker

from Hawai'i real (group B) (Estimate = 1.2294, 95% HPD [0.9804, 1.4644]), and Hawaiian speaker real (group A) and English speaker from Canada real (group B) (Estimate = 1.5886, 95% HPD [1.3931, 1.8002]). From these data, we can interpret that Pidgin speakers differ the least from Hawaiian speakers, followed by English speakers from Hawai'i, and then English speakers from Canada, reflecting the ordered magnitudes of their estimated probit differences from the Hawaiian speakers. Since these ratings reflect confidence that a stimulus item is a real word, the results indicate that Hawaiian speakers' judgments most accurately matched the real-word status of the items, giving higher confidence ratings for real items than every other group. These findings align with the expectation that Hawaiian speakers would outperform the other groups on this task.

Now, in comparing Pidgin speakers to the English speakers, for real words it appears Pidgin speakers are also rating these words with higher confidently real word ratings compared to English speakers from Hawai'i (Estimate = 0.3823, 95% HPD [0.1585, 0.5971]) as well as English speakers from Canada (Estimate = 0.7422, 95% HPD [0.5724, 0.9159]). These are both confident differences as the 95% HPD intervals both exclude zero. This presents compelling evidence that Pidgin speakers exhibited response patterns that corresponded more closely to the real word items compared to both English speaking groups.

Lastly, when comparing English speakers from Hawai'i with English speakers from Canada, we see confident evidence of a closer alignment with real words and higher confidence ratings for the English speakers from Hawai'i compared to those from Canada (Estimate = 0.3617, 95% HPD [0.1693, 0.5722]). This aligns with expectations that English speakers with ambient exposure to Hawaiian would show responses more closely aligned with real stimuli type than those without such exposure.

Moving onto nonwords, the interpretation of the contrasts changes. Since lower ratings indicate greater confidence that an item is nonword, closer alignment with nonword items is reflected in lower ratings. A negative estimate for the contrast between Group A and Group B indicates that Group A's ratings showed closer alignment with the nonword items than Group B's ratings. Once again, Hawaiian speakers demonstrated the closest alignment such that all Hawaiian speakers were providing lower ratings to nonwords compared to all non-Hawaiian speaking groups, including: Pidgin speakers (Estimate = -0.4357, 95% HPD [-0.6461, -0.2130]), English speakers from Hawai'i (Estimate = 0.2910, 95% HPD [-0.5429, -0.0623]), and English speakers from Canada (Estimate = -0.5189, 95% HPD [-0.7108, -0.3110]). Consistent with the negative contrast estimates, Hawaiian speakers assigned

lower confidence ratings to nonword items, reflecting greater accuracy in identifying them as nonword. This pattern parallels their high confidence ratings for real words, highlighting their stronger discrimination compared to non-Hawaiian speakers.

Focusing on nonwords for Pidgin speakers, things become a bit more complicated. When it comes to nonword items, Pidgin speakers do not show confident differences in their ratings of nonwords compared to either English speakers from Hawai'i (Estimate = 0.1416, 95% HPD [-0.0866, 0.3473]) nor English speakers from Canada (Estimate = -0.0837, 95% HPD [-0.2535, 0.0892]). This indicates that there is no evidence of closer alignment for Pidgin speakers in their ratings of nonword items compared to either English from Hawai'i or English from Canada in regard to their mean responses. However, these data do indicate that the English speakers from Hawai'i demonstrated better alignment in their ratings compared to the English speakers from Canada (Estimate = -0.2247, 95% HPD [-0.4181, -0.0113]). When nonword responses are averaged across phonotactic probability, Pidgin speakers do not differ reliably from either English-speaking group, whereas English speakers from Hawai'i continue to show higher confidence ratings for nonwords than English speakers from Canada. However, this pattern reflects an overall mean difference and should be interpreted in light of the significant three-way interaction: Pidgin speakers' nonword ratings vary as a function of phonotactic probability, while English speakers from Canada show little modulation by phonotactic probability. Thus, the apparent similarity between Pidgin and English-speaking groups in the two-way contrast masks differences in how phonotactic structure influences nonword judgments.

5.8. Word Identification Task: Statistical Analyses and Results

Contrast	Estimate	95% HPD
Real words		
Hawaiian speaker – Pidgin speaker	0.0735	0.0042, 0.1413
Hawaiian speaker – English speaker from Hawai'i	0.1052	0.0304, 0.1798
Hawaiian speaker – English speaker from Canada	0.2501	0.1901, 0.3148
Pidgin speaker – English speaker from Hawai'i	0.0310	-0.0347, 0.0986
Pidgin speaker – English speaker from Canada	0.1763	0.1244, 0.2298
English speaker from Hawai'i – English speaker from Canada	0.1455	0.0838, 0.2057
Nonwords		
Hawaiian speaker – Pidgin speaker	0.2009	0.1286, 0.2721
Hawaiian speaker – English speaker from Hawai'i	0.2783	0.2001, 0.3583
Hawaiian speaker – English speaker from Canada	0.3375	0.2750, 0.4047
Pidgin speaker – English speaker from Hawai'i	0.0785	0.0097, 0.1526
Pidgin speaker – English speaker from Canada	0.1373	0.0825, 0.1948
English speaker from Hawai'i – English speaker from Canada	0.0593	-0.0076, 0.1246

Table 5.6: Posterior contrasts from the estimated trends (emmtrends) analysis, showing differences in phonotactic probability slopes across all combinations of language group and item type. Each contrast reports the median posterior difference in slope between the first group listed and the second, along with the associated 95% highest posterior density interval. These results summarize pairwise differences in how confidence ratings change as a function of phonotactic probability for Hawaiian, Pidgin, English from Hawai'i, and English from Canada speakers for both real and nonword items.

Next, we examine the posterior contrasts from the estimated trends (emmtrends) analysis (Table 5.6). These contrasts quantify how phonotactic probability modulated each language groups' ratings of words versus nonwords.

With contrast coding of the slopes, the estimates represent differences in phonotactic probability slopes between group A and group B. Stronger phonotactic knowledge displays stronger influence of phonotactic probabilities in confidence ratings. For real words, Hawaiian speakers exhibit the steepest estimated slopes among all groups, including Pidgin speakers (Estimate = 0.073532, 95% HPD: [0.00415, 0.14127]), English speakers from Hawai'i (Estimate = 0.105205, 95% HPD: [0.03036, 0.17984]), and English speakers from Canada (Estimate = 0.250129, 95% HPD: [0.19007, 0.31482]). This pattern is evident in all contrasts being positive at a credible interval, where Hawaiian speakers are coded as group A and each non-Hawaiian group as group B. Also noteworthy here is that the magnitudes of the contrasts follow the expected pattern with regard to non-Hawaiian speaking groups compared to the Hawaiian speakers. The Pidgin speakers show the smallest

deviation from the Hawaiian speakers, the English speakers from Hawai'i show a larger deviation, and the English speakers from Canada exhibit the largest difference in slope relative to the Hawaiian speakers.

The estimated trends analysis does not indicate evidence that Pidgin speakers are influenced more strongly by phonotactic probability for real words compared to English speakers from Hawai'i, as their 95% HPD interval includes zero (Estimate = 0.031043, 95% HPD: [-0.03468, 0.09861]). However, Pidgin speakers show a clear and reliable difference in their real word phonotactic probability slope compared to that of the English speakers from Canada (Estimate = 0.176264, 95% HPD: [0.12438, 0.22983]). English speakers from Hawai'i also differ reliably from the English speakers from Canada in their real word phonotactic probability slope (Estimate = 0.145519, 95% HPD: [0.08375, 0.20571]).

In comparing the slopes of language groups ratings of nonwords, we would expect phonotactic probability to modulate confidence ratings in the positive direction as with real words if the speaker has strong knowledge of Hawaiian phonotactics. Similarly to real words, Hawaiian speakers are more strongly influenced by phonotactic probability compared to all non-Hawaiian groups for nonwords, including Pidgin speakers (Estimate = 0.200907, 95% HPD: [0.12864, 0.27208]), English speakers from Hawai'i (Estimate = 0.278334, 95% HPD: [0.20007, 0.35832]), and English speakers from Canada (Estimate = 0.337533, 95% HPD: [0.27502, 0.40468]).

For nonwords, Pidgin speakers show a reliably stronger influence of phonotactic probability on their confidence ratings compared to both English-speaking groups, indicating that their judgments are more strongly modulated by phonotactic structure even if mean ratings are similar. This effect is evident when comparing Pidgin speakers to English speakers from Hawai'i (Estimate = 0.078490, 95% HPD: [0.00966, 0.15259]) as well as to English speakers from Canada (Estimate = 0.137274, 95% HPD: [0.08245, 0.19481]). English speakers from Hawai'i do not reliably differ from English speakers from Canada in their phonotactic probability slopes for nonwords (Estimate = 0.059314, 95% HPD: [-0.00762, 0.12462]).

5.8. Word Identification Task: Statistical Analyses and Results

Contrast	Estimate	95% HPD
Hawaiian real - Hawaiian nonword	-0.207070	-0.31603, -0.07911
Pidgin real - Pidgin nonword	-0.080457	-0.18944, 0.03811
English from Hawai'i real - English from Hawai'i nonword	-0.032127	-0.14602, 0.08650
English from Canada real - English from Canada nonword	-0.118881	-0.23068, -0.00914

Table 5.7: Posterior contrasts from the estimated trends (emmtrends) analysis, showing differences in phonotactic probability slopes across all combinations of language group and item type. Each contrast reports the median posterior difference in slope between the first group listed and the second, along with the associated 95% highest posterior density interval. These results summarize pairwise differences in how confidence ratings change as a function of phonotactic probability within each language population for real and nonword stimuli items.

Finally, when examining differences in slope across language groups, we do not anticipate systematic differences between real-word and nonword-word phonotactic probability slopes. Consistent with this expectation, Pidgin speakers and English speakers from Hawai'i show no reliable real–nonword slope differences; their 95% HPD intervals include zero (for Pidgin real – Pidgin nonword Estimate = -0.080457 , 95% HPD: $[-0.18944, 0.03811]$; for English from Hawai'i real – English from Hawai'i nonword Estimate = -0.032127 , 95% HPD: $[-0.14602, 0.08650]$). These patterns suggest that both groups draw on phonotactic probability in a similar manner for real and nonword items.

In contrast, Hawaiian speakers show a real–nonword slope difference: the negative estimate (Estimate = -0.207070 , 95% HPD $[-0.31603, -0.07911]$) indicates that their nonword slopes are slightly steeper than their real-word slopes, reflecting greater reliance on phonotactic probability when evaluating nonwords. English speakers from Canada also show a modest difference in the same direction (Estimate = -0.118881 , 95% HPD: $[-0.23068, -0.00914]$). For Hawaiian speakers, this result indicates that phonotactic probability plays a somewhat larger role in confidence judgments for nonwords than for real words. For English speakers from Canada, the slopes are sufficiently close to zero that they do not appear to reflect meaningful phonotactic knowledge. As such, the small observed difference is unlikely to be interpretable and may simply represent random variation rather than a meaningful response pattern.

Returning to the Bayesian fixed-effects estimates (Table 5.2), Hawaiian speakers exhibited the strongest baseline discrimination between real words and nonwords (Estimate = -2.13 , 95% CI $[-2.25, -2.01]$). Non-Hawaiian

groups showed weaker discrimination, consistent with the emmeans contrasts: Pidgin speakers (Estimate = 1.28, 95% CI [1.22, 1.35]), English speakers from Hawai'i (Estimate = 1.52, 95% CI [1.45, 1.59]), and English speakers from Canada (Estimate = 2.11, 95% CI [2.05, 2.17]) all showed progressively larger deviations relative to Hawaiian speakers in nonword confidence ratings.

The positive main effect of phonotactic probability (Estimate = 0.20, 95% CI [0.11, 0.29]) indicates that, on average, participants' confidence ratings were systematically influenced by phonotactic structure. Interaction terms further clarify group differences: Pidgin speakers show only a slight reduction in phonotactic probability sensitivity relative to Hawaiian speakers (Estimate = -0.07, 95% CI [-0.14, -0.00]), English speakers from Hawai'i show a moderate reduction (Estimate = -0.10, 95% CI [-0.18, -0.03]), and English speakers from Canada show the largest reduction (Estimate = -0.25, 95% CI [-0.31, -0.19]).

Finally, the three-way interactions confirm that phonotactic probability effects are strongest for Hawaiian speakers and less pronounced for non-Hawaiian speakers, particularly for nonwords: Pidgin speakers (Estimate = -0.13, 95% CI [-0.19, -0.06]), English speakers from Hawai'i (Estimate = -0.17, 95% CI [-0.24, -0.10]), and English speakers from Canada (Estimate = -0.09, 95% CI [-0.14, -0.03]). These fixed-effects estimates support and contextualize the patterns observed in the emmeans and emmtrends analyses, showing that both baseline confidence and phonotactic sensitivity vary systematically by language group and item type.

5.9 Word Identification Task: Discussion

Analysis of the Word Identification Task revealed clear differences in how phonotactic probability interacted with language background in shaping word confidence ratings relative to the 'ōlelo Hawai'i speaker reference group. The observed group differences follow the expected direction and indicate gradient variation across speaker groups, consistent with our hypothesis. Specifically, the results align with the predicted continuum of 'ōlelo Hawai'i lexical and proto-lexical knowledge. 'ōlelo Hawai'i speakers demonstrate the strongest and most integrated lexical and proto-lexical knowledge, followed by Pidgin speakers, then English speakers from Hawai'i, and finally English speakers from Canada. This overall pattern is illustrated in Figure 5.3.

Across analyses, 'ōlelo Hawai'i speakers showed the clearest distinction between real words and nonwords, and the steepest positive slopes of all

groups, indicating the strongest alignment between phonotactic probability and confidence ratings for both real words and nonwords. Together, these findings support the construct validity of the assessment, as it appropriately identifies ‘ōlelo Hawai‘i speakers as those with the most robust and systematic knowledge of Hawaiian lexical structure.

Further supporting the validity of the assessment, English speakers from Canada, at the opposite end of the spectrum, did not demonstrate the ability to reliably distinguish between real words and nonwords and demonstrated the weakest interaction with phonotactic probability.

The differences between Pidgin speakers and English speakers from Hawai‘i highlight subtle but meaningful variation in the depth and organization of ‘ōlelo Hawai‘i phonotactic knowledge among non-‘ōlelo Hawai‘i speaking local populations.

Multiple sources of evidence from the Word Identification Task indicate that Pidgin speakers possess a more robust proto-lexicon compared to other non-‘ōlelo Hawai‘i speakers. First, they demonstrate a larger number of Hawaiian word forms than English speakers from Hawai‘i and English speakers from Canada, reflected in greater separation between real words and nonwords on the task (Table 5.4). Second, their performance shows the closest correspondence in slope to ‘ōlelo Hawai‘i speakers for both real words and nonwords, suggesting more speaker-like patterns of lexical sensitivity. Third, Pidgin speakers more systematically utilize phonotactic probabilities to guide confidence rating responses, indicating a sub-lexical evaluation strategy that is better calibrated and more speaker-like than all other non-Hawaiian speaking groups.

With respect to the interaction between phonotactic probability, word type, and confidence ratings, as reflected in the slope differences shown in Figure 5.3, Pidgin speakers also most closely pattern with ‘ōlelo Hawai‘i speakers, followed by English speakers from Hawai‘i, and finally by English speakers from Canada. Results from the emmtrends analysis (Table 5.6) empirically demonstrate that, for both real words and nonwords, the difference between Pidgin and ‘ōlelo Hawai‘i speakers was smaller than the differences observed between other non-Hawaiian speaking groups relative to Hawaiian speakers.

For nonwords, Pidgin speakers differed credibly from both English speakers from Hawai‘i (Estimate = 0.0785, 95% HPD [0.0097, 0.1526]) and English speakers from Canada (Estimate = 0.1373, 95% HPD [0.0825, 0.1948]). English speakers from Hawai‘i did not differ credibly from English speakers from Canada in their ratings of nonwords (Estimate = 0.0593, 95% HPD [-0.0076, 0.1246]). Thus, while English speakers from Hawai‘i patterned similarly to

Pidgin speakers for real words in terms of the relationship between phonotactic probability and confidence ratings, they patterned more like English speakers from Canada for nonwords.

Additional complexity emerges when examining English speakers from Hawai'i. For real words, the emmtrends analysis indicates that the slope of the relationship between phonotactic probabilities and confidence ratings for Pidgin speakers did not differ from that of English speakers from Hawai'i (Estimate = 0.0310, 95% HPD [-0.0347, 0.0986]), indicating overlapping confidence intervals and similar use of phonotactic probability when evaluating real Hawaiian words. Both groups, however, differed credibly from English speakers from Canada, whose alignment between phonotactic probabilities and confidence ratings was weaker (Pidgin speakers – English speakers from Canada: Estimate = 0.1763, 95% HPD [0.1244, 0.2298]; English speakers from Hawai'i – English speakers from Canada: Estimate = 0.1455, 95% HPD [0.0838, 0.2057]). Overall, English speakers from Hawai'i display 'ōlelo Hawai'i knowledge that is intermediate between Pidgin speakers and English speakers from Canada. Compared to English speakers from Canada, they show greater separation between real words and nonwords, indicating a larger repertoire of real Hawaiian word forms, and they are equally sensitive to phonotactic probabilities when processing real words. However, when evaluating nonwords, their use of phonotactic probabilities does not differ from that of English speakers from Canada.

Previous work has shown that real words and nonwords are processed differently. As Vitevitch et al. [1999] discuss, the process of word identification differs fundamentally for nonwords and real words. Inconsistencies in the relationship between phonotactic probability and functional load for words versus nonwords have been well documented. For example, the neighborhood activation model (NAM) (Luce and Pisoni [1998]) and the cohort theory (Marslen-Wilson [1987]) both predict that real words from high density phonological neighborhoods, which are strongly correlated with high phonotactic probability, tend to elicit slower and less accurate responses than those from low density neighborhoods. However, the opposite effect has been observed for nonword stimuli, where nonwords with high phonotactic probability are identified with faster and more accurate responses (e.g., Vitevitch et al. [1997]; Vitevitch and Luce [1999]). Vitevitch and Luce [1999] argue that these inconsistencies reflect the fact that the identification of real words and nonwords draws on two distinct levels of linguistic representation. Real words rely more on lexical level information, whereas nonwords depend more heavily on sub-lexical structure (Vitevitch et al. [1999]). In line with this account, the stronger alignment between phonotactic probability

and confidence ratings for nonwords among Pidgin speakers suggests that their advantage arises at the sub-lexical, proto-lexical level rather than from differences in stored lexical knowledge alone.

The differences observed between Pidgin speakers and English speakers from Hawai'i in their rating behavior with real and nonwords can be interpreted as differences in the strength of the proto-lexicon. The proto-lexicon framework (e.g., Oh et al. [2020]), building on models such as NAM and the Cohort Model, extends theoretical expectations of language processing to listeners with partial acquisition or ambient exposure. It proposes that even without full lexical knowledge, individuals form implicit representations of frequent word forms and sub-word sequences, which then guide judgments about the well formedness of words that may be unknown to the listener (e.g., Panther et al. [2023b]). Within this framework, the stronger alignment between phonotactic probability and confidence ratings for Pidgin speakers suggests more robust and efficiently organized sub-lexical representations, as well as well-developed knowledge of word forms. In contrast, English speakers from Hawai'i show weaker sub-lexical alignment, indicating a less specified proto-lexicon, despite demonstrating greater lexical knowledge than English speakers without Hawaiian exposure. Overall, these findings demonstrate that graded exposure and structural familiarity shape sub-lexical processing strategies, bridging models of L1 unknown word perception with proto-lexical development in partially exposed populations.

Overall, these patterns suggest that Pidgin speakers not only have the most extensive knowledge of word forms but also rely most effectively on proto-lexical knowledge when identifying unfamiliar forms. English speakers from Hawai'i, with sparser representations, show intermediate sensitivity: they perform similarly to Pidgin speakers for real words but rely less effectively on proto-lexical knowledge when evaluating unfamiliar forms, while English speakers from Canada demonstrate the least robust representations overall.

English speakers from Hawai'i do demonstrate a proto-lexical advantage compared to English speakers from Canada. Their performance suggests that they possess both lexical and proto-lexical knowledge of 'ōlelo Hawai'i, although these representations appear to be less finely tuned and less efficiently utilized in word identification than those of Pidgin speakers.

Some limitations of the present task are worth noting, particularly with respect to how “real-ness” is interpreted in the Hawaiian context for this task. The question of whether a form counts as a real word presents particular challenges. For example, three Hawaiian-speaking participants reported uncertainty about how to classify foreign borrowings as real words. One such

example is *pākīpika* ‘Pacific’, as in Pacific Ocean. This form reflects a conventional Hawaiian adaptation of the English word *Pacific*, in which English fricatives /s/ and /f/ are realized as their closest Hawaiian stop counterparts /k/ and /p/, respectively, and an epenthetic vowel is added to avoid coda consonants, which are prohibited by Hawaiian syllable structure. Despite being attested in usage, some participants did not consider such borrowings to be “real” *‘ōlelo Hawai‘i* words, highlighting ongoing debates about authenticity and lexical legitimacy that complicate binary real versus nonword judgments in tasks like the present one. Although unlikely to bias groups systematically, the task of assessing word “real-ness” can be particularly difficult in research on Indigenous or minoritized languages, where speakers may navigate conflicting ideas of what constitutes a “correct” or legitimate word.

Notably, in the Word Identification task, no negative relationship between confidence ratings and phonotactic probabilities was found for any population group, even when stimuli containing Hawaiian specific orthography (*kahakō* and *‘okina*) were included. In contrast, in the Well Formedness task, all non Hawaiian speaking groups showed a negative relationship between well formedness ratings and phonotactic probabilities when nonword items with Hawaiian specific orthography were included. Furthermore, the English from Hawai‘i and English from Canada groups remained negatively aligned even when these items were excluded.

The two tasks included in Experiment 3 asked fundamentally different questions (e.g., rating how “Hawaiian-like” each nonword item was compared to rating confidence that a word was real or not). Thus, behavior in response to these two questions is expected to differ, as participants were evaluating different aspects of the stimuli across tasks. The two tasks differ in what the midpoint of the response scale represents. In the Word Identification task, intermediate responses may reflect uncertainty about whether a token is a word or nonword, as participants are making a categorical lexical decision. In contrast, the Well-Formedness task measures gradient judgments of resemblance to Hawaiian phonological patterns, where intermediate ratings do not necessarily indicate uncertainty but rather that a token is perceived as neither strongly Hawaiian-like nor strongly non-Hawaiian-like. Importantly, “not Hawaiian-like” refers to the degree to which a token resembles Hawaiian phonological structure, rather than whether the token is an attested Hawaiian word. Thus, the Well-Formedness task captures gradient phonological similarity, whereas the Word Identification task captures categorical lexical recognition. This distinction may affect the range of responses expected from participants, particularly among non-Hawaiian-speaking groups, because the Word Identification task permits uncertainty between lexical cat-

egories, whereas the Well-Formedness task requires participants to evaluate a token along a gradient continuum of Hawaiian-like to non-Hawaiian-like.

However, another factor which likely contributed to this difference across tasks is exposure to the stimuli within the experiment itself. Because the Well Formedness task was always performed first, followed by the Word Identification task for all participants, participants may have acquired knowledge about Hawaiian-like phonotactic patterns during the first task and applied this knowledge during the subsequent task. Previous work (e.g., Zervakis and Mazuka [2013]) shows that, even with little to no feedback regarding grammaticality, exposure to novel linguistic patterns can increase familiarization to those patterns and can subsequently influence later judgments.

A future study could address this possibility by counterbalancing the order of the two tasks across participants. Counterbalancing has been used in previous proto-lexical studies in which the same participants completed both tasks to mitigate potential order effects (e.g., Panther et al. [2023b]). Implementing counterbalancing in future work would allow for a clearer distinction between effects arising from the nature of the tasks themselves and effects resulting from participant familiarization over the course of the experiment.

5.10 Conclusion

Previous work done in a related Polynesian context found that non-Māori speaking adults from Aotearoa had a Māori proto-lexicon, and phonotactic sensitivity very similar to fluent speakers of Māori. The nature of the non-Hawaiian speaker from Hawai‘i is more complex than the context investigated in previous work due to the existence of two distinct language populations with differences in exposure to ‘ōlelo Hawai‘i.

Similarly to previous work done in the Māori context, we have found that (a) non-‘ōlelo Hawai‘i speakers from Hawai‘i are sensitive to Hawaiian phonotactic structure, and (b) non-‘ōlelo Hawai‘i speakers from Hawai‘i are able to discriminate a large number of words from nonwords. These findings suggest that non-‘ōlelo Hawai‘i speakers from Hawai‘i, including both English and Pidgin speakers, maintain a repertoire of ‘ōlelo Hawai‘i knowledge that include stored lexical items as well as sensitivity to phonotactic probabilities.

This current study reveals differences across speaker populations, with some groups demonstrating more advanced implicit knowledge than others. In particular, Pidgin speakers show evidence of a more developed proto-lexicon than English speakers from Hawai‘i. In both well formedness and word identification tasks, Pidgin speakers overall demonstrate stronger us-

age compared to English speakers from Hawai'i of phonotactic knowledge as well as knowledge of word form, especially at the sub-lexical level with the ratings of nonwords in both well formedness and word identification tasks. The only exception arises in the relationship between confidence ratings and phonotactic probability for real words, where English speakers from Hawai'i statistically overlap with Pidgin speakers. However, posterior contrast estimates comparing each group to 'ōlelo Hawai'i speakers indicate that Pidgin speakers' slopes are closer to those of 'ōlelo Hawai'i speakers than are the slopes of English speakers from Hawai'i. Overall, Pidgin speakers exhibited the smallest attenuation in well formedness ratings compared to all non-Hawaiian speaking groups and were the only non-Hawaiian speaking group with ratings of word-likeness which increased as phonotactic probability increased when 'okina and kahakō stimuli items were excluded. Pidgin speakers also exhibited the greatest separation between real and nonwords among all non-'ōlelo Hawai'i speaking groups, reflecting the strongest word–nonword discrimination. They further showed the most systematic use of phonotactic probabilities to guide nonword rejection. In contrast, English speakers from Hawai'i demonstrated strong proto-lexical sensitivity to real words, and intermediate lexical knowledge but weaker use of phonotactic probabilities when evaluating nonwords.

Overall, the proto-lexical advantage observed here appears to be gradient: Pidgin speakers outperform English speakers from Hawai'i in most tasks, and English speakers from Hawai'i also outperform English speakers from Canada.

This experiment presents several major novel contributions. First, unlike in any previous work, we demonstrate clear distinctions among language groups in their proto-lexical knowledge, relative to Hawaiian speakers and English speakers with no exposure. Proto-lexical performance followed a gradient across speaker groups, with Pidgin speakers outperforming English speakers from Hawai'i, and both groups outperforming English speakers from Canada. This ordering aligns with differences in language background, though the present study does not isolate the specific factors underlying this pattern.

Second, the subtle differences between Pidgin speakers and English speakers from Hawai'i underscore the interplay between lexical and sublexical competition, revealing consistent differences in how these groups evaluate word and nonword items. More broadly, this pattern highlights the need to develop a model of implicit language perception that accounts for graded representation strength and variability across speaker populations.

Third, we find that neither English speakers from Hawai'i nor Pidgin

speakers perform at the level of Hawaiian speakers on proto-lexical tasks. This contrasts with findings from well formedness tasks in Aotearoa, where participants showed knowledge of Māori (Oh et al. [2020]). Instead, these results align with those of Todd et al. [2023b], who find that Spanish knowledge among Texans and Californians is weaker than that of Spanish speakers. Todd et al. [2023b] present the argument, however, that their findings compared to that of Oh et al. [2020] are consistent with structural differences between Spanish and Māori (Todd et al. [2023b]). Because Hawaiian has a simpler phonotactic system than Māori, and because ambient exposure to Hawaiian in Hawai'i is considerably more limited than exposure to Māori in Aotearoa, we conclude that the weaker proto-lexical patterns among non-Hawaiian speakers in Hawai'i are more plausibly explained by reduced exposure rather than by differences in phonotactic complexity.

Overall, this evidence supports efforts in Hawai'i to increase Hawaiian language exposure for the benefit of non-Hawaiian speakers who may wish to learn the language in the future. The results show that even small increases in ambient exposure can have meaningful effects. Policies that restore Hawaiian place names or expand the presence of spoken Hawaiian, such as airport announcements, directly contribute to the enrichment of the proto-lexicon. This enrichment provides an advantage for listeners regardless of whether they have formally studied Hawaiian and regardless of whether they are Pidgin-speaking.

In addition, this study found clear evidence for a Pidgin advantage. Across both tasks, Pidgin speakers consistently showed stronger use of phonotactic knowledge and a broader set of word forms in their proto-lexicon compared to English speakers from Hawai'i. There are many possible explanations for this. One explanation is that due to linguistic similarities between Pidgin and Hawaiian, knowledge of the Pidgin language itself generalizes to better implicit knowledge of the Hawaiian language than the English language does. Another explanation may be that Pidgin speakers are a part of a community with increased exposure to the Hawaiian language, even compared to the English from Hawai'i language group. Currently, we cannot tease these two possible explanations apart. However, in either case, whether the effect is driven by linguistic experience or cultural associations, these findings suggest that Pidgin knowledge is associated with more robust proto-lexical knowledge than that observed in non-Pidgin speakers from Hawai'i. When considering the long-standing claim that Pidgin imposes a “linguistic deficiency” (Yokota [2008]) among students, thereby hindering their ability to learn additional languages (e.g., Yamauchi et al. [2024]), the present findings instead reveal measurable benefits for Pidgin speakers. Future work, follow-

ing approaches such as those in Mattingley et al. [2024b], may investigate whether this advantage extends to measurable gains in Hawaiian language acquisition or classroom learning outcomes for Pidgin speakers.

The implications extend beyond Hawai‘i and the Pacific. Many communities around the world debate whether increasing the visibility and audibility of Indigenous languages has practical value. The present study demonstrates that it likely does. Even low-level exposure may likely form a foundation for phonotactic knowledge that benefits language learners, strengthens community engagement with the language, and supports the long-term vitality of revitalization movements.

5.11 Mo‘olelo

In conversations with many people in Hawai‘i while recruiting participants for this study, I often heard some variation of, “If your research finds an advantage, it would give me hope and could change the way I think about Pidgin.” I appreciated how deeply people were invested in this question, but it also left me with a strong sense of responsibility. I often found myself wondering, “What if I don’t find that?” And just as importantly, “What if I do?”

This was why it was so important throughout this work to have different kinds of conversations. Rather than suggesting that the value of Pidgin depended on the outcome of our research findings, I tried to emphasize the many reasons we already know Pidgin is valuable, regardless of what any linguistics study might find.

In Experiment 3, we did find evidence of a benefit for Pidgin. That result brought a different kind of responsibility. As a researcher from the community, I felt it was especially important not to embellish or overstate what the findings meant. A single study cannot determine the value of a language, nor should the worth of a community’s language depend on scientific evidence. A single empirical study also necessarily reflects the specific measures, tasks, and methodological choices involved. For this reason, the findings must be interpreted carefully and reported as one piece of evidence within a larger body of knowledge.

So how should we interpret the findings from Experiment 3?

Overall, these findings suggest that there is something about being a Pidgin speaker that supports performance on tasks measuring linguistic

knowledge of ‘ōlelo Hawai‘i compared to English speakers from Hawai‘i. At the same time, there is also something about living in Hawai‘i, regardless of whether someone speaks Pidgin or English, that supports greater ‘ōlelo Hawai‘i knowledge compared to people who do not live in Hawai‘i. Right now, we see two major explanations that may be contributing to this pattern, and these influences cannot yet be fully separated from each other.

One possibility is that Pidgin and ‘ōlelo Hawai‘i share linguistic properties that allow Pidgin speakers to draw upon existing knowledge when encountering Hawaiian. Because Pidgin is a creole language that developed with influence from ‘ōlelo Hawai‘i, it shares more linguistic features with Hawaiian than English does, not only in individual words but also potentially in grammatical structures and sound patterns, such as which sounds are permitted to occur together and where sounds are likely to appear within words. This is similar to how an English speaker may find Spanish easier to learn than Mandarin Chinese because of greater linguistic similarities between English and Spanish. Likewise, speakers of Māori, Tahitian, Tongan, or Sāmoan may have some advantages when learning ‘ōlelo Hawai‘i because these languages share historical relationships as Polynesian languages. Although Pidgin is most closely related to English, its linguistic relationship with Hawaiian may similarly provide speakers with knowledge they can draw upon when engaging with Hawaiian.

If separated from the communities and experiences in which Pidgin exists, this finding might suggest that Pidgin itself provides an inherent linguistic advantage. However, when we consider Pidgin speakers within their lived experiences, we know that this is only one possible explanation.

Another likely explanation is that Pidgin speakers may have greater exposure to Hawaiian than do English speakers from Hawai‘i. Many Pidgin speakers are part of families and communities with longstanding connections to Hawai‘i and may have encountered ‘ōlelo Hawai‘i from childhood through friends, family members, community spaces, and the broader linguistic landscape around them. This may also help explain why English speakers from Hawai‘i demonstrate greater ‘ōlelo Hawai‘i knowledge compared to English speakers from outside of Hawai‘i.

I find this aspect of the findings particularly encouraging. It suggests that participation in Hawaiian language revitalization does not require someone to already be a speaker. By being present in community and engaging with spaces where ‘ōlelo Hawai‘i is spoken, people may be developing linguistic knowledge even before they begin formal learning. Their bodies and minds are responding to the linguistic environment around them. We know more Hawaiian than we realize.

At the same time, this does not mean that adults can become speakers of Hawaiian through exposure alone. Exposure can support learning, but it is not a replacement for the time, effort, and commitment required to learn a language. Experiment 3 demonstrates that ambient exposure can support the development of linguistic knowledge, but language learning still requires intentional engagement with vocabulary, grammar, and the broader cultural knowledge embedded within 'ōlelo Hawai'i. These findings do not mean that Pidgin benefits replace the knowledge gained through the guidance of 'ōlelo Hawai'i kumu who emphasize the importance of attending classes, learning through immersion, and putting in the time to becoming a speaker. Rather, the findings from Experiment 3 suggest that Pidgin knowledge and ambient exposure may provide a foundation or boost that supports later learning.

Research in the Aotearoa context has found that individuals who perform better on similar proto-lexical tasks also demonstrate stronger outcomes in Māori language classrooms. This provides a starting point for exploring whether similar relationships may exist in Hawai'i, while future research will help us better understand how these connections relate to 'ōlelo Hawai'i learning outcomes.

In this study, linguistic similarity and exposure cannot be fully disentangled. While future research may find ways to examine these factors separately, doing so may not fully capture the lived experiences of language in Hawai'i. For example, in reality, linguistic connection and exposure are not independent factors; they develop together through the communities and relationships in which languages are spoken. Pidgin speakers with greater exposure to Hawaiian may also speak a variety of Pidgin with great linguistic overlap with 'ōlelo Hawai'i. In this way, structural relationships between languages and lived experiences of language use are woven together and cannot be naturally separated.

There are also other ways of understanding why Pidgin may support 'ōlelo Hawai'i linguistic knowledge, with explanations sourced directly from the community. As discussed in the mo'olelo chapter, there are several reasons we might expect Pidgin to support 'ōlelo Hawai'i learning. If hearing Pidgin creates a sense of comfort, reduces hesitation, encourages learners to “waiho ka hilahila i ka hale” (leave one's bashfulness at the door; Puk [1993]), connects learners to a collective value system, or supports the practice of breathing life into one's genealogy and 'ōlelo Hawai'i learning journey, these attributes together may create conditions where Pidgin speakers are better supported in their learning of 'ōlelo Hawai'i.

This serves as a reminder that understanding the possible explanations for the Pidgin speaker advantage found in Experiment 3 requires more than

what can be learned through linguistics research alone. It requires looking at language holistically through the lens of the community and the speakers who use it. Future work must continue learning from the mana'o of community members to better understand how these findings can be shared and used in ways that contribute to our collective knowledge and benefit the lāhui.

Should we be encouraged by these findings found in Experiment 3? Absolutely.

But the reasons for this encouragement extend beyond a Pidgin advantage in linguistic knowledge of 'ōlelo Hawai'i. We can be encouraged by the people who have worked to ensure that 'ōlelo Hawai'i continues to be heard throughout Hawai'i. We can be encouraged by those who carried these languages forward despite histories of stigma, marginalization, and devaluation. We can be encouraged by the speakers of Pidgin and 'ōlelo Hawai'i today who continue to demand space for these languages to exist, thrive, and be spoken.

As we continue this work, may we remember to nānā i ke kumu (look to the source). In our search for knowledge and truth, we must continually return to our kūpuna, our communities, and our ancestral mo'olelo, for they are the foundation from which that knowledge continues to grow.

I am grateful for the opportunity, as a researcher and member of the community, to participate in this work and to experience firsthand the value of community and the expression of aloha woven throughout the efforts of so many people in Hawai'i.

Mahalo nui kākou.

Chapter 6

Conclusion

I return to the central question as we bring together the parts of this thesis. Do speakers leverage shared linguistic and sociocultural knowledge across Pidgin and ‘ōlelo Hawai‘i in psycholinguistic processing, and what these patterns reveal when contrasted with English? We situate this broad research question in the context of using Pidgin as an instructional medium for ‘ōlelo Hawai‘i. The answer is complicated.

In Experiment 1, I tested whether trilingual speakers of ‘ōlelo Hawai‘i, English, and Pidgin differ in their processing of Pidgin-to-Hawaiian and English-to-Hawaiian code switches, relative to each language’s unilingual baseline. The results provide suggestive evidence that Pidgin-to-‘ōlelo Hawai‘i switches may involve lower processing costs than English-to-‘ōlelo Hawai‘i switches. This pattern emerges in the follow-up one-tailed paired-samples t-test. However, the Bayesian mixed-effects model does not show a confident interaction between frame language and code-switching condition when predicting reaction time. Together, these results offer weak but suggestive evidence, in favor of Pidgin-to-‘ōlelo Hawai‘i switches reducing code-switching costs. While further work is required to draw strong conclusions, the findings raise important questions about how these three languages relate during real-time processing. In particular, they point to the need to consider the naturalness of different switch directions and whether laboratory settings might privilege the parsing of English rather than Pidgin.

Experiment 2 presents a cross-lingual priming paradigm that investigates whether Pidgin stimuli more successfully prime ‘ōlelo Hawai‘i words than English stimuli in a word identification task. No evidence was found to support the hypothesis that Pidgin primes facilitate processing of ‘ōlelo Hawai‘i targets more effectively than English primes; however, the model also did not show evidence of a priming effect for ‘ōlelo Hawai‘i target words overall. Neither English nor Pidgin primes facilitated processing of ‘ōlelo Hawai‘i target items, and thus these results neither support nor reject predictions concerning a possible Pidgin benefit in ‘ōlelo Hawai‘i priming. These results instead point to limitations inherent in the experimental design, particularly challenges in stimuli selection, task difficulty, and participant recruitment.

In Experiment 3, we find evidence that Pidgin speakers have a more developed ‘ōlelo Hawai‘i proto-lexicon compared to English speakers from Hawai‘i. However, English speakers from Hawai‘i still incur benefits compared to English speakers from Canada. We conclude that gradient differences in ‘ōlelo Hawai‘i proto-lexical development, whether shaped by differences in exposure to the language or by structural overlap between Pidgin and ‘ōlelo Hawai‘i, yield processing benefits for Pidgin speakers. English speakers from Hawai‘i likewise exhibit an intermediate level of proto-lexical development and a corresponding intermediate processing advantage. Pidgin speakers exhibited a larger ‘ōlelo Hawai‘i word form inventory and more robust phonotactic knowledge than English speakers from Hawai‘i. English speakers from Hawai‘i likewise showed an advantage, exhibiting more word form knowledge and stronger phonotactic sensitivity than English speakers from Canada.

Across the three experiments, the primary advantage seen for Pidgin speakers compared to English speakers was in their phonotactic knowledge. Evidence suggesting that Pidgin offers a stronger co-association with ‘ōlelo Hawai‘i than English for trilinguals is preliminary and requires further investigation. When Experiments 1 and 2 are examined in the context of Experiment 3, the cross-experiment relationships emerge more clearly. Although Experiment 3 differs substantially from Experiments 1 and 2, largely because the languages involved are spoken by trilinguals rather than distinct speaker groups, Experiment 3 shows that English speakers from Hawai‘i display advantages not seen among English speakers from Canada. In theory, this suggests that Hawai‘i English in a trilingual context may yield benefits that are not present for speakers without any Hawaiian exposure. The strict comparison of the co-association between Hawai‘i English and Pidgin with ‘ōlelo Hawai‘i in Experiments 1 and 2 is therefore likely more nuanced, as suggested by the broader patterns in Experiment 3.

Experiment 3 results indicate the benefits of both Pidgin and Hawai‘i English on ‘ōlelo Hawai‘i proto-lexical tasks, suggesting the facilitation in the processing of ‘ōlelo Hawai‘i. These processing benefits could potentially extend to learning benefits for ‘ōlelo Hawai‘i, as reported for individuals with a Māori proto-lexicon, who showed advantages in learning Māori (Mattingley et al. [2024b]). These findings suggest that features of both Hawai‘i English and Pidgin support the development of an ‘ōlelo Hawai‘i proto-lexicon with advantages over English speakers from Canada, although to different degrees. This effect may reflect relative differences in exposure to ‘ōlelo Hawai‘i, or possibly also due to gradient linguistic overlap that is featured across the three local languages, Hawai‘i English, Pidgin, and ‘ōlelo Hawai‘i.

However, it is also important to emphasize that Experiment 3 indeed reveals a robust and interpretable Pidgin-benefit. These results suggest that Pidgin speakers have richer representation of phonotactic knowledge and word-forms that more closely approximate the lexicon of an ambient learner than English speakers. Relating this to Experiments 1 and 2, Experiment 3 results suggest that the psycholinguistic benefit associated with Pidgin potentially arises less from dynamic facilitation during cross-lingual switching or priming, and more from the long-term accumulation of sub-lexical and lexical knowledge shaped by linguistic experience. In other words, our results overall indicate that the advantage of Pidgin is most clearly expressed in the organization and density of phonotactic knowledge.

The question arises: why do we observe a Pidgin benefit? Although the current findings do not allow us to identify the precise source of the benefit, several theoretically motivated explanations can be proposed. The first possibility lies in typological similarity. Pidgin overlaps substantially with ‘ōlelo Hawai‘i at both sub-lexical and lexical levels, potentially offering a pathway to linguistic knowledge that is not afforded by English.

Additionally, Pidgin identity is uniquely aligned with ‘ōlelo Hawai‘i identity, shaped both by historical circumstances and by the sociolinguistic landscape of contemporary Hawai‘i. Historically, the two languages have been tied to so-called *local* identity through their shared role in resisting haole exploitation of labor during the plantation era and in opposition of racist policies advanced by haole settlers during the early days of United States occupation. A person who self identifies as Pidgin-speaking may also hold attitudes that support a stronger valuing of ‘ōlelo Hawai‘i, potentially more so than a typical self-identifying English speaker from Hawai‘i. Stronger value and positive attitudes toward a language within a society have previously been shown to be empirically linked to a stronger proto-lexicon (Todd et al. [2023a]).

‘ōlelo Hawai‘i and Pidgin are spoken primarily by Kānaka ‘ōiwi, and their presence in the home increases the likelihood that individuals who have not formally studied Hawaiian, and who do not identify as Hawaiian speaking, still receive meaningful exposure. ‘ōlelo Hawai‘i is fundamental to Kānaka ‘ōiwi identity, thus even if a person does not identify as speaking ‘ōlelo Hawai‘i proficiently, their reverence for the language and their motivation to retain even ambiently learned ‘ōlelo Hawai‘i may support a more developed ‘ōlelo Hawai‘i proto-lexicon.

Pidgin speakerhood also includes individuals who are not Kānaka ‘ōiwi, many of whom are descendants of plantation-era families. Thus, regardless of Kānaka ‘ōiwi identity, such individuals largely come from house-

holds that have lived in Hawai'i across multiple generations. This suggests that Pidgin speakerhood may likewise involve intergenerational exposure to 'ōlelo Hawai'i. Even if an individual does not consider themselves Hawaiian-speaking, they may have received additional exposure in the home through Hawaiian-speaking relatives or through the increased use of Hawaiian words in everyday code-switching, rooted from an era of a more Hawaiian-dominant Hawai'i. Pidgin-speaking may thus entail increased home exposure to 'ōlelo Hawai'i.

In the broader contemporary community, such increased exposure for a given Pidgin speaker to 'ōlelo Hawai'i may arise through social networks, engagement in cultural practices such as hula or mele, and public events where 'ōlelo Hawai'i and Hawaiian cultural practices are prominently represented. A Pidgin speaker may also be more likely than an English speaker to reside in a community or region of Hawai'i where 'ōlelo Hawai'i is spoken at a higher rate in the surrounding neighborhood. Pidgin speakers may thus receive greater exposure by engaging in contexts where 'ōlelo Hawai'i is more salient in ambient input.

Overall, the Pidgin-advantage in the development of an 'ōlelo Hawai'i proto-lexicon likely reflects some combination of linguistic overlap between the two languages and increased opportunities for exposure. This benefit may also stem from other factors, such as stronger identification with Kānaka 'ōiwi identity or a heightened valuation of 'ōlelo Hawai'i.

As revealed through the methodology of mo'ōlelo, however, perhaps the pursuit in finding additional psycholinguistic advantages comes secondary to the community benefits of uplifting Pidgin speakers. In conversations with many members of the lāhui, including teachers, students, and community members in Hawai'i and the diaspora, I have come to understand that community members recognize Pidgin as intrinsically valuable, irrespective of whether it provides measurable linguistic advantages. Across history and in modern times, Pidgin became a language of both survival and friendship. For some, Pidgin evokes nostalgia for an older Hawai'i, and for others it is tied to feelings of home or to memories of family who have passed into the ao 'aumakua. For Pidgin speakers living in the diaspora, hearing Pidgin can evoke a sense of home and a feeling of returning to the land. As is common in the practices of genealogy and aloha 'āina, a language that allows us to call to our ancestors or to the land constitutes a spiritual practice that should be revered. At the same time, Pidgin enables playfulness and is often used in comedic contexts, or to soften the seriousness of speech. The language also serves as a reminder of values upheld in Hawai'i, including humility, collectivism over individualism, centering pilina, and the practice of aloha

‘āina and mālama ‘āina in multiple aspects of daily life.

As relayed by Hawaiian language teachers, the use of Pidgin in the classroom has been observed to help students feel more comfortable. There is something about the nature of Pidgin that brings many people in Hawai‘i a sense of peace and ease. Perhaps this reflects the connection to life at home, or it may be a calling to a broader collective purpose.

When I began this study, I knew it was important to me, personally, to conduct this work alongside members of my community in Hawai‘i. What I did not anticipate was the overwhelming care and support from others in backing this work. The community contributed to this work directly, in their participation in the experimental work or in their conversations with me, and most participants chose to donate their compensation back into the community, in support the ‘Aha Pūnana Leo Hawaiian immersion schools and the next generation of Hawaiian speakers. Witnessing this was profoundly moving for me as a scholar—to see my community offer support and encouragement, motivated not by personal gain, but by something far greater than themselves. Perhaps future research should consider offering participants the opportunity to contribute in this way. Regardless of the impact of the empirical results, one significant outcome was the donations to the schools, made possible by the selflessness of members of the lāhui, a moment I was deeply honored to witness as a scholar.

While Pidgin stigma remains present within the community, based on attitudes expressed in conversations with community members, what is perhaps more important than conducting empirical work is the willingness to engage in uncomfortable conversations. For me as a scholar, this process relied heavily on self-reflection and vulnerability, and in many ways was even more personally challenging than conducting the experimental studies. The broader implications of this work extend beyond the study’s results to include collaboration with others in challenging these notions and identifying them for what they are: remnants of a colonial mindset present within our community.

Expanding these findings beyond Hawai‘i, one way to consider them is through the lens of creole language benefit. Creole and mixed languages across the world deserve recognition and legitimacy in their own right. For individuals or communities that desire to learn related Indigenous language, the present findings raise the possibility that the languages they speak may offer advantages in phonotactic knowledge, rather than being seen as posing obstacles to language acquisition or revitalization efforts. Allowing languages to co-exist within the same linguistic landscape and be mutually recognized as legitimate may help challenge stigma associated with language mixing and

creolization, while also creating conditions under which acquisitional advantages may support engagement with the (re)learning of related Indigenous languages.

This work, unlike previous studies, demonstrated that gradient exposure is directly related to the development of a more robust proto lexicon. Research from Aotearoa (e.g., Oh et al. [2020], Panther et al. [2023b]) has shown that non-Māori speaking participants displayed phonotactic knowledge that was statistically equivalent to that of fluent Māori speakers. In contrast, our study, like work conducted in California and Texas on the Spanish protolexicon (Todd et al. [2023b]), did not find that self-identified non-speakers of ‘ōlelo Hawai‘i from Hawai‘i showed phonotactic knowledge at the same level as fluent speakers. We argue that this difference reflects the lower amount of ambient exposure to ‘ōlelo Hawai‘i compared to the high level of exposure to te reo Māori in Aotearoa. This provides some empirical evidence that increased exposure to the Indigenous language is associated with advantages for language revitalization efforts within the community. Such results further support increasing all public displays of a local Indigenous language so that learners can hear and see the language in their environment, with direct benefits for the development of their proto-lexicon. This includes expanding written language such as street signs and place names in the target language, as well as increasing spoken language in public advertisements, community events, and in everyday public interactions. Even small steps are beneficial, and increasing exposure, whether through a few words spoken at home or brief moments of hearing the language in daily life, can be tremendously impactful. These small acts accumulate and shape both individual learning (e.g., Mattingley et al. [2024a]) and the broader community (e.g., Oh et al. [2020]), lending to the pursuit of revitalization efforts both at local and global scales.

The findings from this study motivate an ideal model of speech perception that incorporates the role of exposure for individuals with proto-lexical knowledge. Greater exposure supports the development of both lexical and sub-lexical representations, which in turn introduces stronger competition at multiple levels. For speakers who rely on a proto-lexicon, lexical representations are comparatively fuzzy, so they may have weaker contextual support for resolving competition between sub-lexical and lexical candidates. For individuals with a less developed proto-lexicon, the benefit of sub-lexical phonotactic knowledge is non-discriminatory. In other words, the strategy utilized by such speakers applies similarly to words and nonwords, because the relatively less developed proto-lexicon offers less competition across multiple levels of representation.

This study began with a research question about the psycholinguistic benefits of Pidgin. However, I came to realize that this question itself may have been searching for legitimacy for both the language and, in a sense, for myself. It struck me that in the context of minoritized languages, there is often a focus on demonstrating their inherent value. No one has ever needed to justify speaking languages like English, so why do we feel compelled to search for benefits, advantages, or other forms of validation for Pidgin? Speakers of Pidgin deserve the right to use their language regardless of any measurable benefits or advantages it may provide in relation to learning Hawaiian or contributing to community values. Pidgin, in its own right, has intrinsic value.

Bibliography

- ‘Ōlelo No‘eau: *Hawaiian proverbs poetical sayings*. Bernice P. Bishop Museum special publication. Bishop Museum Press, Honolulu, Hawai‘i, reprint d. ausg. von 1983 edition, 1993. ISBN 9780910240932.
- David Abercrombie. *Elements of General Phonetics*. Edinburgh University Press, Edinburgh, 1967. ISBN 9780852244517. Reprinted 2022.
- Mark Amengual. Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, 15(3):517–530, July 2012. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728911000460. URL https://www.cambridge.org/core/product/identifier/S1366728911000460/type/journal_article.
- Sally Andrews. Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5):802–814, 1989. doi: 10.1037/0278-7393.15.5.802.
- Sally Andrews. Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2):234–254, 1992. doi: 10.1037/0278-7393.18.2.234.
- Alexander L. Anwyl-Irvine, Jessica Massonnié, Adam Flitton, Natasha Kirkham, and Jo K. Evershed. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1):388–407, 2020. doi: 10.3758/s13428-019-01237-x. URL <https://doi.org/10.3758/s13428-019-01237-x>.
- Xavier Aparicio and Jean-Marc Lavour. Recognising words in three languages: effects of language dominance and language switching. *International Journal of Multilingualism*, 11(2):164–181, April 2014. ISSN 1479-0718, 1747-7530. doi: 10.1080/14790718.2013.783583. URL <http://www.tandfonline.com/doi/abs/10.1080/14790718.2013.783583>.

- S. L. Archer and S. Curtin. Nine-month-olds use frequency of onset clusters to segment novel words. *Journal of Experimental Child Psychology*, 148: 131–141, 2016. doi: 10.1016/j.jecp.2016.04.004. URL <https://doi.org/10.1016/j.jecp.2016.04.004>.
- Ruben F. Asensio. Language policy in the kingdom of hawai‘i: A worldly english approach. *Second Language Studies*, 28(2):1–48, 2010. URL <https://www.hawaii.edu/sls/wp-content/uploads/2014/09/Asensio.pdf>.
- Molly Babel and Benjamin Munson. Producing socially meaningful linguistic variation. In *The Oxford Handbook of Sociolinguistics*. Oxford University Press, 2013a. doi: 10.1093/oxfordhb/9780199735471.013.022. URL <https://academic.oup.com/edited-volume/28004/chapter/211763283>.
- Molly Babel and Benjamin Munson. Producing socially meaningful linguistic variation. In *The Oxford Handbook of Sociolinguistics*. Oxford University Press, December 2013b. doi: 10.1093/oxfordhb/9780199735471.013.022. URL <https://academic.oup.com/edited-volume/28004/chapter/211763283>.
- Todd M. Bailey and Ulrike Hahn. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44(4):568–591, May 2001. ISSN 0749596X. doi: 10.1006/jmla.2000.2756. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X00927566>.
- Tammy Haili‘ōpua Baker. Pidgin on stage: Language, identity, and performance in hawai‘i. *Theatre Journal*, 72(3):341–358, 2020.
- David A. Balota, Michael J. Cortese, Shari D. Sergent-Marshall, Daniel H. Spieler, and Melvin J. Yap. Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133(2):283–316, 2004. doi: 10.1037/0096-3445.133.2.283.
- David A. Balota, Melvin J. Yap, Keith A. Hutchison, Michael J. Cortese, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson, and Rebecca Treiman. The english lexicon project. *Behavior Research Methods*, 39(3):445–459, August 2007a. ISSN 1554-351X, 1554-3528. doi: 10.3758/BF03193014. URL <http://link.springer.com/10.3758/BF03193014>.
- David A Balota, Melvin J Yap, Keith A Hutchison, Michael J Cortese, Brett Kessler, Bjorn Loftis, James H Neely, Douglas L Nelson, Greg B Simpson,

and Rebecca Treiman. The english lexicon project. *Behavior research methods*, 39:445–459, 2007b.

Dale J. Barr. Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4):457–474, November 2008. ISSN 0749-596X. doi: 10.1016/j.jml.2007.09.002. URL <https://www.sciencedirect.com/science/article/pii/S0749596X07001015>.

Dana M. Basnight-Brown and Jeanette Altarriba. Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory Cognition*, 35(5):953–965, July 2007. ISSN 0090-502X, 1532-5946. doi: 10.3758/BF03193468. URL <http://link.springer.com/10.3758/BF03193468>.

Cécile Beauvillain and Jonathan Grainger. Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, 26(6):658–672, December 1987. ISSN 0749596X. doi: 10.1016/0749-596X(87)90108-2. URL <https://linkinghub.elsevier.com/retrieve/pii/0749596X87901082>.

Catherine Becker. *Mana Cards: The Power of Hawaiian Wisdom*. Radiance Network Inc, 1998. ISBN 978-0966014204.

Elika Bergelson and Daniel Swingley. At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9):3253–3258, February 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1113380109. URL <https://pnas.org/doi/full/10.1073/pnas.1113380109>.

Kinsey Bice and Judith F. Kroll. English only? monolinguals in linguistically diverse contexts have an edge in language learning. *Brain and Language*, 196:104644, September 2019. ISSN 0093934X. doi: 10.1016/j.bandl.2019.104644. URL <https://linkinghub.elsevier.com/retrieve/pii/S0093934X18303274>.

Derek Bickerton. The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7(2):173–188, 1984. ISSN 0140-525X, 1469-1825. doi: 10.1017/S0140525X00044149. URL https://www.cambridge.org/core/product/identifier/S0140525X00044149/type/journal_article.

Derek Bickerton and Carol Odo. *General Phonology and Pidgin Syntax*. University of Hawai’i, Honolulu, 1976.

- Marie-Josée Bisson, Walter J. B. van Heuven, Kathy Conklin, and Richard J. Tunney. Incidental acquisition of foreign language vocabulary through brief multi-modal exposure. *PLOS ONE*, 8(4):e60912, 2013. doi: 10.1371/journal.pone.0060912. URL <https://doi.org/10.1371/journal.pone.0060912>.
- Esti Blanco-Elorrieta and Liina Pykkänen. Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *The Journal of Neuroscience*, 37(37):9022–9036, September 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0553-17.2017. URL <https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0553-17.2017>.
- Heather Bortfeld, James L. Morgan, Roberta Michnick Golinkoff, and Karen Rathbun. Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16(4):298–304, April 2005. ISSN 0956-7976, 1467-9280. doi: 10.1111/j.0956-7976.2005.01531.x. URL <https://journals.sagepub.com/doi/10.1111/j.0956-7976.2005.01531.x>.
- Roger W. Brown and Donald C. Hildum. Expectancy and the perception of syllables. *Language*, 32(3):411, July 1956. ISSN 00978507. doi: 10.2307/410561. URL <https://www.jstor.org/stable/410561?origin=crossref>.
- Marc Brysbaert, Geert Van Dyck, and Maarten Van de Poel. Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1):137–148, 1999. doi: 10.1037/0096-1523.25.1.137.
- Marc Brysbaert, Michaël Stevens, Paweł Mandera, and Emmanuel Keuleers. How many words do we know? practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant’s age. *Frontiers in psychology*, 7:1116, 2016.
- Marc Brysbaert, Evelyne Lagrou, and Michaël Stevens. Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, 20(3):530–548, May 2017. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728916000353. URL https://www.cambridge.org/core/product/identifier/S1366728916000353/type/journal_article.

- Marc Brysbaert, Paweł Mandera, and Emmanuel Keuleers. The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27(1):45–50, February 2018. ISSN 0963-7214, 1467-8721. doi: 10.1177/0963721417727521. URL <https://journals.sagepub.com/doi/10.1177/0963721417727521>.
- Dean V. Buonomano, Jennifer Bramen, and Mahsa Khodadadifar. Influence of the interstimulus interval on temporal processing and learning: testing the state-dependent network model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1525):1865–1873, July 2009. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2009.0019. URL <https://royalsocietypublishing.org/doi/10.1098/rstb.2009.0019>.
- Wednesday Bushong and T. Florian Jaeger. Dynamic re-weighting of acoustic and contextual cues in spoken word recognition. *The Journal of the Acoustical Society of America*, 146(2):EL135–EL140, August 2019. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.5119271. URL <https://pubs.aip.org/jasa/article/146/2/EL135/663694/Dynamic-re-weighting-of-acoustic-and-contextual>.
- Krista Byers-Heinlein, Tracey C. Burns, and Janet F. Werker. The roots of bilingualism in newborns. *Psychological Science*, 21:343–348, 2010.
- Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.
- Jason Iota Cabral. E ola ka ‘Ōlelo pa‘i‘ai. Conference Presentation at the Hawaiian-Pidgin Summit, Honolulu, HI, United States, October 25 2018. URL <https://www.hawaii.edu/satocenter/?p=1713>.
- Kathryn Campbell-Kibler. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change*, 22(3):423–441, 2011. doi: 10.1017/S0954394510000177.
- Alfonso Caramazza and Isabel Brones. Semantic classification by bilinguals. *Canadian Journal of Psychology / Revue canadienne de psychologie*, 34(1):77–81, March 1980. ISSN 0008-4255. doi: 10.1037/h0081016. URL <https://doi.apa.org/doi/10.1037/h0081016>.
- Elizabeth Ball Carr. *Da Kine Talk: From Pidgin to Standard English in Hawaii*. University of Hawaii Press, March 2019. ISBN 9780824881245.

doi: 10.2307/j.ctv9zckfd. URL <http://www.jstor.org/stable/10.2307/j.ctv9zckfd>.

Bharath Chandrasekaran, Padma D. Sampath, and Patrick C. M. Wong. Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, 128(1):456–465, 2010a. doi: 10.1121/1.3445785.

Bharath Chandrasekaran, Padma D. Sampath, and Patrick C.M. Wong. Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, 128(1):456–465, July 2010b. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.3445785. URL <https://pubs.aip.org/jasa/article/128/1/456/656030/Individual-variability-in-cue-weighting-and>.

Adel Chaouch-Orozco, Jorge González Alonso, and Jason Rothman. Individual differences in bilingual word recognition: The role of experiential factors and word frequency in cross-language lexical priming. *Applied Psycholinguistics*, 42(2):447–474, March 2021. ISSN 0142-7164, 1469-1817. doi: 10.1017/S014271642000082X. URL https://www.cambridge.org/core/product/identifier/S014271642000082X/type/journal_article.

Hsuan-Chih Chen and Man-Lai Ng. Semantic facilitation and translation priming effects in chinese-english bilinguals. *Memory Cognition*, 17(4):454–462, July 1989. ISSN 0090-502X, 1532-5946. doi: 10.3758/BF03202618. URL <http://link.springer.com/10.3758/BF03202618>.

Yu-Lin Cheng and David Howard. The time cost of mixed-language processing: an investigation. *International Journal of Bilingualism*, 12(3):209–222, September 2008. ISSN 1367-0069, 1756-6878. doi: 10.1177/1367006908098568. URL <https://journals.sagepub.com/doi/10.1177/1367006908098568>.

Steven J. Choy and David H. Dodd. Standard and nonstandard hawaiian english-speaking children: Comprehension of both dialects and teacher’s evaluations. *Journal of Educational Psychology*, 68(2):184–193, April 1976. ISSN 1939-2176, 0022-0663. doi: 10.1037/0022-0663.68.2.184. URL <https://doi.apa.org/doi/10.1037/0022-0663.68.2.184>.

Cynthia G. Clopper and Ann R. Bradlow. Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*,

51(3):175–198, September 2008. ISSN 0023-8309, 1756-6053. doi: 10.1177/0023830908098539. URL <https://journals.sagepub.com/doi/10.1177/0023830908098539>.

Cynthia G. Clopper and Abby Walker. Effects of lexical competition and dialect exposure on phonological priming. *Language and Speech*, 60(1): 85–109, 2017. doi: 10.1177/0023830916665585.

Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Hillsdale, NJ, 2nd edition, 1988.

John S. Coleman and Janet Pierrehumbert. Stochastic phonological grammars and acceptability. In *Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology*, pages 49–56, Somerset, NJ, 1997. Association for Computational Linguistics.

Allan M. Collins and Elizabeth F. Loftus. A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407–428, November 1975. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.82.6.407. URL <https://doi.apa.org/doi/10.1037/0033-295X.82.6.407>.

Rebecca J. Compton, Elizabeth Heaton, and Emily Ozer. Intertrial interval duration affects error monitoring. *Psychophysiology*, 54(8):1151–1162, August 2017. ISSN 0048-5772, 1469-8986. doi: 10.1111/psyp.12877. URL <https://onlinelibrary.wiley.com/doi/10.1111/psyp.12877>.

Svetlana V. Cook and Kira Gor. Lexical access in l2: Representational deficit or processing constraint? *The Mental Lexicon*, 10(2):247–270, September 2015. ISSN 1871-1340, 1871-1375. doi: 10.1075/ml.10.2.04coo. URL <http://www.jbe-platform.com/content/journals/10.1075/ml.10.2.04coo>.

Svetlana V. Cook, Nick B. Pandža, Alia K. Lancaster, and Kira Gor. Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology*, 7, September 2016. ISSN 1664-1078. doi: 10.3389/fpsyg.2016.01345. URL <http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.01345/abstract>.

Vivian Cook. *Effects of the Second Language on the First*. Multilingual Matters, Clevedon, 2003.

Albert Costa and Mikel Santesteban. Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and l2 learners. *Journal of Memory and Language*, 50(4):491–511,

May 2004. ISSN 0749596X. doi: 10.1016/j.jml.2004.02.002. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X04000191>.

Albert Costa, Alfonso Caramazza, and Nuria Sebastian-Galles. The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5):1283–1296, 2000. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.26.5.1283. URL <https://doi.apa.org/doi/10.1037/0278-7393.26.5.1283>.

Albert Costa, Mikel Santesteban, and Iva Ivanova. How do highly proficient bilinguals control their lexicalization process? inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5):1057–1074, 2006. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.32.5.1057. URL <https://doi.apa.org/doi/10.1037/0278-7393.32.5.1057>.

Suzanne Curtin and Stephanie L. Archer. *Speech perception*, page 137–158. Cambridge University Press, 2 edition, November 2015. ISBN 9781316095829. doi: 10.1017/CBO9781316095829.007. URL https://www.cambridge.org/core/product/identifier/9781316095829%23CN-bp-7/type/book_part.

Da Pidgin Coup. Pidgin and education: A position paper. *Educational Perspectives*, 41(1–2):30–39, 2008. URL <https://coe.hawaii.edu/wp-content/uploads/2020/04/Vol141-1-2.pdf>. Position paper by Da Pidgin Coup.

Robert Daland, Bruce Hayes, James White, Marc Garellek, Andrea Davis, and Ingrid Norrmann. Explaining sonority projection effects. *Phonology*, 28(2):197–234, 2011. doi: 10.1017/S0952675711000145.

Richard R. Day. The development of linguistic attitudes and preferences. *TESOL Quarterly*, 14(1):27–37, 1980.

Annette M. B de Groot and Gerard L. J Nas. Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30(1):90–123, February 1991. ISSN 0749-596X. doi: 10.1016/0749-596X(91)90012-9. URL <https://www.sciencedirect.com/science/article/pii/0749596X91900129>.

- David DeCamp. Toward a generative analysis of a post-creole speech continuum. In Dell Hymes, editor, *Pidginization and Creolization of Languages*, pages 349–370. Cambridge University Press, 1971.
- Kevin Diependaele, Kristin Lemhöfer, and Marc Brysbaert. The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, 66(5):843–863, May 2013. ISSN 1747-0218, 1747-0226. doi: 10.1080/17470218.2012.720994. URL <https://journals.sagepub.com/doi/10.1080/17470218.2012.720994>.
- Ton Dijkstra and Walter J. B. van Heuven. The bia model and bilingual word recognition. *Bilingualism: Language and Cognition*, 1(1):51–66, 1998. doi: 10.1017/S1366728998000121.
- Ton Dijkstra and Walter J. B. van Heuven. The architecture of the bilingual word recognition system: From identification to decision. In Jonathan Grainger and Ton Dijkstra, editors, *Bilingualism: Language and Cognition*, pages 175–197. Oxford University Press, 2002.
- Ton Dijkstra, Jonathan Grainger, and Walter J. B. van Heuven. Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41(4):496–518, 1999. doi: 10.1006/jmla.1999.2654.
- Maria Dimitropoulou, Jon Andoni Duñabeitia, and Manuel Carreiras. Masked translation priming effects with low proficient bilinguals. *Memory Cognition*, 39(2):260–275, February 2011. ISSN 0090-502X, 1532-5946. doi: 10.3758/s13421-010-0004-9. URL <http://link.springer.com/10.3758/s13421-010-0004-9>.
- Liam Doherty. [dohliam/hawaiian-corpus](https://github.com/dohliam/hawaiian-corpus), June 2024. URL <https://github.com/dohliam/hawaiian-corpus>.
- Katie Drager. Speaker age and vowel perception. *Language and Speech*, 54(1):99–121, March 2011. ISSN 0023-8309, 1756-6053. doi: 10.1177/0023830910388017. URL <https://journals.sagepub.com/doi/10.1177/0023830910388017>.
- Katie Drager. Pidgin and hawai‘i english: An overview. *International Journal of Language, Translation and Intercultural Communication*, 1(1):61–73, 2012. doi: 10.12681/ijltic.10.

- Janet M. Duchek and James H. Neely. A dissociative word-frequency x levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory Cognition*, 17(2):148–162, March 1989. ISSN 0090-502X, 1532-5946. doi: 10.3758/BF03197065. URL <http://link.springer.com/10.3758/BF03197065>.
- Wouter Duyck. Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6):1340–1359, 2005. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.31.6.1340. URL <https://doi.apa.org/doi/10.1037/0278-7393.31.6.1340>.
- Jon Andoni Duñabeitia, Manuel Perea, and Manuel Carreiras. Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, 57(2):98–107, November 2010. ISSN 1618-3169, 2190-5142. doi: 10.1027/1618-3169/a000013. URL <https://econtent.hogrefe.com/doi/10.1027/1618-3169/a000013>.
- Jan Edwards, Mary E. Beckman, and Benjamin Munson. The interaction between vocabulary size and phonotactic probability effects on children’s production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, 47:421–436, 2004. doi: 10.1044/1092-4388(2004/034). URL [https://doi.org/10.1044/1092-4388\(2004/034\)](https://doi.org/10.1044/1092-4388(2004/034)).
- Samuel H. Elbert and Mary Kawena Pukui. *Hawaiian Grammar*. University Press of Hawai’i, Honolulu, 1979. URL https://books.google.com/books/about/Hawaiian_Grammar.html?id=I2UOAAAAYAAJ.
- Paola Escudero. Linguistic perception of similar l2 sounds. In Paul Boersma and Silke Hamann, editors, *Phonology in Perception*, pages 151–190. Mouton de Gruyter, 2009.
- Paola Escudero, Rachel Hayes-Harb, and Holger Mitterer. Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2):345–360, April 2008. ISSN 00954470. doi: 10.1016/j.wocn.2007.11.002. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447007000502>.
- Larry Fenson, Elizabeth Bates, Philip Dale, Judith Goodman, J. Steven Reznick, and Donna Thal. Measuring variability in early child language:

Don't shoot the messenger. *Child Development*, 71:323–328, 2000. doi: 10.1111/1467-8624.00147.

Daniel Fogerty. Indexical properties influence time-varying amplitude and fundamental frequency contributions of vowels to sentence intelligibility. *Journal of Phonetics*, 52:89–104, September 2015. ISSN 00954470. doi: 10.1016/j.wocn.2015.06.005. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447015000479>.

Kenneth I. Forster and Susan M. Chambers. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12(6):627–635, December 1973. ISSN 00225371. doi: 10.1016/S0022-5371(73)80042-8. URL <https://linkinghub.elsevier.com/retrieve/pii/S0022537173800428>.

Alexander L. Francis, Katherine Baldwin, and Howard C. Nusbaum. Effects of training on attention to acoustic cues. *Perception & Psychophysics*, 62(8):1668–1680, 2000. doi: 10.3758/BF03212164.

Wendy S. Francis, Natasha Tokowicz, and Judith F. Kroll. The consequences of language proficiency and difficulty of lexical access for translation performance and priming. *Memory Cognition*, 42(1):27–40, January 2014. ISSN 0090-502X, 1532-5946. doi: 10.3758/s13421-013-0338-1. URL <https://link.springer.com/10.3758/s13421-013-0338-1>.

Michael C. Frank, Joshua B. Tenenbaum, and Edward Gibson. Learning and long-term retention of large-scale artificial languages. *PLoS ONE*, 8(1):e52500, January 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0052500. URL <https://dx.plos.org/10.1371/journal.pone.0052500>.

Melinda Fricke, Judith F Kroll, and Paola E Dussias. Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of memory and language*, 89:110–137, 2016.

Stefan A. Frisch, Nathan R. Large, and David B. Pisoni. Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, 42(4):481–496, May 2000. ISSN 0749596X. doi: 10.1006/jmla.1999.2692. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X9992692X>.

Stefan A. Frisch, Nathan R. Large, Bushra Adnan Zawaydeh, and David B. Pisoni. *Emergent phonotactic generalizations in English and Arabic*, volume 45, page 159–180. John Benjamins Publishing Company, Amster-

dam, October 2001. ISBN 9789027229472. doi: 10.1075/tsl.45.09fri. URL <https://benjamins.com/catalog/tsl.45.09fri>.

R. Frost, A. Deutsch, O. Gilboa, and K. I. Forster. Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition*, 28(8):1277–1288, 2000. doi: 10.3758/BF03211828.

Candace Fujikane. Reimagining development and the local in lois-ann yamanaka’s saturday night at the pahala theatre. *Social Process in Hawaii*, 38:40–61, 1997. Special issue: Women in Hawaii: Sites, Identities, and Voices, issn = 0737-6871.

Gavin Furukawa. Stylization and language ideologies in pidgin comedic skits. *Discourse, Context Media*, 23:41–52, 2018. ISSN 22116958. doi: 10.1016/j.dcm.2017.06.006. URL <https://linkinghub.elsevier.com/retrieve/pii/S2211695816302197>.

Andrew Gelman, Daniel Simpson, and Michael Betancourt. The prior can often only be understood in the context of the likelihood. *Entropy*, 19, 2017. doi: 10.3390/e19100555. URL <https://www.mdpi.com/1099-4300/19/10/555>.

Stephen D. Goldinger. Echoes of echoes? an episodic theory of lexical access. *Psychological Review*, 105(2):251–279, 1998. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.105.2.251. URL <https://doi.apa.org/doi/10.1037/0033-295X.105.2.251>.

Matthew Goldrick, Elin Runnqvist, and Albert Costa. Language switching makes pronunciation less nativelike. *Psychological Science*, 25(4):1031–1036, April 2014. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797613520014. URL <https://journals.sagepub.com/doi/10.1177/0956797613520014>.

Tamar H. Gollan and Victor S. Ferreira. Should i stay or should i switch? a cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3):640–665, 2009. ISSN 1939-1285, 0278-7393. doi: 10.1037/a0014981. URL <https://doi.apa.org/doi/10.1037/a0014981>.

Tamar H. Gollan, Kenneth I. Forster, and Ram Frost. Translation priming with different scripts: Masked priming with cognates and noncognates in hebrew–english bilinguals. *Journal of Experimental Psychology: Learning*,

Memory, and Cognition, 23(5):1122–1139, 1997. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.23.5.1122. URL <https://doi.apa.org/doi/10.1037/0278-7393.23.5.1122>.

W. Goodrich Smith and C. L. Hudson Kam. Knowing ‘who she is’ based on ‘where she is’: The effect of co-speech gesture on pronoun comprehension. *Language and Cognition*, 4:75–98, 2012. doi: 10.1515/langcog-2012-0005.

Noelani Goodyear-Ka‘ōpua. *The Seeds We Planted: Portraits of a Native Hawaiian Charter School*. First Peoples: New Directions in Indigenous Studies. University of Minnesota Press, Minneapolis, 2013. ISBN 9780816680481.

Kira Gor, Svetlana Cook, Denisa Bordag, Anna Chrabaszcz, and Andreas Opitz. Fuzzy lexical representations in adult second language speakers. *Frontiers in Psychology*, 12:732030, November 2021. ISSN 1664-1078. doi: 10.3389/fpsyg.2021.732030. URL <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.732030/full>.

Katharine Graf Estes, Julia Evans, Martha Alibali, and Jenny R. Saffran. Can infants map meaning to newly segmented words? statistical segmentation and word learning. *Psychological Science*, 18(3):254–260, 2007. doi: 10.1111/j.1467-9280.2007.01885.x. URL <https://doi.org/10.1111/j.1467-9280.2007.01885.x>.

Katharine Graf Estes, Stephanie C.-W. Gluck, and Kevin J. Grimm. Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology*, 146:34–49, 2016. doi: 10.1016/j.jecp.2016.01.012. URL <https://doi.org/10.1016/j.jecp.2016.01.012>.

J. Grainger and C. Beauvillain. Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology Section A*, 39(2):295–319, May 1987. ISSN 0272-4987, 1464-0740. doi: 10.1080/14640748708401788. URL <https://journals.sagepub.com/doi/10.1080/14640748708401788>.

James Grama. Substrate effects in the phonetic realization of pidgin vowels. *Journal of the Acoustical Society of America*, 140:3052–3052, 2016. doi: 10.1121/1.4969484.

- James Grama, Michelle Kamigaki-Baron, and Katie Drager. English in hawai'i. In Andy Kirkpatrick, editor, *The Oxford Handbook of World Englishes*. Oxford University Press, 2022.
- James M Grama. *Variation and change in Hawaii Creole vowels*. PhD thesis, University of Hawai'i at Manoa, 2015.
- David W. Green and Jubin Abutalebi. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5):515–530, August 2013. ISSN 2044-5911, 2044-592X. doi: 10.1080/20445911.2013.796377. URL <http://www.tandfonline.com/doi/abs/10.1080/20445911.2013.796377>.
- François Grosjean. Neurolinguists, beware! the bilingual is not two monolinguals in one person. *Brain and Language*, 36:3–15, 1989.
- François Grosjean. The bilingual individual. *Interpreting. International Journal of Research and Practice in Interpreting*, 2(1–2):163–187, January 1997a. ISSN 1384-6647, 1569-982X. doi: 10.1075/intp.2.1-2.07gro. URL <http://www.jbe-platform.com/content/journals/10.1075/intp.2.1-2.07gro>.
- François Grosjean. The bilingual's language modes. In Janet Nicol, editor, *One Mind, Two Languages: Bilingual Language Processing*, pages 1–22. Blackwell, Oxford, 2001.
- François Grosjean. *Processing Mixed Language: Issues, Findings, and Models*. Psychology Press, 1997b. ISBN 9781315806051.
- Marianne Gullberg, Leah Roberts, Christine Dimroth, Kim Veroude, and Peter Indefrey. Adult language learning after minimal exposure to an unknown natural language. *Language Learning*, 60(s2):5–24, December 2010. ISSN 0023-8333, 1467-9922. doi: 10.1111/j.1467-9922.2010.00598.x. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2010.00598.x>.
- Lauren Hall-Lew, Emma Moore, and Robert J. Podesva. Social meaning and linguistic variation: Theoretical foundations. In Lauren Hall-Lew, Emma Moore, and Robert J. Podesva, editors, *Social Meaning and Linguistic Variation: Theorizing the Third Wave*, pages 1–24. Cambridge University Press, 2021. doi: 10.1017/9781108578684.001.

- Eunice Hargrove and Kent Sakoda. The hegemony of english or *hau kam yu wen kawl wat ai spik ingglish wen yu no no waz*. *Bamboo Ridge*, 75: 48–70, 1999.
- Robert J. Hartsuiker and Martin J. Pickering. Language integration in bilingual sentence production. *Acta Psychologica*, 128(3):479–489, July 2008. ISSN 00016918. doi: 10.1016/j.actpsy.2007.08.005. URL <https://linkinghub.elsevier.com/retrieve/pii/S0001691807000820>.
- Hawaiian Bible Project. Ka baibala hemolele (hawaiian bible). <https://baibala.org>, 2025. Accessed: 2025-10-23.
- Hawai'i Leads Country. Hawai'i leads country in non-english speakers. *Honolulu Star-Advertiser*, May 2016. URL <https://www.staradvertiser.com/2016/05/23/breaking-news/hawaii-leads-country-in-non-english-speakers/>.
- Hawai'i State Data Center Department of Business, Economic Development and Tourism. Detailed languages spoken at home in the state of hawaii: March 2016. Technical report, Research and Economic Analysis Division, Honolulu, HI, 2016. URL https://files.hawaii.gov/dbedt/census/acs/Report/Detailed_Language_March2016.pdf. Based on 2009 to 2013 American Community Survey 5-year estimates.
- Jennifer Hay, Janet Pierrehumbert, and Mary Beckman. Speech perception, well-formedness and the statistics of the lexicon. In John Local, Richard Ogden, and Rosalind Temple, editors, *Papers in Laboratory Phonology VI*, pages 58–74. Cambridge University Press, Cambridge, 2004.
- Jennifer Hay, Aaron Nolan, and Katie Drager. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, 23(3), January 2006. ISSN 0167-6318, 1613-3676. doi: 10.1515/TLR.2006.014. URL <https://www.degruyter.com/document/doi/10.1515/TLR.2006.014/html>.
- Jessica F. Hay, Bruna Pelucchi, Katharine Graf Estes, and Jenny R. Safran. Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, 63(2):93–106, 2011. doi: 10.1016/j.cogpsych.2011.06.002. URL <https://doi.org/10.1016/j.cogpsych.2011.06.002>.
- Bruce Hayes and Colin Wilson. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379–440, 2008.

- Kenneth Heafield. Kenlm: Faster and smaller language model queries. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 187–197, Edinburgh, Scotland, 2011. Association for Computational Linguistics. URL <https://aclanthology.org/W11-2123/>.
- Christina Higgins. Raising critical language awareness in hawai'i at da pidgin coup. In Bettina Migge et al., editors, *Creoles in Education: A Critical Assessment and Comparison of Existing Projects*. John Benjamins, Amsterdam, 2010.
- Anita Hofschneider. Most uh students are people of color but their teachers are mainly white, October 2019. URL <https://www.civilbeat.org/2019/10/most-uh-students-are-people-of-color-but-their-teachers-are-mainly-white/>.
- John Holm. *An Introduction to Pidgins and Creoles*. Cambridge University Press, 1 edition, August 2000. ISBN 9780521584609. doi: 10.1017/CBO9781139164153. URL <https://www.cambridge.org/core/product/identifier/9781139164153/type/book>.
- Lori L. Holt and Andrew J. Lotto. Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119(5):3059–3071, 2006. doi: 10.1121/1.2188377.
- Keith A. Hutchison. Is semantic priming due to association strength or feature overlap? a microanalytic review. *Psychonomic Bulletin Review*, 10(4):785–813, December 2003. ISSN 1069-9384, 1531-5320. doi: 10.3758/BF03196544. URL <http://link.springer.com/10.3758/BF03196544>.
- J. C. L. Ingram and S. G. Park. Cross language vowel perception and production by japanese and korean learners of english. *Journal of Phonetics*, 25:343–370, 1997.
- April Jacobs, Melinda Fricke, and Judith F. Kroll. Cross-language activation begins during speech planning and extends into second language speech. *Language Learning*, 66(2):324–353, April 2016. ISSN 0023-8333, 1467-9922. doi: 10.1111/lang.12148. URL <https://onlinelibrary.wiley.com/doi/10.1111/lang.12148>.
- Gerhard Jäger. Maximum entropy models and stochastic optimality theory. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw,

Joan Maling, and Chris Manning, editors, *Architectures, Rules, and Preferences: Variation on Themes by Joan Bresnan*, pages 467–479. CSLI Publications, Stanford, CA, 2007.

N Jiang. Lexical representation and development in a second language. *Applied Linguistics*, 21(1):47–77, March 2000. ISSN 0142-6001, 1477-450X. doi: 10.1093/applin/21.1.47. URL <https://academic.oup.com/applij/article-lookup/doi/10.1093/applin/21.1.47>.

Nan Jiang. Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, 2(1):59–75, April 1999. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728999000152. URL https://www.cambridge.org/core/product/identifier/S1366728999000152/type/journal_article.

Nan Jiang and Kenneth I. Forster. Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, 44(1):32–51, January 2001. ISSN 0749596X. doi: 10.1006/jmla.2000.2737. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X00927372>.

Elizabeth K. Johnson. Constructing a proto-lexicon: An integrative view of infant language development. *Annual Review of Linguistics*, 2:391–412, 2016. doi: 10.1146/annurev-linguistics-011415-040616. URL <https://doi.org/10.1146/annurev-linguistics-011415-040616>.

Elizabeth K. Johnson, Amanda Seidl, and Michael D. Tyler. The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLoS ONE*, 9(1):e83546, January 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0083546. URL <https://dx.plos.org/10.1371/journal.pone.0083546>.

Keith Johnson and Matthias Sjerps. Speaker normalization in speech perception. *UC Berkeley Phonology Lab Annual Reports*, 14, 2018. ISSN 2768-5047. doi: 10.5070/P7141042474. URL <https://escholarship.org/uc/item/2fc6x1ph>.

Keith Johnson, Elizabeth A Strand, and Mariapaola D’Imperio. Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, 27(4):359–384, October 1999. ISSN 00954470. doi: 10.1006/jpho.1999.0100. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447099901006>.

- Kristin Midgley Jonathan Grainger and Phillip J. Holcomb. Re-thinking the bilingual interactive-activation model from a developmental perspective (bia-d). *Language Acquisition*, 17(3):250–267, 2010. doi: 10.1075/lald.52.18gra.
- Peter W. Jusczyk and Richard N. Aslin. Infants’ detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29(1):1–23, 1995. doi: 10.1006/cogp.1995.1010.
- Peter W. Jusczyk, Angela D. Friederici, Jeanne M. I. Wessels, Vibeke Y. Svenkerud, and Anne M. Jusczyk. Infants’ sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32(3):402–420, 1993. doi: 10.1006/jmla.1993.1022.
- Peter W. Jusczyk, Paul A. Luce, and Charles-Luce Jan. Infants’ sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33(5):630–645, 1994. doi: 10.1006/jmla.1994.1030. URL <https://doi.org/10.1006/jmla.1994.1030>.
- Peter W. Jusczyk, Elizabeth A. Hohne, and Angela Bauman. Infants’ sensitivity to allophonic cues for word segmentation. *Perception Psychophysics*, 61(8):1465–1476, December 1999. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03213111. URL <http://link.springer.com/10.3758/BF03213111>.
- Kahakua et al. v. Hallgren. Kahakua et al. v. hallgren, 1987. No. 86-0434, U.S. District Court for the District of Hawai’i. Cited in Matsuda (1991). No published opinion or summary available.
- K. Kamanā and W. H. Wilson. *Nā Kai ‘Ewalu: Beginning Hawaiian Lessons*. Hale Kuamo’o, 2012.
- E. K. Kauhi. *He Mo’olelo No Kapa’ahu: The Story of Kapa’ahu*. Pili Productions, 2000.
- Margarita Kaushanskaya and Viorica Marian. Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, 57(1):119–163, March 2007. ISSN 0023-8333, 1467-9922. doi: 10.1111/j.1467-9922.2007.00401.x. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9922.2007.00401.x>.
- Margarita Kaushanskaya, Jeewon Yoo, and Stephanie Van Hecke. Word learning in adults with second-language experience: Effects of phonolog-

ical and referent familiarity. *Journal of Speech, Language, and Hearing Research*, 56(2):667–678, April 2013. ISSN 1092-4388, 1558-9102. doi: 10.1044/1092-4388(2012/11-0084). URL <http://pubs.asha.org/doi/10.1044/1092-4388%282012/11-0084%29>.

Catharine W. Keatley, John A. Spinks, and Beatrice De Gelder. Asymmetrical cross-language priming effects. *Memory Cognition*, 22(1):70–84, January 1994a. ISSN 1532-5946. doi: 10.3758/BF03202763. URL <https://doi.org/10.3758/BF03202763>.

Catharine W. Keatley, John A. Spinks, and Beatrice De Gelder. Asymmetrical cross-language priming effects. *Memory Cognition*, 22(1):70–84, January 1994b. ISSN 0090-502X, 1532-5946. doi: 10.3758/BF03202763. URL <http://link.springer.com/10.3758/BF03202763>.

Donghyun Kim, Meghan Clayards, and Heather Goad. A longitudinal study of individual differences in the acquisition of new vowel contrasts. *Journal of Phonetics*, 67:1–20, March 2018. ISSN 00954470. doi: 10.1016/j.wocn.2017.11.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447016301255>.

L. Kimura. The hawaiian language. In *Native Hawaiians Study Commission: Report on the Culture, Needs, and Concerns of Native Hawaiians*, pages 173–224. Department of Interior, 1983.

Larry Lindsey Kauanoë Kimura. Kani‘āina, ulukau. org digital repository of ka haka ‘ula o ke‘elikōlani college of hawaiian language, university of hawai‘i at hilo, ka leo hawai‘i. 1972.

Sachiko Kinoshita, Serje Robidoux, Luke Mills, and Dennis Norris. Visual similarity effects on masked priming. *Memory Cognition*, 42(5):821–833, 2014. ISSN 0090-502X. doi: 10.3758/s13421-013-0388-4. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4055850/>.

Swathi Kiran and Keith R. Lebel. Crosslinguistic semantic and translation priming in normal bilingual individuals and bilingual aphasia. *Clinical Linguistics Phonetics*, 21(4):277–303, January 2007. ISSN 0269-9206, 1464-5076. doi: 10.1080/02699200701243873. URL <http://www.tandfonline.com/doi/full/10.1080/02699200701243873>.

James P. Kirby and Alan C. L. Yu. Lexical and phonotactic effects on word-likeness judgments in cantonese. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 1389–1392, 2007.

- M. Joelle Kirtley, James Grama, Katie Drager, and Sean Simpson. An acoustic analysis of the vowels of hawai'i english. *Journal of the International Phonetic Association*, 46(1):79–97, April 2016. ISSN 0025-1003, 1475-3502. doi: 10.1017/S0025100315000456. URL https://www.cambridge.org/core/product/identifier/S0025100315000456/type/journal_article.
- Megan M. Kittleson, Jessica M. Aguilar, Gry Line Tokerud, Elena Plante, and Arve E. Asbjørnsen. Implicit language learning: Adults' ability to segment words in norwegian. *Bilingualism: Language and Cognition*, 13(4):513–523, October 2010. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728910000039. URL https://www.cambridge.org/core/product/identifier/S1366728910000039/type/journal_article.
- Daniel Kleinman and Tamar H. Gollan. Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, 27(5):700–714, May 2016. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797616634633. URL <https://journals.sagepub.com/doi/10.1177/0956797616634633>.
- Dave F. Kleinschmidt and T. Florian Jaeger. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, 122(2):148–203, April 2015. ISSN 1939-1471, 0033-295X. doi: 10.1037/a0038695. URL <https://doi.apa.org/doi/10.1037/a0038695>.
- John H. Kranzler. *Mental Chronometry*, page 2180–2182. Springer, Boston, MA, 2012. ISBN 9781441914286. doi: 10.1007/978-1-4419-1428-6_238. URL https://link.springer.com/rwe/10.1007/978-1-4419-1428-6_238.
- Jennifer Krizman, Ann R. Bradlow, Silvia Siu-Yin Lam, and Nina Kraus. How bilinguals listen in noise: linguistic and non-linguistic factors. *Bilingualism: Language and Cognition*, 20(4):834–843, August 2017. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728916000444. URL https://www.cambridge.org/core/product/identifier/S1366728916000444/type/journal_article.
- J.F. Kroll and E. Stewart. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2):149–174,

April 1994. ISSN 0749596X. doi: 10.1006/jmla.1994.1008. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X84710084>.

John K. Kruschke. Bayesian analysis reporting guidelines. *Nature Human Behaviour*, 5(10):1282–1291, August 2021. ISSN 2397-3374. doi: 10.1038/s41562-021-01177-7. URL <https://www.nature.com/articles/s41562-021-01177-7>.

Hale Kuamo‘o. Ulukau: The hawaiian electronic library. *Retrieved from*, 2009.

Patricia K Kuhl, Barbara T Conboy, Sharon Coffey-Corina, Denise Padden, Maritza Rivera-Gaxiola, and Tobey Nelson. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (nlm-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493):979–1000, March 2008. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2007.2154. URL <https://royalsocietypublishing.org/doi/10.1098/rstb.2007.2154>.

Jan-Rouke Kuipers and Guillaume Thierry. Event-related brain potentials reveal the time-course of language change detection in early bilinguals. *NeuroImage*, 50(4):1633–1638, May 2010. ISSN 10538119. doi: 10.1016/j.neuroimage.2010.01.076. URL <https://linkinghub.elsevier.com/retrieve/pii/S1053811910000984>.

William Labov. On the adequacy of natural language: The development of tense. In J. V. Singler, editor, *Pidgin and Creole Tense-Mood-Aspect Systems*, pages 1–58. John Benjamins, Amsterdam/Philadelphia, 1971. [1990 reprint].

William Labov. *Principles of linguistic change*. Language in society. Blackwell, Oxford, UK; Cambridge [Mass.], 1994. ISBN 9780631179139.

Sunyoung Lee-Ellis. *Looking into Bilingualism through the Heritage Speaker’s Mind*. Doctoral dissertation, ProQuest Dissertations Publishing, Ann Arbor, MI, USA, 2012. URL <http://www.proquest.com/en-US/products/dissertations/individuals.shtml>.

Kristin Lemhofer, Ton Dijkstra, and Marije Michel. Three languages, one echo: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5):585–611, October 2004. ISSN 0169-0965, 1464-0732. doi: 10.1080/01690960444000007. URL <http://www.tandfonline.com/doi/abs/10.1080/01690960444000007>.

- Russell V. Lenth. *emmeans: Estimated Marginal Means, aka Least-Squares Means*, 2025. URL <https://rvlenth.github.io/emmeans/>. R package version 1.11.1-00001.
- Leigh Lisker. "voicing" in english: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and speech*, 29(1): 3, 1986. URL <https://www.proquest.com/scholarly-journals/voicing-english-catalogue-acoustic-features/docview/1299120945/se-2>. Last updated - 2013-02-23.
- Miquel Llopart. Lexical and phonetic influences on the phonolexical encoding of difficult second-language contrasts: Insights from nonword rejection. *Frontiers in Psychology*, 12:659852, May 2021. ISSN 1664-1078. doi: 10.3389/fpsyg.2021.659852. URL <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.659852/full>.
- Helen M. Lockwood and Scott L. Saft. Shifting language ideologies and the perceptions of hawai'i creole among educators at the university level in hawai'i. *Linguistics and Education*, 33:1–13, 2016.
- Terry L. Lucas. *Language, Schooling, and the Suppression of Hawaiian*. Routledge, 2000.
- Paul A. Luce and David B. Pisoni. Recognizing spoken words: The neighborhood activation model:. *Ear and Hearing*, 19(1):1–36, February 1998. ISSN 0196-0202. doi: 10.1097/00003446-199802000-00001. URL <http://journals.lww.com/00003446-199802000-00001>.
- Stephen J. Lupker and Penny M. Pexman. Making things difficult in lexical decision: The impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5):1267–1289, 2010. ISSN 1939-1285, 0278-7393. doi: 10.1037/a0020125. URL <https://doi.apa.org/doi/10.1037/a0020125>.
- James S. Magnuson, James A. Dixon, Michael K. Tanenhaus, and Richard N. Aslin. The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, 31(1):133–156, February 2007. ISSN 0364-0213, 1551-6709. doi: 10.1080/03640210709336987. URL <https://onlinelibrary.wiley.com/doi/10.1080/03640210709336987>.
- David Uahikeaikalei'ohu Maile. Becoming hoa: Indigenous political thought and the question of settler aloha 'Āina in hawai'i. *American Quarterly*, 76

(3):617–643, 2024. ISSN 1080-6490. doi: 10.1353/aq.2024.a937120. URL <https://muse.jhu.edu/article/937120>.

Viorica Marian and Michael Spivey. Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 6(2):97–115, 2003. doi: 10.1017/S1366728903001068.

Caterina Marino, Carline Bernard, and Judit Gervain. Word frequency is a cue to lexical category for 8-month-old infants. *Current Biology*, 30(8):1380–1386.e3, April 2020. ISSN 0960-9822. doi: 10.1016/j.cub.2020.01.070. URL <https://www.sciencedirect.com/science/article/pii/S0960982220301147>.

Mikaela L. Marlow and Howard Giles. Who you tink you, talkin propah? hawaiian pidgin demarginalized. *Journal of Multicultural Discourses*, 3(1): 53–68, 2008. doi: 10.2167/md038.0.

Mikaela L. Marlow and Howard Giles. ‘we won’t get ahead speaking like that!’ expressing and managing language criticism in hawai’i. *Journal of Multilingual and Multicultural Development*, 31(3):237–251, May 2010a. ISSN 0143-4632, 1747-7557. doi: 10.1080/01434630903582714. URL <http://www.tandfonline.com/doi/abs/10.1080/01434630903582714>.

Mikaela L. Marlow and Howard Giles. ‘we won’t get ahead speaking like that!’ expressing and managing language criticism in hawai’i. *Journal of Multilingual and Multicultural Development*, 31(3):237–251, 2010b. doi: 10.1080/01434631003735452.

William D. Marslen-Wilson. Functional parallelism in spoken word-recognition. *Cognition*, 25(1–2):71–102, March 1987. ISSN 00100277. doi: 10.1016/0010-0277(87)90005-9. URL <https://linkinghub.elsevier.com/retrieve/pii/0010027787900059>.

William D. Marslen-Wilson. Activation, competition, and frequency in lexical access. In Gerry T. M. Altmann, editor, *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, pages 148–172. MIT Press, Cambridge, MA, 1990. ISBN 0262011174.

Andrew Martin, Sharon Peperkamp, and Emmanuel Dupoux. Learning phonemes with a proto-lexicon. *Cognitive Science*, 37(1):103–124, January 2013. ISSN 0364-0213, 1551-6709. doi: 10.1111/j.1551-6709.

2012.01267.x. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1551-6709.2012.01267.x>.

Kalikoaloha Martin. Re-membering and re-claiming pidgin: Pidgin identity in the classroom. Conference Presentation at the Hawaiian-Pidgin Summit, Honolulu, HI, United States, October 25 2018. URL <https://www.hawaii.edu/satocenter/?p=1732>.

María Cruz. Martín, Pedro. Macizo, and Teresa. Bajo. Time course of inhibitory processes in bilingual language processing. *British Journal of Psychology*, 101(4):679–693, November 2010. ISSN 0007-1269, 2044-8295. doi: 10.1348/000712609X480571. URL <https://bpspsychub.onlinelibrary.wiley.com/doi/10.1348/000712609X480571>.

Wakayo Mattingley, Jennifer Hay, Simon Todd, Forrest Panther, Jeanette King, and Peter J. Keegan. Ongoing exposure to an ambient language continues to build implicit knowledge across the lifespan. *Linguistics Vanguard*, 10(1):345–355, December 2024a. ISSN 2199-174X. doi: 10.1515/lingvan-2023-0137. URL <https://www.degruyter.com/document/doi/10.1515/lingvan-2023-0137/html>.

Wakayo Mattingley, Forrest Panther, Simon Todd, Jeanette King, Jennifer Hay, and Peter J. Keegan. Awakening the proto-lexicon: A proto-lexicon gives learning advantages for intentionally learning a language. *Language Learning*, 74(3):744–776, September 2024b. ISSN 0023-8333, 1467-9922. doi: 10.1111/lang.12635. URL <https://onlinelibrary.wiley.com/doi/10.1111/lang.12635>.

Sven L. Mattys and Peter W. Jusczyk. Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance*, 27(3):644–655, 2001. doi: 10.1037/0096-1523.27.3.644. URL <https://doi.org/10.1037/0096-1523.27.3.644>.

Sven L. Mattys, Peter W. Jusczyk, Paul A. Luce, and James L. Morgan. Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38(4):465–494, 1999. doi: 10.1006/cogp.1999.0721. URL <https://doi.org/10.1006/cogp.1999.0721>.

Sven L. Mattys, Matthew H. Davis, Ann R. Bradlow, and Sophie K. Scott. Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7–8):953–978, September 2012. ISSN 0169-0965, 1464-0732. doi: 10.1080/01690965.2012.705006. URL <http://www.tandfonline.com/doi/abs/10.1080/01690965.2012.705006>.

- James L McClelland and Jeffrey L Elman. The trace model of speech perception. *Cognitive Psychology*, 18(1):1–86, January 1986. ISSN 00100285. doi: 10.1016/0010-0285(86)90015-0. URL <https://linkinghub.elsevier.com/retrieve/pii/0010028586900150>.
- Richard McElreath. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC, 2 edition, March 2020. ISBN 9780429029608. doi: 10.1201/9780429029608. URL <https://www.taylorfrancis.com/books/9780429642319>.
- Timothy P. McNamara. *Semantic priming: perspectives from memory and word cognition*. Essays in cognitive psychology. Psychology Press, New York, 2005. ISBN 9781841690797.
- Jacques Mehler, Peter Jusczyk, Gérard Lambertz, Mary Halsted, Josiane Bertoncini, and Claudine Amiel-Tison. A precursor of language acquisition in young infants. *Cognition*, 29:143–178, 1988.
- Merriam-Webster, Inc. Merriam-webster online dictionary, 2025. URL <https://www.merriam-webster.com/>. Accessed: 2025-09-18.
- Renata F.I. Meuter and Alan Allport. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40(1):25–40, January 1999. ISSN 0749596X. doi: 10.1006/jmla.1998.2602. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X9892602X>.
- David E. Meyer and Roger W. Schvaneveldt. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2):227–234, October 1971. ISSN 0022-1015. doi: 10.1037/h0031564. URL <https://doi.apa.org/doi/10.1037/h0031564>.
- Jeff Miller. Estimating the proportions and latencies of reaction time outliers: A pooling method and case study of lexical decision tasks. *Behavior Research Methods*, 56(7):7280–7306, 2024. ISSN 1554-351X. doi: 10.3758/s13428-024-02419-y. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11362516/>.
- Daniel Mirman, James S. Magnuson, Katharine Graf Estes, and James A. Dixon. The link between statistical segmentation and word learning in adults. *Cognition*, 108(1):271–280, 2008. doi: 10.1016/j.cognition.2008.02.003. URL <https://doi.org/10.1016/j.cognition.2008.02.003>.

- Ines M. Miyares. Expressing “local culture” in hawai’i*. *Geographical Review*, 98(4):513–531, October 2008. ISSN 0016-7428, 1931-0846. doi: 10.1111/j.1931-0846.2008.tb00315.x. URL <https://www.tandfonline.com/doi/full/10.1111/j.1931-0846.2008.tb00315.x>.
- Eva M. Moreno, Kara D. Federmeier, and Marta Kutas. Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, 80(2):188–207, February 2002. ISSN 0093934X. doi: 10.1006/brln.2001.2588. URL <https://linkinghub.elsevier.com/retrieve/pii/S0093934X0192588X>.
- Benjamin Munson, Elizabeth C. McDonald, Nancy L. DeBoe, and Aubrey R. White. The acoustic and perceptual bases of judgments of women and men’s sexual orientation from read speech. *Journal of Phonetics*, 34(2):202–240, April 2006. ISSN 00954470. doi: 10.1016/j.wocn.2005.05.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447005000379>.
- Kathleen E. Murphy. *Melodies of Hawai’i: The Relationship Between Hawai’i Creole English and ‘Ōlelo Hawai’i Prosody*. PhD thesis, University of Calgary, Calgary, Alberta, 2013. Publication No. NR96756. ProQuest Dissertations and Theses Global.
- Danielle Navarro. lsr: Companion to “learning statistics with r”, December 2021. URL <https://cran.r-project.org/web/packages/lsr/index.html>.
- Thierry Nazzi, Josiane Bertoncini, and Jacques Mehler. Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3):756–766, 1998.
- James H. Neely. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106(3):226–254, 1977. ISSN 1939-2222, 0096-3445. doi: 10.1037/0096-3445.106.3.226. URL <https://doi.apa.org/doi/10.1037/0096-3445.106.3.226>.
- James H. Neely. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In *Basic Processes in Reading: Visual Word Recognition*, pages 264–336. Routledge, 1991. ISBN 0-8058-0219-3, 978-0-8058-0219-1. doi: 10.4324/9780203052242-9.

- Richard “Keao” Nesmith. *Tutu’s Hawaiian and the Emergence of a Neo-Hawaiian Language*. PhD thesis, University of Hawaii at Manoa, 2002. URL <http://hdl.handle.net/10125/21194>.
- Boris New, Jessica Bourgin, Julien Barra, and Christophe Pallier. Unipseudo: A universal pseudoword generator. *Quarterly Journal of Experimental Psychology*, 77(2):278–286, 2024.
- Elissa L. Newport and Richard N. Aslin. Innately constrained learning: Blending old and new approaches to language acquisition. In *Proceedings of the 24th Annual Boston University Conference on Language Development*, volume 1, pages 1–21, 2000.
- Céline Ngon, Andrew Martin, Emmanuel Dupoux, Dominique Cabrol, Michel Dutat, and Sharon Peperkamp. (non)words, (non)words, (non)words: evidence for a protolexicon during the first year of life. *Developmental Science*, 16(1):24–34, January 2013. ISSN 1363-755X, 1467-7687. doi: 10.1111/j.1467-7687.2012.01189.x. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-7687.2012.01189.x>.
- Nancy Niedzielski. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18(1):62–85, March 1999. ISSN 0261-927X, 1552-6526. doi: 10.1177/0261927X99018001005. URL <https://journals.sagepub.com/doi/10.1177/0261927X99018001005>.
- Georganne Nordstrom. Pidgin as rhetorical sovereignty. In *2018 Hawaiian-Pidgin Summit*, Honolulu, HI, United States, 2018. URL <https://www.hawaii.edu/satocenter/?tag=hawaiian-pidgin-summit>. [Conference Presentation].
- Eleanor C. Nordyke. *The Peopling of Hawai’i*. University of Hawai’i Press, 1989.
- Dennis Norris, James M. McQueen, and Anne Cutler. Perceptual learning in speech. *Cognitive Psychology*, 47(2):204–238, September 2003. ISSN 0010-0285. doi: 10.1016/S0010-0285(03)00006-9. URL <https://www.sciencedirect.com/science/article/pii/S0010028503000069>.
- Y. Oh, S. Todd, C. Beckner, J. Hay, J. King, and J. Needle. Non-māori-speaking new zealanders have a māori proto-lexicon. *Scientific Reports*, 10(1):22318, December 2020. ISSN 2045-2322. doi:

10.1038/s41598-020-78810-4. URL <https://www.nature.com/articles/s41598-020-78810-4>.

Kaleikoa Ka'eo Oliveira. *An Indigenous Politics of Language: Hawaiian Language Revitalization*. University of Hawai'i Press, 2014.

Katrina-Ann R. Kapā'anaokalāokeola Nākoa Oliveira. *Kanaka 'Ōiwi Methodologies: Mo'olelo and Metaphor*. Hawai'inuiākea Series. University of Hawaii Press, Honolulu, 1st ed edition, 2015. ISBN 9780824857516.

Luca Onnis, Padraic Monaghan, Korin Richmond, and Nick Chater. Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, 53(2):225–237, August 2005. ISSN 0749596X. doi: 10.1016/j.jml.2005.02.011. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X05000318>.

Forrest Panther, Wakayo Mattingley, Jen Hay, Simon Todd, Jeanette King, and Peter J. Keegan. Morphological segmentations of non-māori speaking new zealanders match proficient speakers. *Bilingualism: Language and Cognition*, 27(1):1–15, 2023a. doi: 10.1017/s1366728923000329. URL <https://doi.org/10.1017/s1366728923000329>.

Mattingley Wakayo Panther, Forrest Andrew, Simon Todd, Jennifer Hay, and Jeanette King. Proto-lexicon size and phonotactic knowledge are linked in non-māori speaking new zealand adults. *Laboratory Phonology*, 14(1), 2023b.

Alexandra Papoutsaki, Patsorn Sangkloy, James Laskey, Nediya Daskalova, Jeff Huang, and James Hays. Webgazer: scalable webcam eye tracking using user interactions. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16*, page 3839–3845. AAAI Press, 2016. ISBN 9781577357704.

'Ōiwi Parker Jones. Hawaiian. *Journal of the International Phonetic Association*, 48(1):103–115, April 2018. ISSN 0025-1003, 1475-3502. doi: 10.1017/S0025100316000438. URL https://www.cambridge.org/core/product/identifier/S0025100316000438/type/journal_article.

Marcela Pena, Luca L. Bonatti, Marina Nespor, and Jacques Mehler. Signal-driven computations in speech processing. *Science*, 298(5593):604–607, October 2002. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1072901. URL <https://www.science.org/doi/10.1126/science.1072901>.

- Manuel Perea, Eva Rosa, and Consolación Gómez. The frequency effect for pseudowords in the lexical decision task. *Perception Psychophysics*, 67 (2):301–314, February 2005. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03206493. URL <http://link.springer.com/10.3758/BF03206493>.
- P. E. Piccinini and M. Garellek. Prosodic cues to monolingual versus code-switching sentences in english and spanish. In N. Campbell, D. Gibbon, and D. Hirst, editors, *Proceedings of the 7th International Conference on Speech Prosody*, pages 885–889, Dublin, Ireland, 2014. Trinity College Dublin.
- Janet B. Pierrehumbert. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2-3):115–154, 2003a. doi: 10.1177/00238309030460020501. URL <https://doi.org/10.1177/00238309030460020501>.
- Janet B. Pierrehumbert. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2–3):115–154, June 2003b. ISSN 0023-8309, 1756-6053. doi: 10.1177/00238309030460020501. URL <https://journals.sagepub.com/doi/10.1177/00238309030460020501>.
- Shana Poplack. Sometimes i’ll start a sentence in spanish y termino en español: toward a typology of code-switching1. *Linguistics*, 18(7–8), 1980. ISSN 0024-3949, 1613-396X. doi: 10.1515/ling.1980.18.7-8.581. URL <https://www.degruyter.com/document/doi/10.1515/ling.1980.18.7-8.581/html>.
- Mary C. Potter, Kwok-Fai So, Barbara Von Eckardt, and Laurie B. Feldman. Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1): 23–38, February 1984. ISSN 00225371. doi: 10.1016/S0022-5371(84)90489-4. URL <https://linkinghub.elsevier.com/retrieve/pii/S0022537184904894>.
- Mary Kawena Pukui and Samuel H. Elbert. *Hawaiian Dictionary: Hawaiian-English, English-Hawaiian*. University of Hawai’i Press, Honolulu, revised and enlarged edition edition, 1986.
- Mary Kawena Pukui, Samuel H. (Samuel Hoyt) Elbert, and Esther T. Mookini. *The pocket Hawaiian dictionary, with a concise Hawaiian grammar*. Honolulu: University Press of Hawaii, 1975. ISBN 9780824803070. URL <http://archive.org/details/pockethawaiiandi00puku>.

- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.
- Monique Radeau, José Morais, and Juan Segui. Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6):1297–1311, December 1995. ISSN 1939-1277, 0096-1523. doi: 10.1037/0096-1523.21.6.1297. URL <https://doi.apa.org/doi/10.1037/0096-1523.21.6.1297>.
- Roger Ratcliff, Pablo Gomez, and Gail McKoon. A diffusion model account of the lexical decision task. *Psychological Review*, 111(1):159–182, 2004. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.111.1.159. URL <https://doi.apa.org/doi/10.1037/0033-295X.111.1.159>.
- Erik D. Reichle and Charles A. Perfetti. Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, 7(3):219–237, July 2003. ISSN 1088-8438, 1532-799X. doi: 10.1207/S1532799XSSR0703_2. URL http://www.tandfonline.com/doi/abs/10.1207/S1532799XSSR0703_2.
- J. E. Reinecke. ‘pidgin english’ in hawai‘i: A local study in the sociology of language. *American Journal of Sociology*, 43(5):778–789, 1938. doi: 10.1086/217899.
- John E. Reinecke. *Language and Dialect in Hawaii*. Social Science Research Institute, University of Hawai‘i, Honolulu, 1969.
- Julian M. Roberts. Pidgin hawaiian: A sociohistorical study. *Journal of Pidgin and Creole Languages*, 10:1–56, 1995.
- Sarah J. Roberts. The emergence of hawai‘i creole english in the early 20th century: The sociohistorical context. *Language in Society*, 27(4):517–540, 1998.
- Sarah J. Roberts. Nativization and the emergence of hawai‘i creole. In Jeff Siegel, editor, *Processes of Language Contact*, pages 123–150. Cambridge University Press, Cambridge, 2000.
- Sarah J. Roberts. The emergence of hawai‘i creole english in the early 20th century: The demographic and social bases. In Nicholas Faraclas, Thomas Klein, and Christa Schneider, editors, *Pidgin and Creole Linguistics in the Twenty-first Century*, pages 301–330. Lincom Europa, 2005.

- Suzanne Romaine. Changing attitudes to hawai'i creole english: Fo' find one good job, you gotta know how fo' talk like one haole. In John R. Rickford and Suzanne Romaine, editors, *Creole Genesis, Attitudes and Discourse: Studies Celebrating Charlene J. Sato*, pages 287–301. John Benjamins, Amsterdam/Philadelphia, 1999.
- Ali Rozet. Yuzing pidgin foh teeoh hawaiian to mainland an loco needs alike. In *Summit on Pidgin Education*, Honolulu, HI, United States, October 17 2017a. URL <http://www.hawaii.edu/satocenter/?p=1379>. [Conference Presentation].
- Alicia Rozet. Yuzing pidgin foh teeoh hawaiian to mainland an loco needs alike. Conference Presentation at the Summit on Pidgin Education, Honolulu, HI, United States, October 17 2017b. URL <http://www.hawaii.edu/satocenter/?p=1379>.
- Alicia Rozet. Yuzing pidgin foh teeoh hawaiian to mainland an loco needs alike. Conference Presentation at the Hawaiian-Pidgin Summit, Honolulu, HI, United States, October 25 2018. URL <https://www.hawaii.edu/satocenter/?p=1729>.
- Jenny R. Saffran, Richard N. Aslin, and Elissa L. Newport. Statistical learning by 8-month-old infants. *Science*, 274(5294):1926–1928, 1996a. doi: 10.1126/science.274.5294.1926. URL <https://doi.org/10.1126/science.274.5294.1926>.
- Jenny R. Saffran, Richard N. Aslin, and Elissa L. Newport. Statistical learning by 8-month-old infants. *Science*, 274(5294):1926–1928, December 1996b. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.274.5294.1926. URL <https://www.science.org/doi/10.1126/science.274.5294.1926>.
- Jenny R. Saffran, Elissa L. Newport, and Richard N. Aslin. Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4):606–621, 1996c. doi: 10.1006/jmla.1996.0032. URL <https://doi.org/10.1006/jmla.1996.0032>.
- Scott Saft. *Exploring Multilingual Hawai'i: Language Use and Language Ideologies in a Diverse Society*. Lexington Books, Lanham, MD, 2019. ISBN 9781498561198.
- Scott Saft. *Language and Social Justice in Context: Hawai'i As a Case Study*. Springer International Publishing AG, Cham, 2022. ISBN 9783030912505.

- Scott Saft, Gabrielle Tebow, and Rose Santos. Hawai'i creole in the public domain: Humor, emphasis, and heteroglossic language practice in university commencement speeches. *Pragmatics*, 28(3):417–438, 2018. doi: 10.1075/prag.00007.saf.
- Kent Sakoda and Jeff Siegel. *Pidgin grammar: an introduction to the Creole English of Hawai'i*. Bess Press, Honolulu, Hawaii, 2003. ISBN 9781573061698.
- Kent Sakoda and Jeff Siegel. Hawai'i creole: Phonology. In Bernd Kortmann and Edgar W. Schneider, editors, *A Handbook of Varieties of English, Volume 1: Phonology*, pages 729–749. Mouton de Gruyter, Berlin, 2008.
- Arthur G. Samuel and Saioa Larraza. Does listening to non-native speech impair speech perception? *Journal of Memory and Language*, 81:51–71, May 2015. ISSN 0749596X. doi: 10.1016/j.jml.2015.01.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X15000042>.
- Charlene J. Sato. Linguistic inequality in hawaii: The post-creole dilemma. In Nessa Wolfson and Joan Manes, editors, *Language of Inequality*, pages 255–272. De Gruyter, Berlin and Boston, 1985. doi: 10.1515/9783110857320.255.
- Charlene J. Sato. Sociolinguistic variation and language attitudes in hawaii. In Jenny Cheshire, editor, *English around the World*, pages 647–663. Cambridge University Press, 1991a. doi: 10.1017/CBO9780511611889.044. URL https://www.cambridge.org/core/product/identifier/CBO9780511611889A051/type/book_part.
- Charlene J. Sato. Sociolinguistic variation and language attitudes in hawai'i: An exploratory study. *University of Hawai'i Working Papers in ESL*, 10(1):91–135, 1991b.
- Charlene J. Sato. Language change in a creole continuum: Decreolization? In Kenneth Hyltenstam and Åke Viberg, editors, *Progression and Regression in Language: Sociocultural, Neuropsychological and Linguistic Perspectives*, pages 122–143. Cambridge University Press, Cambridge, 1993.
- Jessamyn Schertz, Taehong Cho, Andrew Lotto, and Natasha Warner. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, 52:183–204, September 2015. ISSN 00954470. doi: 10.1016/j.wocn.2015.07.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095447015000534>.

- Hildur E. H. Schilling, Keith Rayner, and James I. Chumbley. Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory Cognition*, 26(6):1270–1281, November 1998. ISSN 0090-502X, 1532-5946. doi: 10.3758/BF03201199. URL <http://link.springer.com/10.3758/BF03201199>.
- Albert J. Schütz. *The Voices of Eden: A History of Hawaiian Language Studies*. University of Hawai‘i Press, Honolulu, 1994. URL <https://archive.org/details/voicesofedenhist0000schu>.
- Albert J. Schütz, Puakea Nogelmeier, and Kenneth L. Cook. *Hawaiian Language: Past, Present, and Future*. University of Hawai‘i Press, Honolulu, 2005. See page 96.
- Alice Shen, Susanne Gahl, and Keith Johnson. Didn’t hear that coming: Effects of withholding phonetic cues to code-switching. *Bilingualism: Language and Cognition*, 23(5):1020–1031, November 2020. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728919000877. URL https://www.cambridge.org/core/product/identifier/S1366728919000877/type/journal_article.
- Lorrin A. Shi. *Educational policy and language shift in Hawai‘i*. University of Hawai‘i, 1990.
- Shun. spectral-subtraction: Spectral subtraction in python. <https://github.com/shun60s/spectral-subtraction>, 2020. Accessed: 2025-07-01.
- Jeff Siegel. *The emergence of pidgin and Creole languages / Jeff Siegel*. Oxford linguistics. Oxford University Press, Oxford ; New York, 2008. ISBN 9780199216673.
- Noenoe K. Silva. *The Power of the Steel-Tipped Pen: Reconstructing Native Hawaiian Intellectual History*. Duke University Press, 2017.
- Douglas Simonson, Pat Sasaki, and Ken Sakata. *Pidgin to da max hana hou*. Honolulu: Bess, 1992. URL <http://archive.org/details/pidgintodamaxhan00simo>.
- Douglas Simonson, Pat Sasaki, and Ken Sakata. *Pidgin to Da Max*. Bess Press, 2005. ISBN 9781573062503. Google-Books-ID: QAcJSXpYoucC.

- Henrik Skovsgaard, Javier San Agustin, Sune Alstrup Johansen, John Paulin Hansen, and Martin Tall. Evaluation of a remote webcam-based eye tracker. In *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*, NGCA '11, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450306805. doi: 10.1145/1983302.1983309. URL <https://doi.org/10.1145/1983302.1983309>.
- Yisrael Smith, Joel Walters, and Anat Prior. Target accessibility contributes to asymmetric priming in translation and cross-language semantic priming in unbalanced bilinguals. *Bilingualism: Language and Cognition*, 22(1):157–176, January 2019. ISSN 1366-7289, 1469-1841. doi: 10.1017/S1366728917000645. URL https://www.cambridge.org/core/product/identifier/S1366728917000645/type/journal_article.
- Ha'alilo Solomon. Attitudes and ideologies surrounding hawaiian language and pidgin. In *2018 Hawaiian-Pidgin Summit*, Honolulu, HI, United States, 2018. URL <https://www.hawaii.edu/satocenter/?tag=hawaiian-pidgin-summit>. [Conference Presentation].
- Rachel Soo and Philip J Monahan. Language dominance and order of acquisition affect auditory translation priming in heritage speakers. *Quarterly Journal of Experimental Psychology*, 76(2):284–293, February 2023. ISSN 1747-0218, 1747-0226. doi: 10.1177/17470218221091753. URL <https://journals.sagepub.com/doi/10.1177/17470218221091753>.
- Michael J. Spivey and Viorica Marian. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10(3):281–284, 1999. doi: 10.1111/1467-9280.00151.
- David E. Stannard. *Before the Horror: The Population of Hawai'i on the Eve of Western Contact*. University of Hawai'i Press, 1989.
- Robert F. Stanners, James E. Jastrzembski, and Allen Westbrook. Frequency and visual quality in a word–nonword classification task. *Journal of Verbal Learning and Verbal Behavior*, 14:259–264, 1975. doi: 10.1016/S0022-5371(75)80069-7.
- G. O. Stone and G. C. Van Orden. Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 19(4):744–774, 1993.

- Holly L. Storkel. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44(6):1321–1337, December 2001. ISSN 1092-4388, 1558-9102. doi: 10.1044/1092-4388(2001/103). URL <http://pubs.asha.org/doi/10.1044/1092-4388%282001/103%29>.
- Holly L. Storkel and Margaret A. Rogers. The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics Phonetics*, 14:407–425, 2000. doi: 10.1080/026992000415859.
- Peter Sulzberger. *Wordlikeness judgements, phonotactic probability, accent and repetition in auditory recognition of novel words in a foreign language*. Unpublished doctoral dissertation, Victoria University of Wellington, 2007.
- Meghan Sumner and Arthur G. Samuel. Lexical inhibition and sublexical facilitation are surprisingly long lasting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4):769–790, 2007. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.33.4.769. URL <https://doi.apa.org/doi/10.1037/0278-7393.33.4.769>.
- Anita Szakay, Molly Babel, and Jeanette King. Social categories are shared across bilinguals’ lexicons. *Journal of Phonetics*, 59:92–109, 2016. ISSN 1095-8576. doi: 10.1016/j.wocn.2016.09.005.
- Marcus Taft and Kenneth I. Forster. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14:638–647, 1975.
- Eileen H. Tamura. Power, status, and hawaii creole english: An example of linguistic intolerance in american history. *Pacific Historical Review*, 65(3): 431–454, 1996.
- Audacity Team. Audacity 3.1. 3. *Audacity Team*, 2021.
- Territory of Hawai‘i. *Reviewed Laws of Hawai‘i*. Hawaiian Gazette Co., Honolulu, 1905.
- Guillaume Thierry and Yan Jing Wu. Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, 104(30):12530–12535, July 2007. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0609927104. URL <https://pnas.org/doi/full/10.1073/pnas.0609927104>.

- Michael S.C Thomas and Alan Allport. Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, 43(1):44–66, July 2000. ISSN 0749596X. doi: 10.1006/jmla.1999.2700. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X99927006>.
- Sarah Grey Thomason and Terrence Kaufman. *Language Contact, Creolization, and Genetic Linguistics*. University of California Press, Berkeley, 1988.
- Simon Todd, Chadi Ben Youssef, and Alonso Vásquez-Aguilar. Language structure, attitudes, and learning from ambient exposure: Lexical and phonotactic knowledge of spanish among non-spanish-speaking californians and texans. *PLOS ONE*, 18(4):e0284919, April 2023a. ISSN 1932-6203. doi: 10.1371/journal.pone.0284919. URL <https://dx.plos.org/10.1371/journal.pone.0284919>.
- Simon Todd, Chadi Ben Youssef, and Alonso Vásquez-Aguilar. Language structure, attitudes, and learning from ambient exposure: Lexical and phonotactic knowledge of spanish among non-spanish-speaking californians and texans. *PLOS ONE*, 18(4):e0284919, 2023b. doi: 10.1371/journal.pone.0284919. URL <https://doi.org/10.1371/journal.pone.0284919>.
- Lee A. Tonouchi. Pidgin, identity, and language legitimacy. In *Language, Culture, and Identity in Hawai‘i*. University of Hawai‘i Press, Honolulu, 2021.
- Haunani-Kay Trask. *From a Native Daughter: Colonialism and Sovereignty in Hawai‘i*. University of Hawai‘i Press, Honolulu, 1999.
- Rebecca Treiman, Brett Kessler, Susan Knewasser, Ruth Tincoff, and Meredith Bowman. English speakers’ sensitivity to phonotactic patterns. In *Proceedings of the Fifth Conference on Laboratory Phonology*, Chicago, May 1996. Paper presented at the Fifth Conference on Laboratory Phonology.
- Endel Tulving, Daniel L. Schacter, and Heather A. Stark. Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(4): 336–342, July 1982. ISSN 1939-1285, 0278-7393. doi: 10.1037/0278-7393.8.4.336. URL <https://doi.apa.org/doi/10.1037/0278-7393.8.4.336>.
- Janet G. Van Hell and Ton Dijkstra. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4):780–789, 2002. doi: 10.3758/BF03196335.

- Walter J.B. Van Heuven, Ton Dijkstra, and Jonathan Grainger. Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39(3):458–483, October 1998. ISSN 0749596X. doi: 10.1006/jmla.1998.2584. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X98925840>.
- Janneke van Hofwegen and Walt Wolfram. Coming of age in african american english: A longitudinal study. *Journal of Sociolinguistics*, 14(4):427–455, 2010.
- Michael S. Vitevitch and Paul A. Luce. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40(3):374–408, April 1999. ISSN 0749596X. doi: 10.1006/jmla.1998.2618. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749596X98926183>.
- Michael S. Vitevitch and Paul A. Luce. Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, 52(2):193–204, 2005. doi: 10.1016/j.jml.2004.10.003. URL <https://doi.org/10.1016/j.jml.2004.10.003>.
- Michael S. Vitevitch, Paul A. Luce, Jan Charles-Luce, and David Kemmerer. Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, 40(1):47–62, January 1997. ISSN 0023-8309, 1756-6053. doi: 10.1177/002383099704000103. URL <https://journals.sagepub.com/doi/10.1177/002383099704000103>.
- Michael S. Vitevitch, Paul A. Luce, David B. Pisoni, and Edward T. Auer. Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and language*, 68(1–2):306–311, 1999. ISSN 0093-934X. doi: 10.1006/brln.1999.2116. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3466467/>.
- Myrte Vos, Serge Minor, and Gillian Catriona Ramchand. Comparing infrared and webcam eye tracking in the visual world paradigm. *Glossa Psycholinguistics*, 1(1), August 2022. ISSN 2767-0279. doi: 10.5070/G6011131. URL <https://escholarship.org/uc/item/3r28x18w>.
- Xuan Wang and Elsa Tragant. The effect of written text on comprehension of spoken english as a foreign language: A replication study. *International Review of Applied Linguistics in Language Teaching*, 60(3): 625–645, September 2022. ISSN 0019-042X, 1613-4141. doi: 10.1515/

iral-2018-0350. URL <https://www.degruyter.com/document/doi/10.1515/iral-2018-0350/html>.

Andrea Weber and Anne Cutler. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50(1):1–25, 2004. doi: 10.1016/S0749-596X(03)00105-0.

Janet F. Werker and Richard C. Tees. Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50:509–535, 1999. doi: 10.1146/annurev.psych.50.1.509.

C.P. Whaley. Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, 17(2):143–154, April 1978. ISSN 00225371. doi: 10.1016/S0022-5371(78)90110-X. URL <https://linkinghub.elsevier.com/retrieve/pii/S002253717890110X>.

Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019. doi: 10.21105/joss.01686. URL <https://doi.org/10.21105/joss.01686>.

Wikipedia contributors. Hawaiian-language wikipedia (‘Ōlelo hawai‘i) — home page. <https://haw.wikipedia.org>, 2025. Accessed: 2025-10-23.

John N. Williams. The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive Psychology*, 6(2):195–220, June 1994. ISSN 0954-1446, 1464-0635. doi: 10.1080/09541449408520143. URL <http://www.tandfonline.com/doi/abs/10.1080/09541449408520143>.

W. Wilson and K. Kamanā. For the interest of the hawaiians themselves: Reclaiming the benefits of hawaiian-medium education. *Hūlili: Multidisciplinary Research on Hawaiian Well-Being*, 3:153–182, 2006.

William H. Wilson. The sociopolitical context of establishing hawaiian-medium education. *Language, Culture, and Curriculum*, 11(3):325–338, 1998a.

Bibliography

- William H. Wilson. I ka ‘ōlelo hawai‘i ke ola: Life is found in the hawaiian language. *International Journal of the Sociology of Language*, 132:123–137, 1998b.
- William H. Wilson and Kauanoē Kamanā. Mai loko mai o ka ‘i‘ini: Proceeding from a dream, the ‘aha pūnana leo connection in hawaiian language revitalization. In Leanne Hinton and Kenneth Hale, editors, *The Green Book of Language Revitalization in Practice*, pages 147–176. Academic Press, 2001.
- William H. Wilson and Kauanoē Kamanā. I kumu lālā: Let there be sources; let there be branches: Teacher education in the college of hawaiian language. *Journal of American Indian Education*, 46(3):37–53, 2007.
- William H. Wilson and Keiki Kamanā. Insights from indigenous language immersion in hawai‘i. In Diane J. Tedick, David Christian, and T. W. Fortune, editors, *Immersion Education: Practices, Policies, Possibilities*, pages 36–57. Multilingual Matters, Clevedon, UK, 2011.
- Donald Winford. *An Introduction to Contact Linguistics*. Blackwell, Oxford, 2003.
- Benjamin Wist. *A century of public education in Hawaii, 1840 to 1940*. Hawaii Educational Review, Honolulu, 1940.
- Alia Wong. Fo teach pidgin o not fo teach pidgin — das da question, March 2013. URL <https://civilbeat.org/2013/03/18498-fo-teach-pidgin-o-not-fo-teach-pidgin-das-da-question/>.
- Alia Wong. The consequences of teaching exclusively in ‘standard english’, November 2015. URL <https://www.theatlantic.com/education/archive/2015/11/hawaiian-pidgin-recognized/416883/>.
- Kevin JP Woods, Max H Siegel, James Traer, and Josh H McDermott. Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, 79(7):2064–2072, 2017.
- O. L. L. Woodside, E. K. P. Reyes, and M. K. Jr. Kim. Nā hulu kūpuna: Living and sharing hawaiian wisdom. *Hūlili: Multidisciplinary Research on Hawaiian Well-Being*, 7:9–23, 2011.
- Yan Jing Wu and Guillaume Thierry. Unconscious translation during incidental foreign language processing. *NeuroImage*, 59(4):

3468–3473, February 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2011.11.049. URL <https://linkinghub.elsevier.com/retrieve/pii/S1053811911013309>.

A. Yacovone, E. Moya, and J. Snedeker. Unexpected words or unexpected languages? two erp effects of code-switching in naturalistic discourse. *Cognition*, 215:104814, 2021. doi: 10.1016/j.cognition.2021.104814.

J. Yamamoto and E. Hargrove. Teachers’ attitudes toward recorded speech samples of elementary school children in hawaii. *University of Hawaii Working Papers in Linguistics*, 4(2):109–134, 1982.

Lois A. Yamauchi, E. Brook Chapman De Sousa, Rebecca I. Ka’anehe, and Bryant Jensen. Instructional conversations for equitable participation to challenge deficit views of pidgin in hawai’i. *Journal of Language, Identity Education*, page 1–16, 2024. ISSN 1534-8458, 1532-7701. doi: 10.1080/15348458.2024.2366382. URL <https://www.tandfonline.com/doi/full/10.1080/15348458.2024.2366382>.

Melvin J. Yap, David A. Balota, David E. Sibley, and Roger Ratcliff. Individual differences in visual word recognition: Insights from the english lexicon project. *Journal of Experimental Psychology: Human Perception and Performance*, 38(1):53–79, 2012. doi: 10.1037/a0024177.

Melvin J. Yap, Daragh E. Sibley, David A. Balota, Roger Ratcliff, and Jay Rueckl. Responding to nonwords in the lexical decision task: Insights from the english lexicon project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3):597–613, 2014. ISSN 1939-1285, 0278-7393. doi: 10.1037/xlm0000064. URL <https://doi.apa.org/doi/10.1037/xlm0000064>.

Thomas Yokota. The “pidgin problem”: Attitudes about hawai’i creole. *Educational Perspectives*, 41:22–29, 2008.

Morris Young. *Minor Re/Visions: Asian American Literacy Narratives as a Rhetoric of Citizenship*. Southern Illinois University Press, Carbondale, 2004.

Tania S. Zamuner. Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research*, 52(1):49–60, February 2009. ISSN 1092-4388, 1558-9102. doi: 10.1044/1092-4388(2008/07-0138). URL <http://pubs.asha.org/doi/10.1044/1092-4388%282008/07-0138%29>.

- J. Zervakis and R. Mazuka. Effect of repeated evaluation and repeated exposure on acceptability ratings of sentences. *Journal of Psycholinguistic Research*, 42(6):505–525, 2013.
- Judy D. Zhu, Esti Blanco-Elorrieta, Yanan Sun, Anita Szakay, and Paul F. Sowman. Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage*, 247:118797, February 2022. ISSN 10538119. doi: 10.1016/j.neuroimage.2021.118797. URL <https://linkinghub.elsevier.com/retrieve/pii/S1053811921010685>.
- Johannes C. Ziegler and Usha Goswami. Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1):3–29, 2005. doi: 10.1037/0033-2909.131.1.3.
- Johannes C. Ziegler, Arthur M. Jacobs, and Dorothee Klüppel. Pseudohomophone effects in lexical decision: Still a challenge for current word recognition models. *Journal of Experimental Psychology: Human Perception and Performance*, 27(3):547–559, 2001. doi: 10.1037/0096-1523.27.3.547.
- Hōkūlani K. ‘Aikau. From malihini to hoa‘āina: Reconnecting people, places, and practices. In Nālani Wilson-Hokowhitu, editor, *The Past before Us: Mo‘okū‘auhau as Methodology*, page 91. University of Hawai‘i Press, Honolulu, 2019.
- ‘Ōiwi Parker Jones. Probabilities improve stress-prediction in a cfg of hawaiian phonology. In Lawrence Cavedon and Ingrid Zukerman, editors, *Proceedings of the Australasian Language Technology Workshop 2006*, pages 161–162, Sydney, Australia, November 2006. URL <https://aclanthology.org/U06-1024/>.

Appendix A

Language Background Questionnaire

Section 1: Life and language in Hawai‘i

Do you identify as being from Hawai‘i, are you Kānaka Maoli, and/or are you a speaker of Hawaiian or Pidgin (Hawai‘i)? Select all that apply. If none apply to you, please select “none” then skip to section 2:

From Hawai‘i

Kānaka Maoli

Speak Hawaiian

Speak Pidgin

None

If you are from Hawai‘i please indicate which city and island:

If you do speak Hawaiian which dialect do you speak? (e.g. Ni‘ihau dialect, Kaua‘i dialect, O‘ahu dialect, etc.)

Are (at least one) of your parents a fluent speaker of:

Pidgin?

Please Select... Yes No

Hawaiian?

Please Select... Yes No

English?

Please Select... Yes No

Are (at least one) of your grandparents a fluent speaker of:

Pidgin?

Please Select... Yes No

Hawaiian?

Please Select... Yes No

English?

Please Select... Yes No

Were you exposed to Pidgin before the age of 5?

Please Select... Yes No

Were you exposed to Hawaiian before the age of 5?

Appendix A. Language Background Questionnaire

Please Select... Yes No

Were you exposed to English before the age of 5?

Please Select... Yes No

Did you attend high school in Hawai'i? If so, which high school did you attend?

Would you describe yourself as diaspora Hawaiian?

Please Select... Yes No

If you are not Kānaka Maoli, how would you describe your ethnicity? List all that apply (e.g. Japanese, Chinese, Filipino etc.)

Which languages did you grow up with in the home? (select all that apply)

Pidgin

Hawaiian

English

Other (please specify)

Where did you receive your Pidgin knowledge? (select all that apply)

School

Family

Friends

Other (please specify)

Where did you receive your Hawaiian knowledge? (select all that apply)

School

Family

Friends

Other (please specify)

Where did you receive your English knowledge? (select all that apply)

School

Family

Friends

Other (please specify)

If applicable, please list the Hawaiian classes you have taken (e.g. Hwn 101, Hwn 102, Hwn 201, Hwn 202, High school Hwn (2 years), etc.)

If applicable, do you speak Pidgin with friends?

Yes always

Yes most of the time

Yes sometimes

Yes rarely

Never

If applicable, do you speak Hawaiian with friends?

Yes always

Appendix A. Language Background Questionnaire

Yes most of the time

Yes sometimes

Yes rarely

Never

If applicable, do you speak English with friends?

Yes always

Yes most of the time

Yes sometimes

Yes rarely

Never

Do you work/go to school in an environment where you often interact with Pidgin speakers?

Yes always

Yes most of the time

Yes sometimes

Yes rarely

No never

Do you work/go to school in an environment where you often interact with Hawaiian speakers?

Yes always

Yes most of the time

Yes sometimes

Yes rarely

No never

Do you work/go to school in an environment where you often interact with English speakers?

Yes always

Yes most of the time

Yes sometimes

Yes rarely

No never

Have you lived in a home that spoke/speaks Pidgin for more than 10 years?

Yes

No

Have you lived in a home that spoke/speaks Hawaiian for more than 10 years?

Yes

No

Appendix A. Language Background Questionnaire

Have you lived in a home that spoke/speaks English for more than 10 years?

Yes

No

Do you participate in any activities where you are surrounded by Pidgin speakers? (e.g. team sports, etc.)

Yes

No

Do you participate in any activities where you are surrounded by Hawaiian speakers? (e.g. team sports, etc.)

Yes

No

Do you participate in any activities where you are surrounded by English speakers? (e.g. team sports, etc.)

Yes

No

If applicable, in an average week, what percentage (%) of the time do you use Pidgin?

If applicable, in an average week, what percentage (%) of the time do you use Hawaiian?

If applicable, in an average week, what percentage (%) of the time do you use English?

If applicable, in an average week, what percentage (%) of the time do you use other language(s)?

How would you describe your relation to Hawaiian identity? (Open ended response)

How would you describe your relation to “local” identity?

Section 2: General Language Experience and Proficiency Questionnaire

What is your age?

Gender identity?

Highest level of education?

Basic questions - language

Please list all the languages you know in order of dominance (e.g. English, Spanish, French, Cantonese, Visayan, Pidgin, etc.)

Appendix A. Language Background Questionnaire

Please list all the languages you know in order of when you learned them (your native language first):

Language questions

For this section, answer each question for all languages you speak.

For each question, please list the language, then answer the question as it pertains to that language (e.g. Question: Please indicate the approximate age you began acquiring the language. Answer: Pidgin: birth, English: age 7, Hawaiian: age 7 etc.)

For each language you speak, please indicate the approximate age...

a. you began acquiring the language? (if learned from birth write “birth”)

b. became fluent in the language? (if learned from birth write “native language”)

On a scale from 0-10 (with 0 being the lowest, 10 being the highest), please select your level of proficiency in speaking in each language you know (write n/a if not applicable):

On a scale from 0-10 (with 0 being the lowest, 10 being the highest), please select your level of proficiency in understanding each language you know (write n/a if not applicable):

On a scale from 0-10 (with 0 being the lowest, 10 being the highest), please select your level of proficiency in reading each language you know (write n/a if not applicable):

In your perception, how much of a foreign accent do you have in each of your languages? (e.g. native, slightly-accented, fairly-accented, heavily-accented)

For each of your languages, please rate how frequently others identify you as a non-native speaker based on your accent? (e.g. frequently, sometimes, rarely, never)

Hawaiian language context

Whether you have ties to Hawai'i or not, describe your exposure to Hawaiian (if no exposure write none)

Whether you have ties to Hawai'i or not, describe your exposure to Pidgin (if no exposure write none)

Even if you do not consider yourself a speaker, describe the domains in which you have been exposed to Hawaiian:

Family

Friends

Appendix A. Language Background Questionnaire

Media (radio; tv; film)

Social media

School

Arts (hula; music etc)

No exposure

Other (please specify)

Even if you do not consider yourself a speaker, describe the domains in which you have been exposed to Pidgin:

Family

Friends

Media (radio; tv; film)

Social media

Cultural gatherings

School

Arts (hula; music etc)

No exposure

Other (please specify)

Language Attitudes Questionnaire

You will be asked questions regarding your language attitudes surrounding: Pidgin, Hawaiian, and English.

Please rate each statement to indicate the extent to which you agree or disagree with the question.

Score guide:

1 = Disagree strongly

2 = Disagree moderately

3 = Disagree a little

4 = Neither agree nor disagree

5 = Agree a little

6 = Agree moderately

7 = Agree strongly

Attitudes on Pidgin

Is it important for people to learn to speak Pidgin?

Is a job candidate whose dominant language is Pidgin less fit for a role as a doctor?

Is it good to teach your children Pidgin as a first language, assuming they have the opportunity to also learn other languages?

Appendix A. Language Background Questionnaire

Is the ability to speak Pidgin important to local identity?

Is a job candidate whose dominant language is Pidgin less fit for a role as a teacher?

Is it good to discourage a student from speaking Pidgin in math class?

Is it good to discourage employees from speaking Pidgin in a private office job?

Do you think it is appropriate for a governor to give a formal address in Pidgin?

Should Pidgin language education be required for keiki in Hawai'i?

Would it be a disservice to youth to teach math in Pidgin?

Open-ended: Where should people be able to speak Pidgin?

Open-ended: Where should Pidgin be forbidden to be spoken?

Open-ended: Would you support a policy designed to punish the use of Pidgin in schools?

Open-ended: Would you support a policy that promoted the hiring of Pidgin-speaking teachers?

Open-ended: What do you think of policies which prevent the use of Pidgin in court?

Attitudes on Hawaiian

Is it important for people to learn to speak Hawaiian?

Is a job candidate whose dominant language is Hawaiian less fit for a role as a doctor?

Is it good to teach your children Hawaiian as a first language, assuming they have the opportunity to also learn other languages?

Is the ability to speak Hawaiian important to local identity?

Is a job candidate whose dominant language is Hawaiian less fit for a role as a teacher?

Is it good to discourage a student from speaking Hawaiian in math class?

Is it good to discourage employees from speaking Hawaiian in a private office job?

Do you think it is appropriate for a governor to give a formal address in Hawaiian?

Should Hawaiian language education be required for keiki in Hawai'i?

Would it be a disservice to youth to teach math in Hawaiian?

Open-ended: Where should people be able to speak Hawaiian?

Open-ended: Where should Hawaiian be forbidden to be spoken?

Open-ended: Would you support a policy designed to punish the use of Hawaiian in schools?

Appendix A. Language Background Questionnaire

Open-ended: Would you support a policy that promoted the hiring of Hawaiian-speaking teachers?

Open-ended: What do you think of policies which prevent the use of Hawaiian in court?

Attitudes on English

Is it important for people to learn to speak English?

Is a job candidate whose dominant language is English less fit for a role as a doctor?

Is it good to teach your children English as a first language, granted they have the opportunity to also learn other languages?

Is the ability to speak English important to local identity?

Is a job candidate whose dominant language is English less fit for a role as a teacher?

Is it good to discourage a student from speaking English in math class?

Is it good to discourage employees from speaking English in a private office job?

Do you think it is appropriate for a governor to give a formal address in English?

Should English language education be required for keiki in Hawai'i?

Would it be a disservice to youth to teach math in English?

Open-ended: Where should people be able to speak English?

Open-ended: Where should English be forbidden to be spoken?

Open-ended: Would you support a policy designed to punish the use of English in schools?

Open-ended: Would you support a policy that promoted the hiring of English-speaking teachers?

Open-ended: What do you think of policies which prevent the use of English in court?

Attitudes on languages in the classroom

If studies demonstrated it would help students perform better, do you think it would be good to use Pidgin instead of English to teach Hawaiian?

If research suggested that students were less stressed when Pidgin was used to teach Hawaiian than when English was used to teach Hawaiian, do you think it would be a good idea to use Pidgin instead of English to teach Hawaiian?

Appendix A. Language Background Questionnaire

Would you find it more fun to learn Hawaiian in Pidgin (rather than in English)?

Would you find it more comfortable to learn Hawaiian in Pidgin (rather than in English)?

Would you find it helpful to learn Hawaiian in Pidgin (rather than in English)?

Would you find it more fun to learn Hawaiian in English (rather than in Pidgin)?

Would you find it more comfortable to learn Hawaiian in English (rather than in Pidgin)?

Would you find it helpful to learn Hawaiian in English (rather than in Pidgin)?

Is Pidgin difficult to learn?

Is English difficult to learn?

Is Hawaiian difficult to learn?

Open-ended: What would you think about a teacher that used Pidgin (instead of English) to teach Hawaiian?

Open-ended: What could you see as benefits of using Pidgin (instead of English) to teach Hawaiian?

Open-ended: Do you see any downsides of using Pidgin (instead of English) to teach Hawaiian?

Appendix B

Experimental Stimuli: Experiment 1

Target Word	Target Language	Frame Language	Choice A	Gloss A	Choice B	Gloss B	Contrast Type	Pair Type
bolo	pid	pid	pua'a	pig	bolo	bald-head	interlingual distractor	morphologically unrelated
kakimochi	pid	pid	kumu	teacher	kakimochi	rice-cracker	interlingual distractor	morphologically unrelated
musubi	pid	pid	musubi	riceball	mo'o	lizard	interlingual distractor	morphologically unrelated
lihingmui	pid	pid	lima	hand	lihingmui	plum-seed	interlingual distractor	morphologically unrelated
kaukau	pid	pid	kaukau	food	keko	monkey	interlingual distractor	morphologically unrelated
lychee	pid	pid	la'au	tree	lychee	lychee	interlingual distractor	morphologically unrelated
moon	eng	eng	moa	chicken	moon	moon	interlingual distractor	morphologically unrelated
nose	eng	eng	nose	nose	niho	tooth	interlingual distractor	morphologically unrelated
mango	eng	eng	mano	shark	mango	mango	interlingual distractor	morphologically unrelated
pen	eng	eng	pepeiao	ears	pen	pen	interlingual distractor	morphologically unrelated
match	eng	eng	match	match	mai'a	banana	interlingual distractor	morphologically unrelated
ant	eng	eng	ant	ant	'anela	angel	interlingual distractor	morphologically unrelated
apples	pid	pid	apple	apple	apples	apples	interlingual distractor	morphologically related
men	eng	eng	father	father	fathers	fathers	interlingual distractor	morphologically related
lizard	pid	pid	lizard	lizard	lizards	lizards	interlingual distractor	morphologically related
women	eng	eng	woman	woman	women	women	interlingual distractor	morphologically related
girl	pid	pid	girl	girl	girls	girls	interlingual distractor	morphologically related
doctors	eng	eng	doctor	doctor	doctors	doctors	interlingual distractor	morphologically related

Table B.1: Unilingual Stimuli List (Control Items, Identical Across Participants)

Appendix C

Experimental Stimuli: Experiment 2

Table C.1: List 1 cross-language priming task stimuli

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	undies	palema‘i	word	word	critical
ENG-‘ōH	big	nui	word	word	critical
ENG-‘ōH	toad	poloka	word	word	critical
ENG-‘ōH	soilderfish	‘ū‘ū	word	word	critical
ENG-‘ōH	idiot	hūpō	word	word	critical
ENG-‘ōH	pompous	ho‘oki‘eki‘e	word	word	critical
ENG-‘ōH	breasts	waiū	word	word	critical
ENG-‘ōH	badass	loea	word	word	critical
ENG-‘ōH	shame	hilahila	word	word	critical
ENG-‘ōH	pig	pua‘a	word	word	critical
ENG-‘ōH	unkempt	‘eke	word	word	critical
ENG-‘ōH	food	‘ai	word	word	critical
ENG-‘ōH	across	palahē	word	word	control
ENG-‘ōH	lent	‘elelū	word	word	control
ENG-‘ōH	nutmeg	mūhe‘e	word	word	control
ENG-‘ōH	damaging	kāpulu	word	word	control
ENG-‘ōH	font	pilo	word	word	control
ENG-‘ōH	walnut	lapu	word	word	control
ENG-‘ōH	repeat	auwē	word	word	control
ENG-‘ōH	squirrel	ho‘opailua	word	word	control
ENG-‘ōH	slouch	wahahe‘e	word	word	control
ENG-‘ōH	joint	wela	word	word	control
ENG-‘ōH	heavy	mele	word	word	control
ENG-‘ōH	greedy	hiki	word	word	control

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	sporty	anu	word	word	filler
ENG-‘ōH	fruit	ānuenu	word	word	filler
ENG-‘ōH	hide	‘āpana	word	word	filler
ENG-‘ōH	ugly	‘āpōpō	word	word	filler
ENG-‘ōH	wooden	aupuni	word	word	filler
ENG-‘ōH	precise	nu‘u	word	word	filler
ENG-‘ōH	jam	‘ōhai	word	word	filler
ENG-‘ōH	bush	‘eleu	word	word	filler
ENG-‘ōH	contest	‘ōiwi	word	word	filler
ENG-‘ōH	jacket	weli	word	word	filler
ENG-‘ōH	toenail	‘awa‘awa	word	word	filler
ENG-‘ōH	daring	anapuni	word	word	filler
ENG-‘ōH	upset	‘iniha	word	word	filler
ENG-‘ōH	laps	paniolo	word	word	filler
ENG-‘ōH	napkin	hōkū	word	word	filler
ENG-‘ōH	blanket	peni	word	word	filler
ENG-‘ōH	vaking	imu	nonword	word	filler
ENG-‘ōH	fluko	puku	nonword	word	filler
ENG-‘ōH	berdict	iwakālua	nonword	word	filler
ENG-‘ōH	heport	kaha	nonword	word	filler
ENG-‘ōH	trahze	kāma‘a	nonword	word	filler
ENG-‘ōH	apess	kauka	nonword	word	filler
ENG-‘ōH	recooled	noho	nonword	word	filler
ENG-‘ōH	pyrics	kua‘āina	nonword	word	filler
ENG-‘ōH	prallpox	lā‘au	nonword	word	filler
ENG-‘ōH	chying	makemake	nonword	word	filler
ENG-‘ōH	suve	kupake‘e	nonword	word	filler
ENG-‘ōH	grists	kāpī	nonword	word	filler
ENG-‘ōH	toffin	hō‘ai	nonword	word	filler
ENG-‘ōH	larned	wai	nonword	word	filler
ENG-‘ōH	sargain	waha	nonword	word	filler
ENG-‘ōH	grampled	moa	nonword	word	filler
ENG-‘ōH	humbup	mai‘a	nonword	word	filler
ENG-‘ōH	drails	maka	nonword	word	filler
ENG-‘ōH	nall	pāpāleo	nonword	word	filler
ENG-‘ōH	gubplot	maka‘aniani	nonword	word	filler
ENG-‘ōH	vate	nane	nonword	word	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	blantly	niho	nonword	word	filler
ENG-‘ōH	vodless	oli	nonword	word	filler
ENG-‘ōH	freelt	pololū	nonword	word	filler
ENG-‘ōH	trush	‘ōma‘oma‘o	nonword	word	filler
ENG-‘ōH	uldock	one	nonword	word	filler
ENG-‘ōH	bives	‘elepani	nonword	word	filler
ENG-‘ōH	stendthrift	‘ahua	nonword	word	filler
ENG-‘ōH	snart	pālule	nonword	word	filler
ENG-‘ōH	warmip	kūkini	nonword	word	filler
ENG-‘ōH	latty	‘ula	nonword	word	filler
ENG-‘ōH	brax	‘ulu	nonword	word	filler
ENG-‘ōH	rainted	kūli‘u	nonword	word	filler
ENG-‘ōH	rigers	huā	nonword	word	filler
ENG-‘ōH	souch	moloā	nonword	word	filler
ENG-‘ōH	loophile	huihui	nonword	word	filler
ENG-‘ōH	larfy	pake	nonword	word	filler
ENG-‘ōH	clat	‘āko‘ako‘a	nonword	word	filler
ENG-‘ōH	splings	pa‘akikī	nonword	word	filler
ENG-‘ōH	torttheast	nalū	nonword	word	filler
ENG-‘ōH	zothy	hemā	nonword	nonword	filler
ENG-‘ōH	ipator	lalo	nonword	nonword	filler
ENG-‘ōH	clitten	lānou	nonword	nonword	filler
ENG-‘ōH	sermanent	noike	nonword	nonword	filler
ENG-‘ōH	pirebreaks	nokoi	nonword	nonword	filler
ENG-‘ōH	rampagins	pumo	nonword	nonword	filler
ENG-‘ōH	pherub	wailei	nonword	nonword	filler
ENG-‘ōH	soalhole	hamule	nonword	nonword	filler
ENG-‘ōH	sleeches	hauaho	nonword	nonword	filler
ENG-‘ōH	selevised	hoika‘i	nonword	nonword	filler
ENG-‘ōH	klereafter	hoikaka	nonword	nonword	filler
ENG-‘ōH	chimena	‘o‘ā‘o	nonword	nonword	filler
ENG-‘ōH	huspicious	‘oika‘i	nonword	nonword	filler
ENG-‘ōH	confussor	kahino	nonword	nonword	filler
ENG-‘ōH	interfice	kahono	nonword	nonword	filler
ENG-‘ōH	dearrange	ke‘aule	nonword	nonword	filler
ENG-‘ōH	pilitant	kekaeo	nonword	nonword	filler
ENG-‘ōH	milak	lālia	nonword	nonword	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	thawl	lawaihae	nonword	nonword	filler
ENG-‘ōH	tinds	manani	nonword	nonword	filler
ENG-‘ōH	crunt	manaula	nonword	nonword	filler
ENG-‘ōH	borm	nolai‘a	nonword	nonword	filler
ENG-‘ōH	dorst	nualo	nonword	nonword	filler
ENG-‘ōH	prall	tahaku	nonword	nonword	filler
ENG-‘ōH	gans	uhoho	nonword	nonword	filler
ENG-‘ōH	geeds	ha‘aliho	nonword	nonword	filler
ENG-‘ōH	slore	ha‘alima	nonword	nonword	filler
ENG-‘ōH	mufts	ho‘ilona	nonword	nonword	filler
ENG-‘ōH	rogged	ho‘olona	nonword	nonword	filler
ENG-‘ōH	hakes	ilahina	nonword	nonword	filler
ENG-‘ōH	plapps	ilaka‘i	nonword	nonword	filler
ENG-‘ōH	bips	ulo	nonword	nonword	filler
ENG-‘ōH	mog	mela	nonword	nonword	filler
ENG-‘ōH	klow	‘iko	nonword	nonword	filler
ENG-‘ōH	yoles	miao	nonword	nonword	filler
ENG-‘ōH	jelt	hanō	nonword	nonword	filler
ENG-‘ōH	hivs	wile	nonword	nonword	filler
ENG-‘ōH	fraft	mape	nonword	nonword	filler
ENG-‘ōH	wope	lewo	nonword	nonword	filler
ENG-‘ōH	velts	pe‘i	nonword	nonword	filler
ENG-‘ōH	wannabee	hokou	word	nonword	filler
ENG-‘ōH	lift	la‘ū	word	nonword	filler
ENG-‘ōH	rash	lauhi	word	nonword	filler
ENG-‘ōH	shut	mō‘ū	word	nonword	filler
ENG-‘ōH	weirdo	nonai	word	nonword	filler
ENG-‘ōH	thickener	poine	word	nonword	filler
ENG-‘ōH	fuss	wepo	word	nonword	filler
ENG-‘ōH	beautiful	halaini	word	nonword	filler
ENG-‘ōH	actually	hauauao	word	nonword	filler
ENG-‘ōH	nuts	hō‘iku	word	nonword	filler
ENG-‘ōH	core	hoikīkī	word	nonword	filler
ENG-‘ōH	forks	‘elihi	word	nonword	filler
ENG-‘ōH	woo	‘oikeia	word	nonword	filler
ENG-‘ōH	plea	iolai	word	nonword	filler
ENG-‘ōH	brook	kaikamoi	word	nonword	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	finger	kehua	word	nonword	filler
ENG-‘ōH	jazz	kilaina	word	nonword	filler
ENG-‘ōH	corpse	kulaine	word	nonword	filler
ENG-‘ōH	farewell	liola	word	nonword	filler
ENG-‘ōH	ramp	maluni	word	nonword	filler
ENG-‘ōH	melt	mawai‘o	word	nonword	filler
ENG-‘ōH	diarrhea	nahiko	word	nonword	filler
ENG-‘ōH	nerdy	nuana	word	nonword	filler
ENG-‘ōH	itchy	pohana	word	nonword	filler
ENG-‘ōH	fault	uka‘a	word	nonword	filler
ENG-‘ōH	centipede	eneloko	word	nonword	filler
ENG-‘ōH	colorful	halaine	word	nonword	filler
ENG-‘ōH	shook	hoapi‘o	word	nonword	filler
ENG-‘ōH	fun	holikoki	word	nonword	filler
ENG-‘ōH	poor	‘oluhele	word	nonword	filler
ENG-‘ōH	flex	ilamo‘o	word	nonword	filler
ENG-‘ōH	slang	u‘o	word	nonword	filler
ENG-‘ōH	forever	aonu	word	nonword	filler
ENG-‘ōH	mesh	kū‘u	word	nonword	filler
ENG-‘ōH	billed	wemo	word	nonword	filler
ENG-‘ōH	slips	lehā	word	nonword	filler
ENG-‘ōH	jewel	mōpa	word	nonword	filler
ENG-‘ōH	deed	kawo	word	nonword	filler
ENG-‘ōH	medicine	lānu	word	nonword	filler
ENG-‘ōH	wings	lipe	word	nonword	filler
PID-‘ōH	schmolkin	iki	word	word	critical
PID-‘ōH	boroboros	weluwelu	word	word	critical
PID-‘ōH	lavalava	kikepa	word	word	critical
PID-‘ōH	hanabata	hūpē	word	word	critical
PID-‘ōH	bolo	‘ōhule	word	word	critical
PID-‘ōH	bakatare	hepa	word	word	critical
PID-‘ōH	tsubu	pipipi	word	word	critical
PID-‘ōH	bocha	‘au‘au	word	word	critical
PID-‘ōH	galas	male	word	word	critical
PID-‘ōH	nainai	hiamoe	word	word	critical
PID-‘ōH	benjo	lua	word	word	critical
PID-‘ōH	tantaran	hō‘oio	word	word	critical

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	haolefied	‘auane‘i	word	word	control
PID-‘ōH	bagoong	hūtā	word	word	control
PID-‘ōH	ala-alas	makepono	word	word	control
PID-‘ōH	futless	nuha	word	word	control
PID-‘ōH	had-it	mimi	word	word	control
PID-‘ōH	trow	kapakahi	word	word	control
PID-‘ōH	horrahz	pohō	word	word	control
PID-‘ōH	oddah	hoaloha	word	word	control
PID-‘ōH	nuff	‘eha	word	word	control
PID-‘ōH	hanakuso	‘u‘uku	word	word	control
PID-‘ōH	kamaboko	pehea	word	word	control
PID-‘ōH	moke	mea	word	word	control
PID-‘ōH	popo	lani	word	word	filler
PID-‘ōH	braddah	wini	word	word	filler
PID-‘ōH	musubi	uluna	word	word	filler
PID-‘ōH	usetato	no‘ono‘o	word	word	filler
PID-‘ōH	lihingmui	kūkaepele	word	word	filler
PID-‘ōH	humbug	‘īlio	word	word	filler
PID-‘ōH	kden	kō‘ai	word	word	filler
PID-‘ōH	kotonk	kane	word	word	filler
PID-‘ōH	buddahead	honu	word	word	filler
PID-‘ōH	yobo	nohea	word	word	filler
PID-‘ōH	malasada	ho‘oilo	word	word	filler
PID-‘ōH	li‘dat	kūpuna	word	word	filler
PID-‘ōH	bento	‘i‘ike	word	word	filler
PID-‘ōH	obakechocho	lelele	word	word	filler
PID-‘ōH	no‘mo	hālau	word	word	filler
PID-‘ōH	planny	‘ahi	word	word	filler
PID-‘ōH	bito	kāhoaka	nonword	word	filler
PID-‘ōH	momong	pupule	nonword	word	filler
PID-‘ōH	banak	lapuwale	nonword	word	filler
PID-‘ōH	mauki	ka‘a	nonword	word	filler
PID-‘ōH	dinngah	wiwo‘ole	nonword	word	filler
PID-‘ōH	uk‘ume	iwi	nonword	word	filler
PID-‘ōH	rehut	laha	nonword	word	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	behua	‘ele‘ele	nonword	word	filler
PID-‘ōH	goke	‘ukulele	nonword	word	filler
PID-‘ōH	tanah	makika	nonword	word	filler
PID-‘ōH	bossabossa	liliko‘i	nonword	word	filler
PID-‘ōH	darehloco	alia	nonword	word	filler
PID-‘ōH	kaoyu	paka	nonword	word	filler
PID-‘ōH	yote	ho‘olana	nonword	word	filler
PID-‘ōH	lélau	momona	nonword	word	filler
PID-‘ōH	hakakattack	niu	nonword	word	filler
PID-‘ōH	ainez	‘a‘ā	nonword	word	filler
PID-‘ōH	tapax	hau‘oli	nonword	word	filler
PID-‘ōH	raoks	kīna‘u	nonword	word	filler
PID-‘ōH	haoka	‘ōpū	nonword	word	filler
PID-‘ōH	rakeakai	apo	nonword	word	filler
PID-‘ōH	tana-hina-loa	piapia	nonword	word	filler
PID-‘ōH	bakut	naonao	nonword	word	filler
PID-‘ōH	makayoinz	kūpa‘a	nonword	word	filler
PID-‘ōH	zone	pika	nonword	word	filler
PID-‘ōH	racha	‘ili‘ili	nonword	word	filler
PID-‘ōH	‘aukenui	manini	nonword	word	filler
PID-‘ōH	‘ajahz	mano	nonword	word	filler
PID-‘ōH	hahap	pāpale	nonword	word	filler
PID-‘ōH	hinakahi	wāwahi	nonword	word	filler
PID-‘ōH	rakus	hāleu	nonword	word	filler
PID-‘ōH	manoz	haumāna	nonword	word	filler
PID-‘ōH	irratoa	huluhulu	nonword	word	filler
PID-‘ōH	blaloha	ikaika	nonword	word	filler
PID-‘ōH	tingahkali	kakani	nonword	word	filler
PID-‘ōH	kenat	kālai	nonword	word	filler
PID-‘ōH	‘aohalaniz	ke‘ena	nonword	word	filler
PID-‘ōH	rapax	pehu	nonword	word	filler
PID-‘ōH	pinat	ka‘ā	nonword	word	filler
PID-‘ōH	majax	lomi	nonword	word	filler
PID-‘ōH	okola	‘ā‘au	nonword	nonword	filler
PID-‘ōH	paco	kānoi	nonword	nonword	filler
PID-‘ōH	majai	maula	nonword	nonword	filler
PID-‘ōH	mokat	mo‘u	nonword	nonword	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	lelax	numo	nonword	nonword	filler
PID-‘ōH	niebe	poihi	nonword	nonword	filler
PID-‘ōH	tanax	a‘oua	nonword	nonword	filler
PID-‘ōH	rajak	halaili	nonword	nonword	filler
PID-‘ōH	bapah	hauaula	nonword	nonword	filler
PID-‘ōH	heffahead	hīmena	nonword	nonword	filler
PID-‘ōH	latonk	hukoa	nonword	nonword	filler
PID-‘ōH	yonore	‘ekane	nonword	nonword	filler
PID-‘ōH	kachi	‘oko	nonword	nonword	filler
PID-‘ōH	nehualoa	ileia	nonword	nonword	filler
PID-‘ōH	retang	kalawi	nonword	nonword	filler
PID-‘ōH	beno	kehea	nonword	nonword	filler
PID-‘ōH	gote	kilawai	nonword	nonword	filler
PID-‘ōH	humotrillz	kuhiai	nonword	nonword	filler
PID-‘ōH	maoki	loawa	nonword	nonword	filler
PID-‘ōH	lenau	malauao	nonword	nonword	filler
PID-‘ōH	scrai	moake	nonword	nonword	filler
PID-‘ōH	okebe	muliwao	nonword	nonword	filler
PID-‘ōH	haolakea	ohoko	nonword	nonword	filler
PID-‘ōH	ohakaz	pilawao	nonword	nonword	filler
PID-‘ōH	hanataran	unena	nonword	nonword	filler
PID-‘ōH	mevai	auahina	nonword	nonword	filler
PID-‘ōH	baike	halauea	nonword	nonword	filler
PID-‘ōH	gobotani	hoāli‘i	nonword	nonword	filler
PID-‘ōH	lotulofa	holu‘olo	nonword	nonword	filler
PID-‘ōH	banai	‘olowale	nonword	nonword	filler
PID-‘ōH	mokau	ilauloa	nonword	nonword	filler
PID-‘ōH	hapak	ewi	nonword	nonword	filler
PID-‘ōH	kokat	pāmo	nonword	nonword	filler
PID-‘ōH	tanakuso	oi‘u	nonword	nonword	filler
PID-‘ōH	latona	neli	nonword	nonword	filler
PID-‘ōH	pakagi	homi	nonword	nonword	filler
PID-‘ōH	pakamaka	ao‘u	nonword	nonword	filler
PID-‘ōH	mokiwiki	wiai	nonword	nonword	filler
PID-‘ōH	lauku	wemi	nonword	nonword	filler
PID-‘ōH	blammajangah	hiho	nonword	nonword	filler
PID-‘ōH	potagee	‘āpau	word	nonword	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	killahweeffah	‘ihu	word	nonword	filler
PID-‘ōH	shaka	mauou	word	nonword	filler
PID-‘ōH	saimin	mō‘ao	word	nonword	filler
PID-‘ōH	arare	pamae	word	nonword	filler
PID-‘ōH	locomoco	pemu	word	nonword	filler
PID-‘ōH	chair	anaila	word	nonword	filler
PID-‘ōH	ume	eleli	word	nonword	filler
PID-‘ōH	sleepahs	hawai‘o	word	nonword	filler
PID-‘ōH	aiyah	hikili	word	nonword	filler
PID-‘ōH	buckaloose	‘a‘ane	word	nonword	filler
PID-‘ōH	figga	‘ekaeo	word	nonword	filler
PID-‘ōH	jankenpo	‘ōlepa	word	nonword	filler
PID-‘ōH	manapua	ilawā	word	nonword	filler
PID-‘ōH	mokemobile	kānua	word	nonword	filler
PID-‘ōH	nails	kaweni	word	nonword	filler
PID-‘ōH	pukashirt	kohalo	word	nonword	filler
PID-‘ōH	pukahead	kuhalo	word	nonword	filler
PID-‘ōH	puating	lō‘ike	word	nonword	filler
PID-‘ōH	ainokea	mailoha	word	nonword	filler
PID-‘ōH	shootz	mohua	word	nonword	filler
PID-‘ōH	shoowah	molomau	word	nonword	filler
PID-‘ōH	rajah	oihaka	word	nonword	filler
PID-‘ōH	chout	pilai‘o	word	nonword	filler
PID-‘ōH	ukubillionz	wahike	word	nonword	filler
PID-‘ōH	cah	waihale	word	nonword	filler
PID-‘ōH	faddah	hale‘ele	word	nonword	filler
PID-‘ōH	termometah	hilinahi	word	nonword	filler
PID-‘ōH	brokdamout	huaoli‘e	word	nonword	filler
PID-‘ōH	hashi	‘oiapa‘a	word	nonword	filler
PID-‘ōH	mo‘bettah	kahaila	word	nonword	filler
PID-‘ōH	buggah	kape	word	nonword	filler
PID-‘ōH	girigiri	miwo	word	nonword	filler
PID-‘ōH	bondance	ha‘ī	word	nonword	filler
PID-‘ōH	geevum	wano	word	nonword	filler
PID-‘ōH	calabash	aoku	word	nonword	filler
PID-‘ōH	irraz	me‘u	word	nonword	filler
PID-‘ōH	manong	lapi	word	nonword	filler
PID-‘ōH	sole	nule	word	nonword	filler

Table C.1 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	vog	māne	word	nonword	filler

Table C.2: List 2 cross-language priming task stimuli

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	bibadeez	palema‘i	word	word	critical
PID-‘ōH	bambucha	nui	word	word	critical
PID-‘ōH	bufo	poloka	word	word	critical
PID-‘ōH	menpachi	‘ū‘ū	word	word	critical
PID-‘ōH	babooze	hūpō	word	word	critical
PID-‘ōH	highmakamaka	ho‘oki‘eki‘e	word	word	critical
PID-‘ōH	chichis	waiū	word	word	critical
PID-‘ōH	hammah	loea	word	word	critical
PID-‘ōH	halā	hilahila	word	word	critical
PID-‘ōH	buta	pua‘a	word	word	critical
PID-‘ōH	hash	‘eke	word	word	critical
PID-‘ōH	kaukau	‘ai	word	word	critical
PID-‘ōH	haolefied	palahē	word	word	control
PID-‘ōH	bagoong	‘elelū	word	word	control
PID-‘ōH	ala-alas	mūhe‘e	word	word	control
PID-‘ōH	futless	kāpulu	word	word	control
PID-‘ōH	had-it	pilo	word	word	control
PID-‘ōH	trow	lapu	word	word	control
PID-‘ōH	horrahz	auwē	word	word	control
PID-‘ōH	oddah	ho‘opailua	word	word	control
PID-‘ōH	nuff	wahahe‘e	word	word	control
PID-‘ōH	hanakuso	wela	word	word	control
PID-‘ōH	kamaboko	mele	word	word	control
PID-‘ōH	moke	hiki	word	word	control
ENG-‘ōH	sporty	anu	word	word	filler
ENG-‘ōH	fruit	ānuenue	word	word	filler
ENG-‘ōH	hide	‘āpana	word	word	filler
ENG-‘ōH	ugly	‘āpōpō	word	word	filler
ENG-‘ōH	wooden	aupuni	word	word	filler
ENG-‘ōH	precise	nu‘u	word	word	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	jam	‘ōhai	word	word	filler
ENG-‘ōH	bush	‘eleu	word	word	filler
ENG-‘ōH	contest	‘ōiwi	word	word	filler
ENG-‘ōH	jacket	weli	word	word	filler
ENG-‘ōH	toenail	‘awa‘awa	word	word	filler
ENG-‘ōH	daring	anapuni	word	word	filler
ENG-‘ōH	upset	‘iniha	word	word	filler
ENG-‘ōH	laps	paniolo	word	word	filler
ENG-‘ōH	napkin	hōkū	word	word	filler
ENG-‘ōH	blanket	peni	word	word	filler
ENG-‘ōH	vaking	imu	nonword	word	filler
ENG-‘ōH	fluko	puku	nonword	word	filler
ENG-‘ōH	berdict	iwakālua	nonword	word	filler
ENG-‘ōH	heport	kaha	nonword	word	filler
ENG-‘ōH	trahze	kāma‘a	nonword	word	filler
ENG-‘ōH	apess	kauka	nonword	word	filler
ENG-‘ōH	recooled	noho	nonword	word	filler
ENG-‘ōH	pyrics	kua‘āina	nonword	word	filler
ENG-‘ōH	prallpox	lā‘au	nonword	word	filler
ENG-‘ōH	chying	makemake	nonword	word	filler
ENG-‘ōH	suve	kupake‘e	nonword	word	filler
ENG-‘ōH	grists	kāpī	nonword	word	filler
ENG-‘ōH	toffin	hō‘ai	nonword	word	filler
ENG-‘ōH	larned	wai	nonword	word	filler
ENG-‘ōH	sargain	waha	nonword	word	filler
ENG-‘ōH	grampled	moa	nonword	word	filler
ENG-‘ōH	humbup	mai‘a	nonword	word	filler
ENG-‘ōH	drails	maka	nonword	word	filler
ENG-‘ōH	nall	pāpāleo	nonword	word	filler
ENG-‘ōH	gubplot	maka‘aniani	nonword	word	filler
ENG-‘ōH	vate	nane	nonword	word	filler
ENG-‘ōH	blantly	niho	nonword	word	filler
ENG-‘ōH	vodless	oli	nonword	word	filler
ENG-‘ōH	freelt	pololū	nonword	word	filler
ENG-‘ōH	trush	‘ōma‘oma‘o	nonword	word	filler
ENG-‘ōH	uldock	one	nonword	word	filler
ENG-‘ōH	bives	‘elepani	nonword	word	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	stendthrift	‘ahua	nonword	word	filler
ENG-‘ōH	snart	pālule	nonword	word	filler
ENG-‘ōH	warmip	kūkini	nonword	word	filler
ENG-‘ōH	latty	‘ula	nonword	word	filler
ENG-‘ōH	brax	‘ulu	nonword	word	filler
ENG-‘ōH	rainted	kūli‘u	nonword	word	filler
ENG-‘ōH	rigers	hua	nonword	word	filler
ENG-‘ōH	souch	moloā	nonword	word	filler
ENG-‘ōH	loophile	huihui	nonword	word	filler
ENG-‘ōH	larfy	pake	nonword	word	filler
ENG-‘ōH	clat	‘āko‘ako‘a	nonword	word	filler
ENG-‘ōH	splings	pa‘akikī	nonword	word	filler
ENG-‘ōH	torttheast	nalu	nonword	word	filler
ENG-‘ōH	zothy	hemā	nonword	nonword	filler
ENG-‘ōH	ipatord	laloī	nonword	nonword	filler
ENG-‘ōH	clitten	lānou	nonword	nonword	filler
ENG-‘ōH	sermanent	noike	nonword	nonword	filler
ENG-‘ōH	pirebreaks	nokoi	nonword	nonword	filler
ENG-‘ōH	rampagins	pumo	nonword	nonword	filler
ENG-‘ōH	pherub	wailei	nonword	nonword	filler
ENG-‘ōH	soalhole	hamule	nonword	nonword	filler
ENG-‘ōH	sleeches	hauaho	nonword	nonword	filler
ENG-‘ōH	selevised	hoika‘i	nonword	nonword	filler
ENG-‘ōH	klereafter	hoikaka	nonword	nonword	filler
ENG-‘ōH	chimena	‘o‘ā‘o	nonword	nonword	filler
ENG-‘ōH	huspicious	‘oika‘i	nonword	nonword	filler
ENG-‘ōH	confussor	kahino	nonword	nonword	filler
ENG-‘ōH	interfice	kahono	nonword	nonword	filler
ENG-‘ōH	dearrange	ke‘aule	nonword	nonword	filler
ENG-‘ōH	pilitant	kekaeo	nonword	nonword	filler
ENG-‘ōH	milak	lālia	nonword	nonword	filler
ENG-‘ōH	thawl	lawaihae	nonword	nonword	filler
ENG-‘ōH	tinds	manani	nonword	nonword	filler
ENG-‘ōH	crunt	manaula	nonword	nonword	filler
ENG-‘ōH	borm	nolai‘a	nonword	nonword	filler
ENG-‘ōH	dorst	nualo	nonword	nonword	filler
ENG-‘ōH	prall	tahaku	nonword	nonword	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	gans	uhoho	nonword	nonword	filler
ENG-‘ōH	geeds	ha‘aliho	nonword	nonword	filler
ENG-‘ōH	slore	ha‘alima	nonword	nonword	filler
ENG-‘ōH	mufts	ho‘ilona	nonword	nonword	filler
ENG-‘ōH	rogged	ho‘olona	nonword	nonword	filler
ENG-‘ōH	hakes	ilahina	nonword	nonword	filler
ENG-‘ōH	plapps	ilaka‘i	nonword	nonword	filler
ENG-‘ōH	bips	ulo	nonword	nonword	filler
ENG-‘ōH	mog	melā	nonword	nonword	filler
ENG-‘ōH	klow	‘iko	nonword	nonword	filler
ENG-‘ōH	yoles	miao	nonword	nonword	filler
ENG-‘ōH	jelt	hanō	nonword	nonword	filler
ENG-‘ōH	hivs	wile	nonword	nonword	filler
ENG-‘ōH	fraft	mapē	nonword	nonword	filler
ENG-‘ōH	wope	lewo	nonword	nonword	filler
ENG-‘ōH	velts	pe‘i	nonword	nonword	filler
ENG-‘ōH	wannabee	hokou	word	nonword	filler
ENG-‘ōH	lift	la‘ū	word	nonword	filler
ENG-‘ōH	rash	lauhi	word	nonword	filler
ENG-‘ōH	shut	mō‘ū	word	nonword	filler
ENG-‘ōH	weirdo	nonai	word	nonword	filler
ENG-‘ōH	thickener	poine	word	nonword	filler
ENG-‘ōH	fuss	wepo	word	nonword	filler
ENG-‘ōH	beautiful	halaini	word	nonword	filler
ENG-‘ōH	actually	hauauao	word	nonword	filler
ENG-‘ōH	nuts	hō‘iku	word	nonword	filler
ENG-‘ōH	core	hoikīkī	word	nonword	filler
ENG-‘ōH	forks	‘elihi	word	nonword	filler
ENG-‘ōH	woo	‘oikeia	word	nonword	filler
ENG-‘ōH	plea	iolai	word	nonword	filler
ENG-‘ōH	brook	kaikamoi	word	nonword	filler
ENG-‘ōH	finger	kehua	word	nonword	filler
ENG-‘ōH	jazz	kilaina	word	nonword	filler
ENG-‘ōH	corpse	kulaine	word	nonword	filler
ENG-‘ōH	farewell	liola	word	nonword	filler
ENG-‘ōH	ramp	maluni	word	nonword	filler
ENG-‘ōH	melt	mawai‘o	word	nonword	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	diarrhea	nahiko	word	nonword	filler
ENG-‘ōH	nerdy	nuana	word	nonword	filler
ENG-‘ōH	itchy	pohana	word	nonword	filler
ENG-‘ōH	fault	uka‘a	word	nonword	filler
ENG-‘ōH	centipede	eneloko	word	nonword	filler
ENG-‘ōH	colorful	halaine	word	nonword	filler
ENG-‘ōH	shook	hoapi‘o	word	nonword	filler
ENG-‘ōH	fun	holikoki	word	nonword	filler
ENG-‘ōH	poor	‘oluhele	word	nonword	filler
ENG-‘ōH	flex	ilamo‘o	word	nonword	filler
ENG-‘ōH	slang	u‘o	word	nonword	filler
ENG-‘ōH	forever	aonu	word	nonword	filler
ENG-‘ōH	mesh	kū‘u	word	nonword	filler
ENG-‘ōH	billed	wemo	word	nonword	filler
ENG-‘ōH	slips	lehā	word	nonword	filler
ENG-‘ōH	jewel	mōpa	word	nonword	filler
ENG-‘ōH	deed	kawo	word	nonword	filler
ENG-‘ōH	medicine	lānu	word	nonword	filler
ENG-‘ōH	wings	lipe	word	nonword	filler
ENG-‘ōH	little	iki	word	word	critical
ENG-‘ōH	rags	weluwelu	word	word	critical
ENG-‘ōH	sarong	kīkepa	word	word	critical
ENG-‘ōH	mucus	hūpē	word	word	critical
ENG-‘ōH	baldy	‘ōhule	word	word	critical
ENG-‘ōH	stupid	hepa	word	word	critical
ENG-‘ōH	seasnail	pipipi	word	word	critical
ENG-‘ōH	bath	‘au‘au	word	word	critical
ENG-‘ōH	phlegm	male	word	word	critical
ENG-‘ōH	sleep	hiamoe	word	word	critical
ENG-‘ōH	toilet	lua	word	word	critical
ENG-‘ōH	cocky	hō‘oio	word	word	critical
ENG-‘ōH	across	‘auane‘i	word	word	control
ENG-‘ōH	lent	hūtā	word	word	control
ENG-‘ōH	nutmeg	makepono	word	word	control
ENG-‘ōH	damaging	nuha	word	word	control
ENG-‘ōH	font	mimi	word	word	control

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	walnut	kapakahi	word	word	control
ENG-‘ōH	repeat	pohō	word	word	control
ENG-‘ōH	squirrel	hoaloha	word	word	control
ENG-‘ōH	slouch	‘eha	word	word	control
ENG-‘ōH	joint	‘u‘uku	word	word	control
ENG-‘ōH	heavy	pehea	word	word	control
ENG-‘ōH	greedy	mea	word	word	control
PID-‘ōH	popo	lani	word	word	filler
PID-‘ōH	braddah	wini	word	word	filler
PID-‘ōH	musubi	uluna	word	word	filler
PID-‘ōH	usetato	no‘ono‘o	word	word	filler
PID-‘ōH	lihingmui	kūkaepele	word	word	filler
PID-‘ōH	humbug	‘ilio	word	word	filler
PID-‘ōH	kden	kō‘ai	word	word	filler
PID-‘ōH	kotonk	kane	word	word	filler
PID-‘ōH	buddahead	honu	word	word	filler
PID-‘ōH	yobo	nohea	word	word	filler
PID-‘ōH	malasada	ho‘oilo	word	word	filler
PID-‘ōH	li‘dat	kūpuna	word	word	filler
PID-‘ōH	bento	‘i‘ike	word	word	filler
PID-‘ōH	obakechocho	lelele	word	word	filler
PID-‘ōH	no‘mo	hālau	word	word	filler
PID-‘ōH	planny	‘ahi	word	word	filler
PID-‘ōH	bito	kāhoaka	nonword	word	filler
PID-‘ōH	momong	pupule	nonword	word	filler
PID-‘ōH	banak	lapuwale	nonword	word	filler
PID-‘ōH	mauki	ka‘a	nonword	word	filler
PID-‘ōH	dinngah	wiwo‘ole	nonword	word	filler
PID-‘ōH	uk‘ume	iwi	nonword	word	filler
PID-‘ōH	rehut	laha	nonword	word	filler
PID-‘ōH	behua	‘ele‘ele	nonword	word	filler
PID-‘ōH	goke	‘ukulele	nonword	word	filler
PID-‘ōH	tanah	makika	nonword	word	filler
PID-‘ōH	bossabossa	liliko‘i	nonword	word	filler
PID-‘ōH	darehloco	alia	nonword	word	filler
PID-‘ōH	kaoyu	paka	nonword	word	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	yote	ho‘olana	nonword	word	filler
PID-‘ōH	lelau	momona	nonword	word	filler
PID-‘ōH	hakakattack	niu	nonword	word	filler
PID-‘ōH	ainez	‘a‘ā	nonword	word	filler
PID-‘ōH	tapax	hau‘oli	nonword	word	filler
PID-‘ōH	raoks	kīna‘u	nonword	word	filler
PID-‘ōH	haoka	‘ōpū	nonword	word	filler
PID-‘ōH	rakeakai	apo	nonword	word	filler
PID-‘ōH	tana-hina-loa	piapia	nonword	word	filler
PID-‘ōH	bakut	naonao	nonword	word	filler
PID-‘ōH	makayoinz	kūpa‘a	nonword	word	filler
PID-‘ōH	zone	pika	nonword	word	filler
PID-‘ōH	racha	‘ili‘ili	nonword	word	filler
PID-‘ōH	‘aukenui	manini	nonword	word	filler
PID-‘ōH	‘ajahz	mano	nonword	word	filler
PID-‘ōH	hahap	pāpale	nonword	word	filler
PID-‘ōH	hinakahi	wāwahi	nonword	word	filler
PID-‘ōH	rakus	hāleu	nonword	word	filler
PID-‘ōH	manoz	haumāna	nonword	word	filler
PID-‘ōH	irratoa	huluhulu	nonword	word	filler
PID-‘ōH	blaloha	ikaika	nonword	word	filler
PID-‘ōH	tingahkali	kakani	nonword	word	filler
PID-‘ōH	kenat	kālai	nonword	word	filler
PID-‘ōH	‘aohalaniz	ke‘ena	nonword	word	filler
PID-‘ōH	rapax	pehu	nonword	word	filler
PID-‘ōH	pinat	ka‘ā	nonword	word	filler
PID-‘ōH	majax	lomi	nonword	word	filler
PID-‘ōH	okola	‘ā‘au	nonword	nonword	filler
PID-‘ōH	paco	kānoi	nonword	nonword	filler
PID-‘ōH	majai	maula	nonword	nonword	filler
PID-‘ōH	mokat	mo‘u	nonword	nonword	filler
PID-‘ōH	lelax	numo	nonword	nonword	filler
PID-‘ōH	niebe	poihi	nonword	nonword	filler
PID-‘ōH	tanax	a‘oua	nonword	nonword	filler
PID-‘ōH	rajak	halaili	nonword	nonword	filler
PID-‘ōH	bapah	hauaula	nonword	nonword	filler
PID-‘ōH	heffahead	hīmena	nonword	nonword	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	latonk	hukoa	nonword	nonword	filler
PID-‘ōH	yonore	‘ekane	nonword	nonword	filler
PID-‘ōH	kachi	‘okoke	nonword	nonword	filler
PID-‘ōH	nehualoa	ileia	nonword	nonword	filler
PID-‘ōH	retang	kalawi	nonword	nonword	filler
PID-‘ōH	beno	kehea	nonword	nonword	filler
PID-‘ōH	gote	kilawai	nonword	nonword	filler
PID-‘ōH	humotrillz	kuhiai	nonword	nonword	filler
PID-‘ōH	maoki	loawa	nonword	nonword	filler
PID-‘ōH	lenau	malauao	nonword	nonword	filler
PID-‘ōH	scrai	moake	nonword	nonword	filler
PID-‘ōH	okebe	muliwao	nonword	nonword	filler
PID-‘ōH	haolakea	ohoko	nonword	nonword	filler
PID-‘ōH	ohakaz	pilawao	nonword	nonword	filler
PID-‘ōH	hanataran	unena	nonword	nonword	filler
PID-‘ōH	mevai	auahina	nonword	nonword	filler
PID-‘ōH	baike	halauea	nonword	nonword	filler
PID-‘ōH	gobotani	hoāli‘i	nonword	nonword	filler
PID-‘ōH	lotulofa	holu‘olo	nonword	nonword	filler
PID-‘ōH	banai	‘olowale	nonword	nonword	filler
PID-‘ōH	mokau	ilauloa	nonword	nonword	filler
PID-‘ōH	hapak	ewi	nonword	nonword	filler
PID-‘ōH	kokat	pāmo	nonword	nonword	filler
PID-‘ōH	tanakuso	oi‘u	nonword	nonword	filler
PID-‘ōH	latona	neli	nonword	nonword	filler
PID-‘ōH	pakagi	homi	nonword	nonword	filler
PID-‘ōH	pakamaka	ao‘u	nonword	nonword	filler
PID-‘ōH	mokiwiki	wiai	nonword	nonword	filler
PID-‘ōH	lauku	wemi	nonword	nonword	filler
PID-‘ōH	blammajangah	hiho	nonword	nonword	filler
PID-‘ōH	potagee	‘āpau	word	nonword	filler
PID-‘ōH	killahweeffah	‘ihu	word	nonword	filler
PID-‘ōH	shaka	mauou	word	nonword	filler
PID-‘ōH	saimin	mō‘ao	word	nonword	filler
PID-‘ōH	arare	pamae	word	nonword	filler
PID-‘ōH	locomoco	pemu	word	nonword	filler
PID-‘ōH	chair	anaila	word	nonword	filler

Table C.2 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	ume	eleli	word	nonword	filler
PID-‘ōH	sleepahs	hawai‘o	word	nonword	filler
PID-‘ōH	aiyah	hikili	word	nonword	filler
PID-‘ōH	buckaloose	‘a‘ane	word	nonword	filler
PID-‘ōH	figga	‘ekaeo	word	nonword	filler
PID-‘ōH	jankenpo	‘ōlepa	word	nonword	filler
PID-‘ōH	manapua	ilawā	word	nonword	filler
PID-‘ōH	mokemobile	kānua	word	nonword	filler
PID-‘ōH	nails	kaweni	word	nonword	filler
PID-‘ōH	pukashirt	kohalo	word	nonword	filler
PID-‘ōH	pukahead	kuhalo	word	nonword	filler
PID-‘ōH	puating	lō‘ike	word	nonword	filler
PID-‘ōH	ainokea	mailoha	word	nonword	filler
PID-‘ōH	shootz	mohua	word	nonword	filler
PID-‘ōH	shoowah	molomau	word	nonword	filler
PID-‘ōH	rajah	oihaka	word	nonword	filler
PID-‘ōH	chout	pilai‘o	word	nonword	filler
PID-‘ōH	ukubillionz	wahike	word	nonword	filler
PID-‘ōH	cah	waihale	word	nonword	filler
PID-‘ōH	faddah	hale‘ele	word	nonword	filler
PID-‘ōH	termometah	hilinahi	word	nonword	filler
PID-‘ōH	brokdamout	huaoli‘e	word	nonword	filler
PID-‘ōH	hashi	‘oiapa‘a	word	nonword	filler
PID-‘ōH	mo‘bettah	kahaila	word	nonword	filler
PID-‘ōH	buggah	kape	word	nonword	filler
PID-‘ōH	girigiri	miwo	word	nonword	filler
PID-‘ōH	bondance	ha‘ī	word	nonword	filler
PID-‘ōH	geevum	wano	word	nonword	filler
PID-‘ōH	calabash	aoku	word	nonword	filler
PID-‘ōH	irraz	me‘u	word	nonword	filler
PID-‘ōH	manong	lapi	word	nonword	filler
PID-‘ōH	sole	nule	word	nonword	filler
PID-‘ōH	vog	māne	word	nonword	filler

Appendix C. Experimental Stimuli: Experiment 2

Table C.3: List 3 cross-language priming task stimuli

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	worn	palahē	word	word	critical
ENG-‘ōH	cockroach	‘elelū	word	word	critical
ENG-‘ōH	squid	mūhe‘e	word	word	critical
ENG-‘ōH	sloppy	kāpulu	word	word	critical
ENG-‘ōH	putrid	pilo	word	word	critical
ENG-‘ōH	ghost	lapu	word	word	critical
ENG-‘ōH	goodness	auwē	word	word	critical
ENG-‘ōH	disgusting	ho‘opailua	word	word	critical
ENG-‘ōH	liar	wahahe‘e	word	word	critical
ENG-‘ōH	hot	wela	word	word	critical
ENG-‘ōH	uptempo	mele	word	word	critical
ENG-‘ōH	surely	hiki	word	word	critical
ENG-‘ōH	across	palema‘i	word	word	control
ENG-‘ōH	lent	nui	word	word	control
ENG-‘ōH	nutmeg	poloka	word	word	control
ENG-‘ōH	damaging	‘ū‘ū	word	word	control
ENG-‘ōH	font	hūpō	word	word	control
ENG-‘ōH	walnut	ho‘oki‘eki‘e	word	word	control
ENG-‘ōH	repeat	waiū	word	word	control
ENG-‘ōH	squirrel	loea	word	word	control
ENG-‘ōH	slouch	hilahila	word	word	control
ENG-‘ōH	joint	pua‘a	word	word	control
ENG-‘ōH	heavy	‘eke	word	word	control
ENG-‘ōH	greedy	‘ai	word	word	control
ENG-‘ōH	sporty	anu	word	word	filler
ENG-‘ōH	fruit	ānuenue	word	word	filler
ENG-‘ōH	hide	‘āpana	word	word	filler
ENG-‘ōH	ugly	‘āpōpō	word	word	filler
ENG-‘ōH	wooden	aupuni	word	word	filler
ENG-‘ōH	precise	nu‘u	word	word	filler
ENG-‘ōH	jam	‘ōhai	word	word	filler
ENG-‘ōH	bush	‘eleu	word	word	filler
ENG-‘ōH	contest	‘ōiwi	word	word	filler
ENG-‘ōH	jacket	weli	word	word	filler
ENG-‘ōH	toenail	‘awa‘awa	word	word	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	daring	anapuni	word	word	filler
ENG-‘ōH	upset	‘īniha	word	word	filler
ENG-‘ōH	laps	paniolo	word	word	filler
ENG-‘ōH	napkin	hōkū	word	word	filler
ENG-‘ōH	blanket	peni	word	word	filler
ENG-‘ōH	vaking	imu	nonword	word	filler
ENG-‘ōH	fluko	puku	nonword	word	filler
ENG-‘ōH	berdict	iwakālua	nonword	word	filler
ENG-‘ōH	heport	kaha	nonword	word	filler
ENG-‘ōH	trahze	kāma‘a	nonword	word	filler
ENG-‘ōH	apess	kauka	nonword	word	filler
ENG-‘ōH	recooled	noho	nonword	word	filler
ENG-‘ōH	pyrics	kua‘āina	nonword	word	filler
ENG-‘ōH	prallpox	lā‘au	nonword	word	filler
ENG-‘ōH	chying	makemake	nonword	word	filler
ENG-‘ōH	suve	kupake‘e	nonword	word	filler
ENG-‘ōH	grists	kāpī	nonword	word	filler
ENG-‘ōH	toffin	hō‘ai	nonword	word	filler
ENG-‘ōH	larned	wai	nonword	word	filler
ENG-‘ōH	sargain	waha	nonword	word	filler
ENG-‘ōH	grampled	moa	nonword	word	filler
ENG-‘ōH	humbup	mai‘a	nonword	word	filler
ENG-‘ōH	drails	maka	nonword	word	filler
ENG-‘ōH	nall	pāpāleo	nonword	word	filler
ENG-‘ōH	gubplot	maka‘aniani	nonword	word	filler
ENG-‘ōH	vate	nane	nonword	word	filler
ENG-‘ōH	blantly	niho	nonword	word	filler
ENG-‘ōH	vodless	oli	nonword	word	filler
ENG-‘ōH	freelt	pololū	nonword	word	filler
ENG-‘ōH	trush	‘ōma‘oma‘o	nonword	word	filler
ENG-‘ōH	uldock	one	nonword	word	filler
ENG-‘ōH	bives	‘elepani	nonword	word	filler
ENG-‘ōH	stendthrift	‘ahua	nonword	word	filler
ENG-‘ōH	snart	pālule	nonword	word	filler
ENG-‘ōH	warmip	kūkini	nonword	word	filler
ENG-‘ōH	latty	‘ula	nonword	word	filler
ENG-‘ōH	brax	‘ulu	nonword	word	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	rainted	kūli‘u	nonword	word	filler
ENG-‘ōH	rigers	huā	nonword	word	filler
ENG-‘ōH	souch	moloā	nonword	word	filler
ENG-‘ōH	loophile	huihui	nonword	word	filler
ENG-‘ōH	larfy	pake	nonword	word	filler
ENG-‘ōH	clat	‘āko‘ako‘a	nonword	word	filler
ENG-‘ōH	splings	pa‘akiki	nonword	word	filler
ENG-‘ōH	torttheast	nalu	nonword	word	filler
ENG-‘ōH	zothy	hemā	nonword	nonword	filler
ENG-‘ōH	ipator	lalo	nonword	nonword	filler
ENG-‘ōH	clitten	lānou	nonword	nonword	filler
ENG-‘ōH	sermanent	noike	nonword	nonword	filler
ENG-‘ōH	pirebreaks	nokoi	nonword	nonword	filler
ENG-‘ōH	rampagins	pumo	nonword	nonword	filler
ENG-‘ōH	pherub	wailei	nonword	nonword	filler
ENG-‘ōH	soalhole	hamule	nonword	nonword	filler
ENG-‘ōH	sleeches	hauaho	nonword	nonword	filler
ENG-‘ōH	selevised	hoika‘i	nonword	nonword	filler
ENG-‘ōH	klereafter	hoikaka	nonword	nonword	filler
ENG-‘ōH	chimena	‘o‘ā‘o	nonword	nonword	filler
ENG-‘ōH	huspicious	‘oika‘i	nonword	nonword	filler
ENG-‘ōH	confussor	kahino	nonword	nonword	filler
ENG-‘ōH	interfice	kahono	nonword	nonword	filler
ENG-‘ōH	dearrange	ke‘aule	nonword	nonword	filler
ENG-‘ōH	pilitant	kekaeo	nonword	nonword	filler
ENG-‘ōH	milak	lālia	nonword	nonword	filler
ENG-‘ōH	thawl	lawaihae	nonword	nonword	filler
ENG-‘ōH	tinds	manani	nonword	nonword	filler
ENG-‘ōH	crunt	manaula	nonword	nonword	filler
ENG-‘ōH	borm	nolai‘a	nonword	nonword	filler
ENG-‘ōH	dorst	nualo	nonword	nonword	filler
ENG-‘ōH	prall	tahaku	nonword	nonword	filler
ENG-‘ōH	gans	uhoho	nonword	nonword	filler
ENG-‘ōH	geeds	ha‘aliho	nonword	nonword	filler
ENG-‘ōH	slore	ha‘alima	nonword	nonword	filler
ENG-‘ōH	mufts	ho‘ilona	nonword	nonword	filler
ENG-‘ōH	rogged	ho‘olona	nonword	nonword	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	hakes	ilahina	nonword	nonword	filler
ENG-‘ōH	plapps	ilaka‘i	nonword	nonword	filler
ENG-‘ōH	bips	ulo	nonword	nonword	filler
ENG-‘ōH	mog	mela	nonword	nonword	filler
ENG-‘ōH	klow	‘iko	nonword	nonword	filler
ENG-‘ōH	yoles	miao	nonword	nonword	filler
ENG-‘ōH	jelt	hanō	nonword	nonword	filler
ENG-‘ōH	hivs	wile	nonword	nonword	filler
ENG-‘ōH	fraft	mape	nonword	nonword	filler
ENG-‘ōH	wope	lewo	nonword	nonword	filler
ENG-‘ōH	velts	pe‘i	nonword	nonword	filler
ENG-‘ōH	wannabee	hokou	word	nonword	filler
ENG-‘ōH	lift	la‘ū	word	nonword	filler
ENG-‘ōH	rash	lauhi	word	nonword	filler
ENG-‘ōH	shut	mō‘ū	word	nonword	filler
ENG-‘ōH	weirdo	nonai	word	nonword	filler
ENG-‘ōH	thickener	poine	word	nonword	filler
ENG-‘ōH	fuss	wepo	word	nonword	filler
ENG-‘ōH	beautiful	halaini	word	nonword	filler
ENG-‘ōH	actually	hauauao	word	nonword	filler
ENG-‘ōH	nuts	hō‘iku	word	nonword	filler
ENG-‘ōH	core	hoikikī	word	nonword	filler
ENG-‘ōH	forks	‘elihi	word	nonword	filler
ENG-‘ōH	woo	‘oikeia	word	nonword	filler
ENG-‘ōH	plea	iolai	word	nonword	filler
ENG-‘ōH	brook	kaikamoi	word	nonword	filler
ENG-‘ōH	finger	kehua	word	nonword	filler
ENG-‘ōH	jazz	kilaina	word	nonword	filler
ENG-‘ōH	corpse	kulaine	word	nonword	filler
ENG-‘ōH	farewell	liola	word	nonword	filler
ENG-‘ōH	ramp	maluni	word	nonword	filler
ENG-‘ōH	melt	mawai‘o	word	nonword	filler
ENG-‘ōH	diarrhea	nahiko	word	nonword	filler
ENG-‘ōH	nerdy	nuana	word	nonword	filler
ENG-‘ōH	itchy	pohana	word	nonword	filler
ENG-‘ōH	fault	uka‘a	word	nonword	filler
ENG-‘ōH	centipede	eneloko	word	nonword	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	colorful	halaine	word	nonword	filler
ENG-‘ōH	shook	hoapi‘o	word	nonword	filler
ENG-‘ōH	fun	holikoki	word	nonword	filler
ENG-‘ōH	poor	‘oluhele	word	nonword	filler
ENG-‘ōH	flex	ilamo‘o	word	nonword	filler
ENG-‘ōH	slang	u‘o	word	nonword	filler
ENG-‘ōH	forever	aonu	word	nonword	filler
ENG-‘ōH	mesh	kū‘u	word	nonword	filler
ENG-‘ōH	billed	wemo	word	nonword	filler
ENG-‘ōH	slips	lehā	word	nonword	filler
ENG-‘ōH	jewel	mōpa	word	nonword	filler
ENG-‘ōH	deed	kawo	word	nonword	filler
ENG-‘ōH	medicine	lānu	word	nonword	filler
ENG-‘ōH	wings	lipe	word	nonword	filler
PID-‘ōH	bumbai	‘auane‘i	word	word	critical
PID-‘ōH	cheehoo	hūtā	word	word	critical
PID-‘ōH	charreh	makepono	word	word	critical
PID-‘ōH	habuteru	nuha	word	word	critical
PID-‘ōH	shishi	mimi	word	word	critical
PID-‘ōH	hammajang	kapakahi	word	word	critical
PID-‘ōH	bachī	pohō	word	word	critical
PID-‘ōH	gangies	hoaloha	word	word	critical
PID-‘ōH	itai	‘eha	word	word	critical
PID-‘ōH	minahs	‘u‘uku	word	word	critical
PID-‘ōH	howzit	pehea	word	word	critical
PID-‘ōH	dakine	mea	word	word	critical
PID-‘ōH	haolefied	iki	word	word	control
PID-‘ōH	bagoong	weluwelu	word	word	control
PID-‘ōH	ala-alas	kīkepa	word	word	control
PID-‘ōH	futless	hūpē	word	word	control
PID-‘ōH	had-it	‘ōhule	word	word	control
PID-‘ōH	trow	hepa	word	word	control
PID-‘ōH	horrahz	pipipi	word	word	control
PID-‘ōH	oddah	‘au‘au	word	word	control
PID-‘ōH	nuff	male	word	word	control
PID-‘ōH	hanakuso	hiamoe	word	word	control

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	kamaboko	lua	word	word	control
PID-‘ōH	moke	hō‘oio	word	word	control
PID-‘ōH	popo	lani	word	word	filler
PID-‘ōH	braddah	wini	word	word	filler
PID-‘ōH	musubi	uluna	word	word	filler
PID-‘ōH	usetato	no‘ono‘o	word	word	filler
PID-‘ōH	lihingmui	kūkaepele	word	word	filler
PID-‘ōH	humbug	‘īlio	word	word	filler
PID-‘ōH	kden	kō‘ai	word	word	filler
PID-‘ōH	kotonk	kane	word	word	filler
PID-‘ōH	buddahead	honu	word	word	filler
PID-‘ōH	yobo	nohea	word	word	filler
PID-‘ōH	malasada	ho‘oilo	word	word	filler
PID-‘ōH	li‘dat	kūpuna	word	word	filler
PID-‘ōH	bento	‘ī‘ike	word	word	filler
PID-‘ōH	obakechocho	lelele	word	word	filler
PID-‘ōH	no‘mo	hālau	word	word	filler
PID-‘ōH	planny	‘ahi	word	word	filler
PID-‘ōH	bito	kāhoaka	nonword	word	filler
PID-‘ōH	momong	pupule	nonword	word	filler
PID-‘ōH	banak	lapuwale	nonword	word	filler
PID-‘ōH	mauki	ka‘a	nonword	word	filler
PID-‘ōH	dinngah	wiwo‘ole	nonword	word	filler
PID-‘ōH	uk‘ume	iwi	nonword	word	filler
PID-‘ōH	rehut	laha	nonword	word	filler
PID-‘ōH	behua	‘ele‘ele	nonword	word	filler
PID-‘ōH	goke	‘ukulele	nonword	word	filler
PID-‘ōH	tanah	makika	nonword	word	filler
PID-‘ōH	bossabossa	liliko‘i	nonword	word	filler
PID-‘ōH	darehloco	alia	nonword	word	filler
PID-‘ōH	kaoyu	paka	nonword	word	filler
PID-‘ōH	yote	ho‘olana	nonword	word	filler
PID-‘ōH	latau	momona	nonword	word	filler
PID-‘ōH	hakakattack	niu	nonword	word	filler
PID-‘ōH	ainez	‘a‘ā	nonword	word	filler
PID-‘ōH	tapax	hau‘oli	nonword	word	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	raoks	kīna‘u	nonword	word	filler
PID-‘ōH	haoka	‘ōpū	nonword	word	filler
PID-‘ōH	rakeakai	apo	nonword	word	filler
PID-‘ōH	tana-hina-loa	piapia	nonword	word	filler
PID-‘ōH	bakut	naonao	nonword	word	filler
PID-‘ōH	makayoinz	kūpa‘a	nonword	word	filler
PID-‘ōH	zone	pika	nonword	word	filler
PID-‘ōH	racha	‘ili‘ili	nonword	word	filler
PID-‘ōH	‘aukenui	manini	nonword	word	filler
PID-‘ōH	‘ajahz	mano	nonword	word	filler
PID-‘ōH	hahap	pāpale	nonword	word	filler
PID-‘ōH	hinakahi	wāwahi	nonword	word	filler
PID-‘ōH	rakus	hāleu	nonword	word	filler
PID-‘ōH	manoz	haumāna	nonword	word	filler
PID-‘ōH	irratoa	huluhulu	nonword	word	filler
PID-‘ōH	blaloha	ikaika	nonword	word	filler
PID-‘ōH	tingahkali	kakani	nonword	word	filler
PID-‘ōH	kenat	kālai	nonword	word	filler
PID-‘ōH	‘aohalaniz	ke‘ena	nonword	word	filler
PID-‘ōH	rapax	pehu	nonword	word	filler
PID-‘ōH	pinat	ka‘ā	nonword	word	filler
PID-‘ōH	majax	lomi	nonword	word	filler
PID-‘ōH	okola	‘ā‘au	nonword	nonword	filler
PID-‘ōH	paco	kānoi	nonword	nonword	filler
PID-‘ōH	majai	maula	nonword	nonword	filler
PID-‘ōH	mokat	mo‘u	nonword	nonword	filler
PID-‘ōH	lelax	numo	nonword	nonword	filler
PID-‘ōH	niebe	poihi	nonword	nonword	filler
PID-‘ōH	tanax	a‘oua	nonword	nonword	filler
PID-‘ōH	rajak	halaili	nonword	nonword	filler
PID-‘ōH	bapah	hauaula	nonword	nonword	filler
PID-‘ōH	heffahead	hīmena	nonword	nonword	filler
PID-‘ōH	latonk	hukoa	nonword	nonword	filler
PID-‘ōH	yonore	‘ekane	nonword	nonword	filler
PID-‘ōH	kachi	‘okoke	nonword	nonword	filler
PID-‘ōH	nehualoa	ileia	nonword	nonword	filler
PID-‘ōH	retang	kalawi	nonword	nonword	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	beno	kehea	nonword	nonword	filler
PID-‘ōH	gote	kilawai	nonword	nonword	filler
PID-‘ōH	humotrillz	kuhiai	nonword	nonword	filler
PID-‘ōH	maoki	loawa	nonword	nonword	filler
PID-‘ōH	lenau	malauao	nonword	nonword	filler
PID-‘ōH	scrai	moake	nonword	nonword	filler
PID-‘ōH	okebe	muliwao	nonword	nonword	filler
PID-‘ōH	haolakea	ohoko	nonword	nonword	filler
PID-‘ōH	ohakaz	pilawao	nonword	nonword	filler
PID-‘ōH	hanataran	unena	nonword	nonword	filler
PID-‘ōH	mevai	auahina	nonword	nonword	filler
PID-‘ōH	baike	halauea	nonword	nonword	filler
PID-‘ōH	gobotani	hoāli‘i	nonword	nonword	filler
PID-‘ōH	lotulofa	holu‘olo	nonword	nonword	filler
PID-‘ōH	banai	‘olowale	nonword	nonword	filler
PID-‘ōH	mokau	ilauloa	nonword	nonword	filler
PID-‘ōH	hapak	ewi	nonword	nonword	filler
PID-‘ōH	kokat	pāmo	nonword	nonword	filler
PID-‘ōH	tanakuso	oi‘u	nonword	nonword	filler
PID-‘ōH	latona	neli	nonword	nonword	filler
PID-‘ōH	pakagi	homi	nonword	nonword	filler
PID-‘ōH	pakamaka	ao‘u	nonword	nonword	filler
PID-‘ōH	mokiwiki	wiai	nonword	nonword	filler
PID-‘ōH	lauku	wemi	nonword	nonword	filler
PID-‘ōH	blammajangah	hiho	nonword	nonword	filler
PID-‘ōH	potagee	‘āpau	word	nonword	filler
PID-‘ōH	killahweeffah	‘ihu	word	nonword	filler
PID-‘ōH	shaka	mauou	word	nonword	filler
PID-‘ōH	saimin	mō‘ao	word	nonword	filler
PID-‘ōH	arare	pamae	word	nonword	filler
PID-‘ōH	locomoco	pemu	word	nonword	filler
PID-‘ōH	chair	anaila	word	nonword	filler
PID-‘ōH	ume	eleli	word	nonword	filler
PID-‘ōH	sleepahs	hawai‘o	word	nonword	filler
PID-‘ōH	aiyah	hikili	word	nonword	filler
PID-‘ōH	buckaloose	‘a‘ane	word	nonword	filler
PID-‘ōH	figga	‘ekaeo	word	nonword	filler

Table C.3 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	jankenpo	‘ōlepa	word	nonword	filler
PID-‘ōH	manapua	ilawā	word	nonword	filler
PID-‘ōH	mokemobile	kānua	word	nonword	filler
PID-‘ōH	nails	kaweni	word	nonword	filler
PID-‘ōH	pukashirt	kohalo	word	nonword	filler
PID-‘ōH	pukahead	kuhalo	word	nonword	filler
PID-‘ōH	puating	lō‘ike	word	nonword	filler
PID-‘ōH	ainokea	mailoha	word	nonword	filler
PID-‘ōH	shootz	mohua	word	nonword	filler
PID-‘ōH	shoowah	molomau	word	nonword	filler
PID-‘ōH	rajah	oihaka	word	nonword	filler
PID-‘ōH	chout	pilai‘o	word	nonword	filler
PID-‘ōH	ukubillionz	wahike	word	nonword	filler
PID-‘ōH	cah	waihale	word	nonword	filler
PID-‘ōH	faddah	hale‘ele	word	nonword	filler
PID-‘ōH	termometah	hilinahi	word	nonword	filler
PID-‘ōH	brokdamout	huaoli‘e	word	nonword	filler
PID-‘ōH	hashi	‘oiapa‘a	word	nonword	filler
PID-‘ōH	mo‘bettah	kahaila	word	nonword	filler
PID-‘ōH	buggah	kape	word	nonword	filler
PID-‘ōH	girigiri	miwo	word	nonword	filler
PID-‘ōH	bondance	ha‘ī	word	nonword	filler
PID-‘ōH	geevum	wano	word	nonword	filler
PID-‘ōH	calabash	aoku	word	nonword	filler
PID-‘ōH	irraz	me‘u	word	nonword	filler
PID-‘ōH	manong	lapi	word	nonword	filler
PID-‘ōH	sole	nule	word	nonword	filler
PID-‘ōH	vog	māne	word	nonword	filler

Table C.4: List 4 cross-language priming task stimuli

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	worn	palahē	word	word	critical
ENG-‘ōH	cockroach	‘elelū	word	word	critical
ENG-‘ōH	squid	mūhe‘e	word	word	critical
ENG-‘ōH	sloppy	kāpulu	word	word	critical
ENG-‘ōH	putrid	pilo	word	word	critical

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	ghost	lapu	word	word	critical
ENG-‘ōH	goodness	auwē	word	word	critical
ENG-‘ōH	disgusting	ho‘opailua	word	word	critical
ENG-‘ōH	liar	wahahe‘e	word	word	critical
ENG-‘ōH	hot	wela	word	word	critical
ENG-‘ōH	uptempo	mele	word	word	critical
ENG-‘ōH	surely	hiki	word	word	critical
ENG-‘ōH	across	palema‘i	word	word	control
ENG-‘ōH	lent	nui	word	word	control
ENG-‘ōH	nutmeg	poloka	word	word	control
ENG-‘ōH	damaging	‘ū‘ū	word	word	control
ENG-‘ōH	font	hūpō	word	word	control
ENG-‘ōH	walnut	ho‘oki‘eki‘e	word	word	control
ENG-‘ōH	repeat	waiū	word	word	control
ENG-‘ōH	squirrel	loea	word	word	control
ENG-‘ōH	slouch	hilahila	word	word	control
ENG-‘ōH	joint	pua‘a	word	word	control
ENG-‘ōH	heavy	‘eke	word	word	control
ENG-‘ōH	greedy	‘ai	word	word	control
ENG-‘ōH	sporty	anu	word	word	filler
ENG-‘ōH	fruit	ānuenue	word	word	filler
ENG-‘ōH	hide	‘āpana	word	word	filler
ENG-‘ōH	ugly	‘āpōpō	word	word	filler
ENG-‘ōH	wooden	aupuni	word	word	filler
ENG-‘ōH	precise	nu‘u	word	word	filler
ENG-‘ōH	jam	‘ōhai	word	word	filler
ENG-‘ōH	bush	‘eleu	word	word	filler
ENG-‘ōH	contest	‘ōiwi	word	word	filler
ENG-‘ōH	jacket	weli	word	word	filler
ENG-‘ōH	toenail	‘awa‘awa	word	word	filler
ENG-‘ōH	daring	anapuni	word	word	filler
ENG-‘ōH	upset	‘iniha	word	word	filler
ENG-‘ōH	laps	paniolo	word	word	filler
ENG-‘ōH	napkin	hōkū	word	word	filler
ENG-‘ōH	blanket	peni	word	word	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	vaking	imu	nonword	word	filler
ENG-‘ōH	fluko	puku	nonword	word	filler
ENG-‘ōH	berdict	iwakālua	nonword	word	filler
ENG-‘ōH	heport	kaha	nonword	word	filler
ENG-‘ōH	trahze	kāma‘a	nonword	word	filler
ENG-‘ōH	apess	kauka	nonword	word	filler
ENG-‘ōH	recooled	noho	nonword	word	filler
ENG-‘ōH	pyrics	kua‘aina	nonword	word	filler
ENG-‘ōH	prallpox	lā‘au	nonword	word	filler
ENG-‘ōH	chying	makemake	nonword	word	filler
ENG-‘ōH	suve	kupake‘e	nonword	word	filler
ENG-‘ōH	grists	kāpī	nonword	word	filler
ENG-‘ōH	toffin	hō‘ai	nonword	word	filler
ENG-‘ōH	larned	wai	nonword	word	filler
ENG-‘ōH	sargain	waha	nonword	word	filler
ENG-‘ōH	grampled	moa	nonword	word	filler
ENG-‘ōH	humbup	mai‘a	nonword	word	filler
ENG-‘ōH	drails	maka	nonword	word	filler
ENG-‘ōH	nall	pāpāleo	nonword	word	filler
ENG-‘ōH	gubplot	maka‘aniani	nonword	word	filler
ENG-‘ōH	vate	nane	nonword	word	filler
ENG-‘ōH	blantly	niho	nonword	word	filler
ENG-‘ōH	vodless	oli	nonword	word	filler
ENG-‘ōH	freelt	pololū	nonword	word	filler
ENG-‘ōH	trush	‘ōma‘oma‘o	nonword	word	filler
ENG-‘ōH	uldock	one	nonword	word	filler
ENG-‘ōH	bives	‘elepani	nonword	word	filler
ENG-‘ōH	stendthrift	‘ahua	nonword	word	filler
ENG-‘ōH	snart	pālule	nonword	word	filler
ENG-‘ōH	warmip	kūkini	nonword	word	filler
ENG-‘ōH	latty	‘ula	nonword	word	filler
ENG-‘ōH	brax	‘ulu	nonword	word	filler
ENG-‘ōH	rainted	kūli‘u	nonword	word	filler
ENG-‘ōH	rigers	huā	nonword	word	filler
ENG-‘ōH	souch	moloā	nonword	word	filler
ENG-‘ōH	loophile	huihui	nonword	word	filler
ENG-‘ōH	larfy	pake	nonword	word	filler
ENG-‘ōH	clat	‘āko‘ako‘a	nonword	word	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	splings	pa‘akikī	nonword	word	filler
ENG-‘ōH	torttheast	nalū	nonword	word	filler
ENG-‘ōH	zothy	hemā	nonword	nonword	filler
ENG-‘ōH	ipatorđ	laloi	nonword	nonword	filler
ENG-‘ōH	clitten	lānou	nonword	nonword	filler
ENG-‘ōH	sermanent	noike	nonword	nonword	filler
ENG-‘ōH	pirebreaks	nokoi	nonword	nonword	filler
ENG-‘ōH	rampagins	pumo	nonword	nonword	filler
ENG-‘ōH	pherub	wailei	nonword	nonword	filler
ENG-‘ōH	soalhole	hamule	nonword	nonword	filler
ENG-‘ōH	sleeches	hauaho	nonword	nonword	filler
ENG-‘ōH	selevised	hoika‘i	nonword	nonword	filler
ENG-‘ōH	klereafter	hoikaka	nonword	nonword	filler
ENG-‘ōH	chimena	‘o‘ā‘o	nonword	nonword	filler
ENG-‘ōH	huspicious	‘oika‘i	nonword	nonword	filler
ENG-‘ōH	confussor	kahino	nonword	nonword	filler
ENG-‘ōH	interfice	kahono	nonword	nonword	filler
ENG-‘ōH	dearrange	ke‘aule	nonword	nonword	filler
ENG-‘ōH	pilitant	kekaeo	nonword	nonword	filler
ENG-‘ōH	milak	lālia	nonword	nonword	filler
ENG-‘ōH	thawl	lawaihae	nonword	nonword	filler
ENG-‘ōH	tinds	manani	nonword	nonword	filler
ENG-‘ōH	crunt	manaula	nonword	nonword	filler
ENG-‘ōH	borm	nolai‘a	nonword	nonword	filler
ENG-‘ōH	dorst	nualo	nonword	nonword	filler
ENG-‘ōH	prall	tahaku	nonword	nonword	filler
ENG-‘ōH	gans	uhoho	nonword	nonword	filler
ENG-‘ōH	geeds	ha‘aliho	nonword	nonword	filler
ENG-‘ōH	slore	ha‘alima	nonword	nonword	filler
ENG-‘ōH	mufts	ho‘ilona	nonword	nonword	filler
ENG-‘ōH	rogged	ho‘olona	nonword	nonword	filler
ENG-‘ōH	hakes	ilahina	nonword	nonword	filler
ENG-‘ōH	plapps	ilaka‘i	nonword	nonword	filler
ENG-‘ōH	bips	ulo	nonword	nonword	filler
ENG-‘ōH	mog	melā	nonword	nonword	filler
ENG-‘ōH	klow	‘iko	nonword	nonword	filler
ENG-‘ōH	yoles	miao	nonword	nonword	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	jelt	hanō	nonword	nonword	filler
ENG-‘ōH	hivs	wile	nonword	nonword	filler
ENG-‘ōH	fraft	mape	nonword	nonword	filler
ENG-‘ōH	wope	lewo	nonword	nonword	filler
ENG-‘ōH	velts	pe‘i	nonword	nonword	filler
ENG-‘ōH	wannabee	hokou	word	nonword	filler
ENG-‘ōH	lift	la‘ū	word	nonword	filler
ENG-‘ōH	rash	lauhi	word	nonword	filler
ENG-‘ōH	shut	mō‘ū	word	nonword	filler
ENG-‘ōH	weirdo	nonai	word	nonword	filler
ENG-‘ōH	thickener	poine	word	nonword	filler
ENG-‘ōH	fuss	wepo	word	nonword	filler
ENG-‘ōH	beautiful	halaini	word	nonword	filler
ENG-‘ōH	actually	hauauao	word	nonword	filler
ENG-‘ōH	nuts	hō‘iku	word	nonword	filler
ENG-‘ōH	core	hoikīkī	word	nonword	filler
ENG-‘ōH	forks	‘elihi	word	nonword	filler
ENG-‘ōH	woo	‘oikeia	word	nonword	filler
ENG-‘ōH	plea	iolai	word	nonword	filler
ENG-‘ōH	brook	kaikamoi	word	nonword	filler
ENG-‘ōH	finger	kehua	word	nonword	filler
ENG-‘ōH	jazz	kilaina	word	nonword	filler
ENG-‘ōH	corpse	kulaine	word	nonword	filler
ENG-‘ōH	farewell	liola	word	nonword	filler
ENG-‘ōH	ramp	maluni	word	nonword	filler
ENG-‘ōH	melt	mawai‘o	word	nonword	filler
ENG-‘ōH	diarrhea	nahiko	word	nonword	filler
ENG-‘ōH	nerdy	nuana	word	nonword	filler
ENG-‘ōH	itchy	pohana	word	nonword	filler
ENG-‘ōH	fault	uka‘a	word	nonword	filler
ENG-‘ōH	centipede	eneloko	word	nonword	filler
ENG-‘ōH	colorful	halaine	word	nonword	filler
ENG-‘ōH	shook	hoapi‘o	word	nonword	filler
ENG-‘ōH	fun	holikoki	word	nonword	filler
ENG-‘ōH	poor	‘oluhele	word	nonword	filler
ENG-‘ōH	flex	ilamo‘o	word	nonword	filler
ENG-‘ōH	slang	u‘o	word	nonword	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
ENG-‘ōH	forever	aonu	word	nonword	filler
ENG-‘ōH	mesh	kū‘u	word	nonword	filler
ENG-‘ōH	billed	wemo	word	nonword	filler
ENG-‘ōH	slips	lehā	word	nonword	filler
ENG-‘ōH	jewel	mōpa	word	nonword	filler
ENG-‘ōH	deed	kawo	word	nonword	filler
ENG-‘ōH	medicine	lānu	word	nonword	filler
ENG-‘ōH	wings	lipe	word	nonword	filler
PID-‘ōH	bumbai	‘auane‘i	word	word	critical
PID-‘ōH	cheehoo	hūtā	word	word	critical
PID-‘ōH	charreh	makepono	word	word	critical
PID-‘ōH	habuteru	nuha	word	word	critical
PID-‘ōH	shishi	mimi	word	word	critical
PID-‘ōH	hammajang	kapakahi	word	word	critical
PID-‘ōH	bachī	pohō	word	word	critical
PID-‘ōH	gangies	hoaloha	word	word	critical
PID-‘ōH	itai	‘eha	word	word	critical
PID-‘ōH	minahs	‘u‘uku	word	word	critical
PID-‘ōH	howzit	pehea	word	word	critical
PID-‘ōH	dakine	mea	word	word	critical
PID-‘ōH	haolefied	iki	word	word	control
PID-‘ōH	bagoong	weluwelu	word	word	control
PID-‘ōH	ala-alas	kīkepa	word	word	control
PID-‘ōH	futless	hūpē	word	word	control
PID-‘ōH	had-it	‘ōhule	word	word	control
PID-‘ōH	trow	hepa	word	word	control
PID-‘ōH	horrahz	pipipi	word	word	control
PID-‘ōH	oddah	‘au‘au	word	word	control
PID-‘ōH	nuff	male	word	word	control
PID-‘ōH	hanakuso	hiamoe	word	word	control
PID-‘ōH	kamaboko	lua	word	word	control
PID-‘ōH	moke	hō‘oio	word	word	control
PID-‘ōH	popo	lani	word	word	filler
PID-‘ōH	braddah	wini	word	word	filler
PID-‘ōH	musubi	uluna	word	word	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	usetato	no‘ono‘o	word	word	filler
PID-‘ōH	lihingmui	kūkaepele	word	word	filler
PID-‘ōH	humbug	‘īlio	word	word	filler
PID-‘ōH	kden	kō‘ai	word	word	filler
PID-‘ōH	kotonk	kane	word	word	filler
PID-‘ōH	buddahead	honu	word	word	filler
PID-‘ōH	yobo	nohea	word	word	filler
PID-‘ōH	malasada	ho‘oilō	word	word	filler
PID-‘ōH	li‘dat	kūpuna	word	word	filler
PID-‘ōH	bento	‘i‘ike	word	word	filler
PID-‘ōH	obakechocho	lelele	word	word	filler
PID-‘ōH	no‘mo	hālau	word	word	filler
PID-‘ōH	planny	‘ahi	word	word	filler
PID-‘ōH	bito	kāhoaka	nonword	word	filler
PID-‘ōH	momong	pupule	nonword	word	filler
PID-‘ōH	banak	lapuwale	nonword	word	filler
PID-‘ōH	mauki	ka‘a	nonword	word	filler
PID-‘ōH	dinngah	wiwo‘ole	nonword	word	filler
PID-‘ōH	uk‘ume	iwi	nonword	word	filler
PID-‘ōH	rehut	laha	nonword	word	filler
PID-‘ōH	behua	‘ele‘ele	nonword	word	filler
PID-‘ōH	goke	‘ukulele	nonword	word	filler
PID-‘ōH	tanah	makika	nonword	word	filler
PID-‘ōH	bossabossa	liliko‘i	nonword	word	filler
PID-‘ōH	darehloco	alia	nonword	word	filler
PID-‘ōH	kaoyu	paka	nonword	word	filler
PID-‘ōH	yote	ho‘olana	nonword	word	filler
PID-‘ōH	lelau	momona	nonword	word	filler
PID-‘ōH	hakakattack	niu	nonword	word	filler
PID-‘ōH	ainez	‘a‘ā	nonword	word	filler
PID-‘ōH	tapax	hau‘oli	nonword	word	filler
PID-‘ōH	raoks	kīna‘u	nonword	word	filler
PID-‘ōH	haoka	‘ōpū	nonword	word	filler
PID-‘ōH	rakeakai	apo	nonword	word	filler
PID-‘ōH	tana-hina-loa	piapia	nonword	word	filler
PID-‘ōH	bakut	naonao	nonword	word	filler
PID-‘ōH	makayoinz	kūpa‘a	nonword	word	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	zone	pika	nonword	word	filler
PID-‘ōH	racha	‘ili‘ili	nonword	word	filler
PID-‘ōH	‘aukenui	manini	nonword	word	filler
PID-‘ōH	‘ajahz	mano	nonword	word	filler
PID-‘ōH	hahap	pāpale	nonword	word	filler
PID-‘ōH	hinakahi	wāwahi	nonword	word	filler
PID-‘ōH	rakus	hāleu	nonword	word	filler
PID-‘ōH	manoz	haumāna	nonword	word	filler
PID-‘ōH	irratoa	huluhulu	nonword	word	filler
PID-‘ōH	blaloha	ikaika	nonword	word	filler
PID-‘ōH	tingahkali	kakani	nonword	word	filler
PID-‘ōH	kenat	kālai	nonword	word	filler
PID-‘ōH	‘aohalaniz	ke‘ena	nonword	word	filler
PID-‘ōH	rapax	pehu	nonword	word	filler
PID-‘ōH	pinat	ka‘ā	nonword	word	filler
PID-‘ōH	majax	lomi	nonword	word	filler
PID-‘ōH	okola	‘ā‘au	nonword	nonword	filler
PID-‘ōH	paco	kānoi	nonword	nonword	filler
PID-‘ōH	majai	maula	nonword	nonword	filler
PID-‘ōH	mokat	mo‘u	nonword	nonword	filler
PID-‘ōH	lelax	numo	nonword	nonword	filler
PID-‘ōH	niebe	poihi	nonword	nonword	filler
PID-‘ōH	tanax	a‘oua	nonword	nonword	filler
PID-‘ōH	rajak	halaili	nonword	nonword	filler
PID-‘ōH	bapah	hauaula	nonword	nonword	filler
PID-‘ōH	heffahead	hīmena	nonword	nonword	filler
PID-‘ōH	latonk	hukoa	nonword	nonword	filler
PID-‘ōH	yonore	‘ekane	nonword	nonword	filler
PID-‘ōH	kachi	‘okoke	nonword	nonword	filler
PID-‘ōH	nehualoa	ileia	nonword	nonword	filler
PID-‘ōH	retang	kalawi	nonword	nonword	filler
PID-‘ōH	beno	kehea	nonword	nonword	filler
PID-‘ōH	gote	kilawai	nonword	nonword	filler
PID-‘ōH	humotrillz	kuhiai	nonword	nonword	filler
PID-‘ōH	maoki	loawa	nonword	nonword	filler
PID-‘ōH	lenau	malauao	nonword	nonword	filler
PID-‘ōH	scrai	moake	nonword	nonword	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	okebe	muliwao	nonword	nonword	filler
PID-‘ōH	haolakea	ohoko	nonword	nonword	filler
PID-‘ōH	ohakaz	pilawao	nonword	nonword	filler
PID-‘ōH	hanataran	unena	nonword	nonword	filler
PID-‘ōH	mevai	auahina	nonword	nonword	filler
PID-‘ōH	baike	halauea	nonword	nonword	filler
PID-‘ōH	gobotani	hoāli‘i	nonword	nonword	filler
PID-‘ōH	lotulofa	holu‘olo	nonword	nonword	filler
PID-‘ōH	banai	‘olowale	nonword	nonword	filler
PID-‘ōH	mokau	ilauloa	nonword	nonword	filler
PID-‘ōH	hapak	ewi	nonword	nonword	filler
PID-‘ōH	kokat	pāmo	nonword	nonword	filler
PID-‘ōH	tanakuso	oi‘u	nonword	nonword	filler
PID-‘ōH	latona	neli	nonword	nonword	filler
PID-‘ōH	pakagi	homi	nonword	nonword	filler
PID-‘ōH	pakamaka	ao‘u	nonword	nonword	filler
PID-‘ōH	mokiwiki	wiai	nonword	nonword	filler
PID-‘ōH	lauku	wemi	nonword	nonword	filler
PID-‘ōH	blammajangah	hiho	nonword	nonword	filler
PID-‘ōH	potagee	‘āpau	word	nonword	filler
PID-‘ōH	killahweeffah	‘ihu	word	nonword	filler
PID-‘ōH	shaka	mauou	word	nonword	filler
PID-‘ōH	saimin	mō‘ao	word	nonword	filler
PID-‘ōH	arare	pamae	word	nonword	filler
PID-‘ōH	locomoco	pemu	word	nonword	filler
PID-‘ōH	chair	anaila	word	nonword	filler
PID-‘ōH	ume	eleli	word	nonword	filler
PID-‘ōH	sleepahs	hawai‘o	word	nonword	filler
PID-‘ōH	aiyah	hikili	word	nonword	filler
PID-‘ōH	buckaloose	‘a‘ane	word	nonword	filler
PID-‘ōH	figga	‘ekaeo	word	nonword	filler
PID-‘ōH	jankenpo	‘ōlepa	word	nonword	filler
PID-‘ōH	manapua	ilawā	word	nonword	filler
PID-‘ōH	mokemobile	kānua	word	nonword	filler
PID-‘ōH	nails	kaweni	word	nonword	filler
PID-‘ōH	pukashirt	kohalo	word	nonword	filler
PID-‘ōH	pukahead	kuhalo	word	nonword	filler

Table C.4 continued from previous page

Language Pairing	Prime	Target	Prime Category	Target Category	Trial Type
PID-‘ōH	puating	lō‘ike	word	nonword	filler
PID-‘ōH	ainokea	mailoha	word	nonword	filler
PID-‘ōH	shootz	mohua	word	nonword	filler
PID-‘ōH	shoowah	molomau	word	nonword	filler
PID-‘ōH	rajah	oihaka	word	nonword	filler
PID-‘ōH	chout	pilai‘o	word	nonword	filler
PID-‘ōH	ukubillionz	wahike	word	nonword	filler
PID-‘ōH	cah	waihale	word	nonword	filler
PID-‘ōH	faddah	hale‘ele	word	nonword	filler
PID-‘ōH	termometah	hilinahi	word	nonword	filler
PID-‘ōH	brokdamout	huaoli‘e	word	nonword	filler
PID-‘ōH	hashi	‘oiapa‘a	word	nonword	filler
PID-‘ōH	mo‘bettah	kahaila	word	nonword	filler
PID-‘ōH	buggah	kape	word	nonword	filler
PID-‘ōH	girigiri	miwo	word	nonword	filler
PID-‘ōH	bondance	ha‘ī	word	nonword	filler
PID-‘ōH	geevum	wano	word	nonword	filler
PID-‘ōH	calabash	aoku	word	nonword	filler
PID-‘ōH	irraz	me‘u	word	nonword	filler
PID-‘ōH	manong	lapi	word	nonword	filler
PID-‘ōH	sole	nule	word	nonword	filler
PID-‘ōH	vog	māne	word	nonword	filler

Appendix D

Experimental Stimuli: Experiment 3 Task 1

Table D.1: Participants saw only the nonword stimuli. Phonotactic probabilities for the nonwords and their matched real words are included for comparison; matched real words were selected to minimize phonotactic probability differences. Note that phonotactic probability values are reported on a log scale; negative values reflect the logarithmic transformation of probabilities less than 1.

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
kamea	-3.268741131	kahua	-3.670741081
maia	-3.328451633	kaua	-3.212371111
konui	-3.451542377	honua	-4.413266659
‘ani	-3.497308731	kona	-2.673646927
ame	-3.523716927	aku	-3.694455147
amea	-3.550576687	anei	-5.485807896
palana	-3.626607895	maluna	-5.239580631
kanui	-3.670220137	makua	-3.823079824
ma‘o	-3.671542406	nihi	-5.10601902
aku	-3.694455147	ala	-3.082314014
o‘e	-3.699463606	ala	-3.082314014
‘oku	-3.705592155	moku	-3.456752777
loana	-3.717121124	liona	-4.437847614
kaloa	-3.748192787	walea	-5.48173666
ike	-3.751540184	iho	-3.815015793
melo	-3.776590586	mele	-3.167592049
malana	-3.779800892	palena	-5.224745274

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
meana	-3.80603075	moana	-4.155353069
uale	-3.807802677	ualo	-4.230238438
‘ika	-3.819750071	lani	-3.484110832
mokua	-3.841442347	makua	-3.823079824
anui	-3.873233557	anei	-5.485807896
li’e	-3.88121748	hema	-3.120645523
konu	-3.881311417	kona	-2.673646927
aloa	-3.889064312	alia	-4.190722466
kaku	-3.894194603	kela	-4.151501656
konaka	-3.895303011	kanaka	-3.552060366
olu	-3.899982452	aku	-3.694455147
wika	-3.902826309	maka	-2.368447542
pa‘ana	-3.906187057	palena	-5.224745274
amai	-3.906431675	anei	-5.485807896
pui	-3.910889626	mai	-1.99143064
o‘i	-3.923931837	ola	-3.879148006
akou	-3.926837206	akua	-4.079144478
awai	-3.948923349	ahai	-4.039509773
pa’e	-3.967787743	hema	-3.120645523
muka	-3.983209133	hele	-2.472830296
a‘ia	-3.995309591	alia	-4.190722466
makaka	-3.999584675	manaka	-3.708699942
moloa	-4.000946999	melia	-4.489686012
lelu	-4.002071381	helu	-2.528495312
‘a‘i	-4.002404213	moku	-3.456752777
‘ake	-4.007361889	make	-3.155767918
molana	-4.031961441	palika	-4.586669445
melo	-4.034823895	melia	-4.489686012
moli	-4.052256584	maka	-2.368447542
‘oko	-4.06704998	moku	-3.456752777
lailu	-4.067248821	laila	-4.085686207
mela	-4.079708576	kela	-4.151501656
paloa	-4.080876827	walea	-5.48173666
me‘i	-4.096039295	kela	-4.151501656
uana	-4.110608578	ualo	-4.230238438
mo‘e	-4.134906769	manu	-4.256628513
lanaka	-4.138295174	manaka	-3.708699942

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
ua'o	-4.14872694	iune	-4.170080185
'olele	-4.155188084	helene	-4.285511017
meke	-4.156024933	hiki	-3.887227058
mailu	-4.157550335	maile	-4.420339584
nuia	-4.161841393	kaua	-3.212371111
'elu	-4.165024281	helu	-2.528495312
konua	-4.166888714	honua	-4.413266659
lui	-4.167178631	nui	-2.24919939
mema	-4.172280788	mele	-3.167592049
moki	-4.195721149	moku	-3.456752777
wike	-4.196361065	wale	-3.971321583
walana	-4.197721481	kalana	-3.293923378
maikana	-4.199129581	kaikana	-4.77451849
oloa	-4.200614452	ulia	-5.16528511
mike	-4.208728313	noho	-4.304669857
kamana	-4.213485718	kalana	-3.293923378
kalei	-4.214359283	walea	-5.48173666
peli	-4.21656847	hiki	-3.887227058
maloa	-4.234069824	holoi	-5.310965061
ma'ani	-4.248548985	makani	-4.073209763
ele	-4.253139973	ana	-2.533417702
meakana	-4.257837772	kaikana	-4.77451849
maki'i	-4.269412518	makiki	-4.432230473
noana	-4.277801037	luana	-5.347952843
kakua	-4.278883934	kahua	-3.670741081
loani	-4.284593105	kaona	-4.511769772
pākou	-3.561922312	kākou	-3.317111254
kānā	-3.110583067	nānā	-2.901767254
lāka	-3.624388933	pāka	-3.94278121
nāne	-3.707261562	kāne	-3.916077375
pā'oi	-3.739451408	pō'ai	-4.711065292
kūmola	-3.857042551	kūmole	-3.013247252
kūmolu	-3.877876997	kūmole	-3.013247252
pā'u	-3.961532116	pāka	-3.94278121
lānā	-4.023416042	nānā	-2.901767254
mō'ai	-4.047272682	pō'ai	-4.711065292
nāwa	-4.078777779	nāna	-3.131410122

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
‘ōlana	-4.091734409	‘ōlena	-4.62236166
māme	-4.117119312	kāne	-3.916077375
‘ōlelu	-4.195127487	‘ōlelo	-4.748461246
kāli	-4.270960808	kāne	-3.916077375
pō‘ī	-4.294777393	mō‘ī	-3.630984783
māne	-4.310482025	kōna	-4.799674034
lō‘ī	-4.31303072	mō‘ī	-3.630984783
‘ōleo	-4.35143137	kāloa	-5.0218153
lā‘ana	-4.391882896	lākana	-4.761678696
āka	-4.392560005	āna	-4.286595821
maku‘āina	-4.490555286	moku‘āina	-4.508917809
mō‘a	-4.508872509	pāka	-3.94278121
‘ōloa	-4.53903389	kōloa	-5.678159714
makikī	-4.551741123	pukikī	-5.049109936
pā‘o	-4.56499958	kāne	-3.916077375
kū‘ole	-4.567825317	kūmole	-3.013247252
‘ōna	-4.569571972	kōna	-4.799674034
ma‘ī	-4.570030212	ma‘ō	-5.053075314
kūkua	-4.577953815	kūlua	-5.350237846
ipū	-4.579567909	inā	-2.951788664
pūkā	-4.584023476	pēlā	-5.097847462
mālā	-4.584133148	māmā	-6.58194828
‘ākana	-4.596969604	‘āpana	-3.332066536
kāno	-4.601796627	kāne	-3.916077375
lāwai	-4.603275776	lā‘au	-4.909724712
lālai	-4.607047558	lālau	-5.636786461
kīka	-4.619978905	kōna	-4.799674034
pūka	-4.626369476	pāka	-3.94278121
‘āla	-4.629929543	pāka	-3.94278121
kū‘ana	-4.644133091	kū‘ono	-6.245885372
mā‘a	-4.675792694	pāka	-3.94278121
‘āpaka	-4.678221703	‘ālaka	-5.922054291
nākou	-4.694833755	mākou	-3.25188756
pōna	-4.753421783	kōna	-4.799674034
wēlā	-4.755131245	kēlā	-3.727903128
kūlama	-4.788482189	kūlana	-3.499276161
nā‘ō	-4.792551041	mō‘ī	-3.828407288

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
wā'o	-4.794668198	lā'o	-4.820008755
pa'i	-4.79514122	ka'ū	-5.570735931
pō'ia	-4.8015728	pā'ia	-4.369996071
pānā	-4.802838802	nānā	-2.901767254
'āpe	-4.807551384	kāne	-3.916077375
lō'ia	-4.819826126	līhau	-6.184019566
'āpū	-4.82542038	kālā	-4.654706478
lāne	-4.828910828	kāne	-3.916077375
'ōlema	-4.840841293	'ōlena	-4.62236166
nānui	-4.847671509	lānui	-5.969320297
kū'i	-4.850428104	lā'o	-4.820008755
lāku	-4.850823879	pāka	-3.94278121
kāni	-4.855392933	kāne	-3.916077375
kūpa	-4.855840683	kōna	-4.799674034
ka'akā	-4.860626698	kanakā	-4.341937065
kūke	-4.864521027	nāna	-3.131410122
ōla	-4.880220413	āna	-4.286595821
nānana	-4.880224228	'āpana	-3.332066536
kamakā	-4.884766102	kanakā	-4.341937065
kānai	-4.889944553	kākau	-5.127821922
pōlana	-4.890574455	pōlani	-5.458046913
nā'u	-4.896357536	nāna	-3.131410122
kūlame	-4.901875019	kūlani	-4.066748142
kēkiu	-4.916338444	wēkiu	-4.810349941
'ōmai	-4.917953014	'āhui	-4.626906872
mō'u	-4.920133591	wā'o	-4.794668198
pōpū	-4.940585136	pāpā	-6.938903809
'ōleli	-4.942463398	'ōlelo	-4.748461246
awā	-4.943463326	inā	-2.951788664
kūmai	-4.947850227	kāloa	-5.0218153
māloa	-4.95124197	mākua	-5.243870735
'āna	-4.951655388	hāna	-3.612905979
mānaka	-4.95628643	kānaka	-4.561882496
mōna	-4.960637093	nāna	-3.131410122
pāke	-4.976390839	wā'o	-4.794668198
'āpano	-4.990444183	'āpana	-3.332066536
wāme	-4.991808414	hāna	-3.612905979

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
mā'o	-4.993599892	māpu	-6.907806873
kīlani	-5.007264137	kūlani	-4.066748142
nāma	-5.014060497	nāna	-3.131410122
malikī	-5.01502943	pukikī	-5.049109936
kalakā	-5.023443699	kanakā	-4.341937065
kipū	-5.028586388	ka'ū	-5.570735931
'ōlaka	-5.031378269	'ālaka	-5.922054291
kūwa'o	-5.035361767	kūwaho	-4.071207523
ukā	-5.037948132	akā	-3.414389133
mō'ole	-5.049776554	mā'ili	-6.488866806
nāpila	-7.930813313	lālelo	-6.879790783
pākīpila	-7.93097353	pākīpika	-7.628818035
panekima	-7.930976391	papekema	-7.797260761
'ekukana	-7.931558132	wakikana	-7.310508251
polipine	-7.932451725	kalipuna	-6.508081436
panalou	-7.933216572	punahou	-6.455976009
kalepoko	-7.933567047	kaleponi	-7.287604809
'apuko	-7.933726311	'a'ohe	-6.370930672
anuape	-7.934185982	ahiahi	-7.190309048
peapea	-7.934697151	puapua	-6.79300499
kanolima	-7.934885025	kanalima	-6.609313965
punamā	-7.934966564	panamā	-7.85876894
kōpoa	-7.935008526	kāmoa	-6.795035362
kīmolo	-7.935010433	kū'ono	-6.245885372
mēki	-7.935324669	māpu	-6.907806873
līlili	-7.935396194	lāna'i	-6.465153217
'eieu	-7.935679436	heiau	-6.839509964
nalekana	-7.936900139	palekana	-6.663931847
malema'e	-7.937048435	ma'ema'e	-8.2374897
pololike	-7.938527584	pololika	-7.644993305
kalepeni	-7.938705921	kaleponi	-7.287604809
pukelala	-7.93922472	pokepola	-9.513386726
wekiko	-7.939230919	mekiko	-5.968534946
ponomoka	-7.939363956	pololika	-7.644993305
'ēlaka	-7.93947649	kālepa	-7.317207336
ne'apa	-7.939541817	'alepa	-6.550013542
pipoke	-7.940264225	popohe	-7.504107475

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
pu‘ipū	-7.940361023	po‘ipū	-6.43969059
ka‘i‘ino	-7.940548897	kalipuna	-6.508081436
memema	-7.940639019	‘elima	-6.357162476
aleane	-7.941295624	aheahe	-8.386638641
naiapa	-7.941696167	nuioka	-6.685477734
polekena	-7.942164898	pokekona	-7.917643547
mikiwala	-7.942203045	miki‘ala	-7.730043888
me‘alino	-7.943013191	pelelina	-7.760923862
kūpiku	-7.943178654	kāhiko	-7.487306595
nali‘a	-7.944824219	hali‘a	-6.678161621
nama‘ama	-7.94515419	lamalama	-7.872804165
puka‘a	-7.945397854	punahou	-6.455976009
hama	-7.945425987	‘umi	-6.59074688
polipola	-7.945621967	pokepola	-9.513386726
lanikila	-7.945687294	lanakila	-6.958368301
kenalima	-7.946442604	kanalima	-6.609313965
‘ialuni	-7.946522236	‘iolani	-8.500228882
e‘epo	-7.946623325	o‘ahu	-7.070436478
‘īnole	-7.94689846	pikake	-7.524001598
likekeo	-7.947291851	mokekao	-8.084990501
kapelaku	-7.947652817	kaka‘ako	-7.213842869
kilaima	-7.947672367	kalaima	-6.959157944
kalomako	-7.947751045	kololako	-9.461911201
‘akolaka	-7.949045658	‘awelika	-7.650775433
‘ekelike	-7.949525356	‘ewelaka	-8.673282623
pelapalu	-7.949552536	palapola	-7.943061829
kuleponi	-7.949931622	kaleponi	-7.287604809
kololaka	-7.94995594	kololako	-9.461911201
wālelo	-7.9502244	lālelo	-6.879790783
lāpako	-7.950743675	lāpaki	-7.362885475
kamalenī	-7.951173782	kamali‘i	-6.495099545
nukako	-7.951174259	nalala	-6.439413548
‘āne‘i	-7.951442719	‘ōpu‘u	-6.812529087
kaiapuki	-7.952628136	kaiapuni	-7.049773693
mōmalu	-7.953267574	mōhalu	-6.648550987
‘ilinui	-7.953606606	‘ilinoe	-8.441153526
lepalala	-7.953748226	hapawalu	-8.940691948

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
ponokī	-7.955700397	ponoī	-7.120977402
kekone	-7.955992699	kikinē	-7.308529854
niwala	-7.956850529	newaka	-8.129976273
walakila	-7.957438469	palakila	-7.386324883
nune'e	-7.957528114	'one'i	-7.265221119
kekekapa	-7.958508492	kekemapa	-8.239406586
nakiamā	-7.958736897	makeima	-7.139750004
lōkila	-7.95946312	lōkihi	-8.577402115
'ukeana	-7.959840775	kuleana	-7.046266556
Take	-7.959960461	lāpaki	-7.362885475
wikeō	-7.960010529	wikiō	-6.590454578
'anakoli	-7.961147785	'anakola	-7.588371754
pewalo	-7.962566853	nepala	-6.735029697
po'ola'u	-7.963046551	ho'opi'i	-7.856572628
palanowa	-7.963083744	palaholo	-7.798202038
'ono'a	-7.963407993	'oko'a	-7.127515793
lī'iu	-7.964395046	lā'ie	-6.252262115
'ulelena	-7.964961529	'ukulele	-8.130132675
pulipilo	-7.965137959	palapola	-7.943061829
ni'iamā	-7.965794086	hi'iaka	-7.040477753
palokila	-7.968442917	palakila	-7.386324883
kemopele	-7.968498707	hemolele	-5.999778271
wamo	-7.968556404	'umi	-6.59074688
wākoi	-7.968644619	kāpae	-6.518949032
neane	-7.969232559	kuini	-6.607920647
kakokaka	-7.970151424	kanekaka	-6.579940319
mokekio	-7.970282078	mokekao	-8.084990501
palalako	-7.970525265	pololika	-7.644993305
neme'e	-7.970915318	nenehe	-8.052018166
mupali	-7.97141552	nepala	-6.735029697
āwea	-7.97213316	ākea	-6.171949387
'ulelani	-7.972494125	'ilelani	-8.592634201
momoli'i	-7.97283411	kamali'i	-6.495099545
kīlone	-7.973020077	kāko'o	-7.073477745
ki'iū	-7.97373867	wikiō	-6.590454578
po'oleu	-7.974479198	ko'olau	-6.577877045
'iniona	-7.975576878	'iniana	-8.170416832

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
pōpo‘u	-7.97563076	pōpoki	-7.789032459
malike‘i	-7.976119995	moloka‘i	-7.274825096
kupepo	-7.976205349	kupelo	-6.629128933
mekipale	-7.976500988	kemikala	-7.068134785
numumea	-7.976544857	kumumea	-6.85439539
pomui	-7.977065563	pinao	-6.12219429
wimako	-7.977319241	kikako	-6.717649937
waikipū	-7.97827816	waikapū	-7.538689137
lepunika	-7.97831583	lipekika	-8.090015411
lia‘ae	-7.979575634	mea‘ai	-6.711965084
pelokine	-7.979593754	pelekane	-8.093774796
anāwa	-7.98064661	awāwa	-6.36200285
mi‘ili‘i	-7.982379436	li‘ili‘i	-6.643495083
moloki‘o	-7.983438015	moloka‘i	-7.274825096
ma‘iki‘iki	-7.98388958	ma‘ika‘ika	-8.136527061
maimolo	-7.983912945	maopopo	-9.799553871
nome‘i	-7.984494209	‘one‘i	-7.265221119
walanola	-7.984623909	palapola	-7.943061829
iemona	-7.984797955	iemene	-8.951075554
wipui	-7.985105038	kipuu	-7.127011299
kūhou	-7.985309124	wēlau	-7.324225426
waielu	-7.985528946	waiolu	-6.340688705
lālepo	-7.986968994	lālelo	-6.879790783
pū‘apu	-7.986987591	pū‘ulu	-6.559227943
nāolana	-7.986988068	nāihuna	-6.756262302
lolāna	-7.987010479	polēna	-7.150882244
ānoku	-7.987057686	ānehe	-7.370239258
kuluwa	-7.98745966	kakoka	-6.339014053
mēno	-7.988201141	māpu	-6.907806873
‘okakole	-7.988280296	‘anakola	-7.588371754
‘ipeku	-7.988317013	‘ewalu	-7.555225372
mū‘i	-7.988673687	māpu	-6.907806873
pianu	-7.988680363	hoaka	-5.990481377
nukinoa	-7.988893986	‘ilinoe	-8.441153526
lono‘ī	-7.989073753	pono‘ī	-7.120977402
waleluku	-7.989285946	waleluka	-6.965236664
pīka‘u	-7.989535809	hāla‘i	-6.653727055

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
pila'i	-7.98960495	pono'i	-7.120977402
kululai	-7.990474701	kahului	-6.127697468
pimakai	-7.990522385	pa'akai	-5.999601841
kuluilo	-7.990855694	kalaima	-6.959157944
kanamolo	-7.991237164	kanapela	-6.979746342
pekakui	-7.99148798	pa'akai	-5.999601841
'aie'ame	-7.991528511	kaiahome	-7.790527344
lamakila	-7.992318153	lanakila	-6.958368301
pēpule	-7.992427826	pōpule	-6.42163372
malamolo	-7.992591381	malama'o	-7.338502407
kakamoli	-7.992753983	ka'awale	-7.068377972
nakelaka	-7.992815018	'ukelena	-8.709891319
kemekala	-7.992854595	kemikala	-7.068134785
kanelaka	-7.99307394	waleluka	-6.965236664
po'omo'a	-7.993329048	ho'oku'i	-6.308495998
mo'oluli	-7.993829727	ho'oluli	-6.656894684
ma'wāme	-7.993917942	ma'wāma	-8.61067009
wake'i	-7.995135307	'one'i	-7.265221119
'ewelika	-7.995270252	'ewelaka	-8.673282623
ka'awa'e	-7.99531126	ka'awale	-7.068377972
'eulalu	-7.99532795	'eulopa	-6.54446888
kekau	-7.995977402	hekau	-6.051301956
poe'ea	-7.99606657	pai'ea	-8.715615273
pū'eli	-7.99620533	pū'ali	-6.651563644
lukale'e	-7.99652338	'ukulele	-8.130132675
pinamana	-7.997340202	pihemanu	-9.121997833
kunina	-7.997420311	kuhina	-6.5382967
eneana	-7.997578621	enaena	-7.440856934
kuinu	-7.998484135	kuini	-6.607920647
pāpoki	-7.998511791	pōpoki	-7.789032459
manekikā	-7.998545647	menekikā	-8.239958763
wekikana	-7.998556614	wakikana	-7.310508251
kenepaka	-7.999167442	kenemaka	-6.84204483
'alelewa	-7.999224186	'ilelani	-8.592634201
ni'aku	-7.999812126	ne'ena	-7.184824944
ka'ulawa	-7.999944687	ka'awale	-7.068377972
mepuno	-7.999960899	kupelo	-6.629128933

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
niloe	-8.000141144	lolii	-6.450511456
lepalene	-8.000190735	lenalena	-8.624716759
kekapiko	-8.000303268	kaka'ako	-7.213842869
manuo	-8.00083828	maloo	-6.613012791
maepae	-8.000968933	maemae	-6.105203152
nu'eluna	-8.000988007	ne'e'ana	-8.461318016
'alika	-8.001485825	'alāha	-7.653284073
kēkanu	-8.001606941	mōhalu	-6.648550987
maōe	-8.001867294	niōi	-6.964528084
nakiwila	-8.002048492	makawela	-7.876414776
kemikela	-8.002753258	kemikala	-7.068134785
pīnoe	-8.002780914	līhau	-6.184019566
ka'apou	-8.002904892	kahakai	-6.146406174
po'ao	-8.00299263	pehea	-6.680996895
palakola	-8.003129959	palakila	-7.386324883
lāwo'o	-8.00326252	'ōma'o	-6.464613438
pōni'i	-8.003328323	pōki'i	-6.837197304
'oko'u	-8.003661156	'oko'a	-7.127515793
kakinū	-8.005005836	kikinē	-7.308529854
'olalai	-8.005117416	'ilikai	-7.059216022
kolawale	-8.005282402	ka'awale	-7.068377972
kalukila	-8.005485535	palakila	-7.386324883
mikikipo	-8.005598068	mikikipi	-8.439543724
lūkene	-8.005933762	lūkini	-7.756094933
pa'akoli	-8.006046295	palaholo	-7.798202038
u'eli	-8.006345749	a'ela	-6.502534866
mūpu	-8.006733894	māpu	-6.907806873
kipikali	-8.006997108	kemikala	-7.068134785
'opekema	-8.007725716	'epekema	-9.590374947
mīnei	-8.007977486	līhau	-6.184019566
kaēnaeak	-15.05570602	naīnaua	-13.27703285
naāneeil	-15.03612709	naīnaua	-13.27703285
kīīneual	-15.03536797	naīnaua	-13.27703285
nuīnaua	-15.01224709	naīnaua	-13.27703285
naīnoua'	-14.98139668	naīnaua	-13.27703285
'aīnaua	-14.8433857	naīnaua	-13.27703285
'olekuonao	-14.83700371	'alekuoneo	-13.56740952

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
maīnaual	-14.79654121	naīnaual	-13.27703285
moīmeual	-14.72577763	naīnaual	-13.27703285
‘ilokuotea	-14.71629524	‘alekuoneo	-13.56740952
‘alekuonie	-14.71441269	‘alekuoneo	-13.56740952
minoupomu	-14.60880852	mineapoli	-13.06374836
nineeloli	-14.046731	mineapoli	-13.06374836
nāmiakēi	-14.03221226	nāmiakūa	-12.59550476
miteupoli	-13.9499073	mineapoli	-13.06374836
naōnaeal	-13.94691086	naīnaual	-13.27703285
napatune	-13.94392204	nepekune	-12.07018375
lāmiakāi	-13.80917931	nāmiakūa	-12.59550476
kuīnaial	-13.79942989	naīnaual	-13.27703285
tepekeni	-13.56864548	nepekune	-12.07018375
nāhekekā	-13.53894424	nāpekekā	-11.8182888
ilātela	-13.5095253	ilāmuku	-11.51108932
ehāmuli	-13.50686741	ilāmuku	-11.51108932
ulātiki	-13.42320251	ilāmuku	-11.51108932
ilātuke	-13.38434124	ilāmuku	-11.51108932
i‘āwuka	-13.35398102	ilāmuku	-11.51108932
nā‘eku‘ā	-13.29251099	nāpekekā	-11.8182888
luelho	-13.23863029	luelmo	-11.46405697
wuemmo	-13.14244366	luelmo	-11.46405697
lue‘hi	-13.12607479	luelmo	-11.46405697
uānamap	-13.09845734	uānahel	-11.2203064
iwūmuki	-13.09622097	ilāmuku	-11.51108932
uōlanel	-13.05619431	uānahel	-11.2203064
uēnalel	-13.05010033	uānahel	-11.2203064
uānakut	-13.04871655	uānahel	-11.2203064
uīnamul	-13.02931499	uānahel	-11.2203064
loemhi	-12.99316502	luelmo	-11.46405697
pineipoli	-12.97138977	mineapoli	-13.06374836
nāpekakē	-12.94064236	nāpekekā	-11.8182888
i‘āmuku	-12.92388916	ilāmuku	-11.51108932
lehekene	-12.91545486	nepekune	-12.07018375
ilānuku	-12.89814758	ilāmuku	-11.51108932
kōki‘ene‘i	-12.88641357	kīke‘eke‘e	-11.02880096
wuulmu	-12.85770893	luelmo	-11.46405697

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
ilāhoku	-12.83277035	ilāmuku	-11.51108932
itāmuku	-12.83147812	ilāmuku	-11.51108932
taikakeni	-12.81224251	naikokene	-10.84278011
luelwu	-12.79074669	luelmo	-11.46405697
uāwawel	-12.78368568	uānahel	-11.2203064
‘ūpekekā	-12.78362465	nāpekekā	-11.8182888
uānamul	-12.77120495	uānahel	-11.2203064
uāka‘ew	-12.76971245	uānahel	-11.2203064
luatho	-12.76482582	luelmo	-11.46405697
alīmimu	-12.74082661	ilāmuku	-11.51108932
ni‘emune	-12.72046947	nepekune	-12.07018375
ilāmuko	-12.71388721	ilāmuku	-11.51108932
luelwo	-12.65550613	luelmo	-11.46405697
luutko	-12.60319519	luelmo	-11.46405697
luulmo	-12.55175686	luelmo	-11.46405697
luelko	-12.54091835	luelmo	-11.46405697
nipekune	-12.52974033	nepekune	-12.07018375
paulto	-12.51755142	luelmo	-11.46405697
ta‘ekane	-12.51615429	nepekune	-12.07018375
ilāpuko	-12.51602364	ilāmuku	-11.51108932
pepekunu	-12.49917221	nepekune	-12.07018375
luukme	-12.49199486	luelmo	-11.46405697
naekonepi	-12.48391056	naikokene	-10.84278011
nepakune	-12.48255825	nepekune	-12.07018375
lāpekekē	-12.48243332	nāpekekā	-11.8182888
kuahto	-12.46526051	luelmo	-11.46405697
nāhekikā	-12.42647362	nāpekekā	-11.8182888
līwēnapa	-12.42383575	kēkēmapa	-10.44022846
popemune	-12.42270851	nepekune	-12.07018375
naulokine	-12.41222858	naikokene	-10.84278011
upāmuke	-12.35920238	ilāmuku	-11.51108932
ihāluki	-12.30985546	ilāmuku	-11.51108932
luulmi	-12.30640507	luelmo	-11.46405697
‘epekune	-12.29273701	nepekune	-12.07018375
nipekime	-12.27223682	nepekune	-12.07018375
elāmuki	-12.24293423	ilāmuku	-11.51108932
namlia‘e	-12.23185539	ha‘liaha	-11.0277195

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
uāpakel	-12.22179031	uānahel	-11.2203064
hulimaki	-12.20765305	puhipaka	-10.27510071
pimuēki	-12.19470978	pamiāli	-10.2598877
nāpekenā	-12.1864357	nāpekekā	-11.8182888
‘uopmi	-12.18008137	luelmo	-11.46405697
pitoāla	-12.17588329	pamiāli	-10.2598877
uāna‘el	-12.17303467	uānahel	-11.2203064
no‘uanu	-12.16575241	nu‘uanu	-10.19327259
nelekune	-12.16307735	nepekune	-12.07018375
a‘apaha	-12.15442848	awapuhi	-10.18260002
nunuanu	-12.15264511	nu‘uanu	-10.19327259
awapupi	-12.1463604	awapuhi	-10.18260002
leulto	-12.1261549	luelmo	-11.46405697
leelmo	-12.08379745	luelmo	-11.46405697
laanmo	-12.07462502	luelmo	-11.46405697
ne‘uamu	-12.07390785	nu‘uanu	-10.19327259
wauhmo	-12.06923962	luelmo	-11.46405697
kēkētapā	-12.06480026	kēkēmapa	-10.44022846
lau‘mo	-12.05114746	luelmo	-11.46405697
ilānoki	-12.03838444	ilāmuku	-11.51108932
uānalul	-12.03800297	uānahel	-11.2203064
‘emomapa	-12.03444004	wehimala	-10.07026577
woha‘ala	-12.02367783	wehimala	-10.07026577
kēkē‘upi	-12.01727295	kēkēmapa	-10.44022846
naīmuā	-12.01530361	wuīmuā	-10.08550072
naenokene	-12.00840759	naikokene	-10.84278011
iwāmuku	-12.00108337	ilāmuku	-11.51108932
pamiilo	-11.99679089	pamiāli	-10.2598877
wu‘uako	-11.9855442	nu‘uanu	-10.19327259
luulno	-11.95953751	luelmo	-11.46405697
nuepmo	-11.93987942	luelmo	-11.46405697
uip‘iā	-11.93868542	uil‘iā	-9.966914177
wimoeli	-11.91823387	mikouli	-10.00983238
uin‘iā	-11.90638447	uil‘iā	-9.966914177
nākulekō	-11.90556431	nāpekekā	-11.8182888
hāikeike	-11.90151119	hōikeike	-9.92951107
powuuli	-11.90008831	mikouli	-10.00983238

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
ne'uepu	-11.89989567	nu'uanu	-10.19327259
pumiōli	-11.89677429	pamiāli	-10.2598877
luenmo	-11.89081955	luelmo	-11.46405697
ualniū	-11.88717556	uil'ia	-9.966914177
kēikaite	-11.88714886	hōikeike	-9.92951107
ne'uanu	-11.88398838	nu'uanu	-10.19327259
kēlēmpa	-11.88322735	kēkēmapa	-10.44022846
wenumala	-11.8803587	wehimala	-10.07026577
alātoke	-11.87490749	ilāmuku	-11.51108932
'enoeli	-11.86529541	mikouli	-10.00983238
itīkulu	-11.85812759	ilāmuku	-11.51108932
'imalopi	-11.85108471	'okakopa	-10.02038383
ualhoō	-11.84479713	uil'ia	-9.966914177
'eku'opi	-11.84074211	'okakopa	-10.02038383
ewepe'i	-11.83965492	awapuhi	-10.18260002
wēimuā	-11.82232475	wuīmuā	-10.08550072
loekmo	-11.82143116	luelmo	-11.46405697
ketīmoka	-11.81728077	kanāmika	-9.829124451
kanūmike	-11.80491638	kanāmika	-9.829124451
alelohi	-11.79549599	awapuhi	-10.18260002
leewpe	-11.79413891	luelmo	-11.46405697
aip'io	-11.78130341	uil'ia	-9.966914177
alk'maku	-11.76165867	alk'makakā	-10.00196838
pu'iawi	-11.7612257	pamiāli	-10.2598877
pawiūne	-11.76006222	pamiāli	-10.2598877
pamoāli	-11.75821018	pamiāli	-10.2598877
luamta	-11.75516891	luelmo	-11.46405697
e'ēmeku	-11.7536993	ilāmuku	-11.51108932
iwapumi	-11.75298119	awapuhi	-10.18260002
nuīmuū	-11.74985218	wuīmuā	-10.08550072
'okakopi	-11.74809456	'okakopa	-10.02038383
elouni	-11.74650764	eloelo	-9.767097473
ouh'ia	-11.7434473	uil'ia	-9.966914177
iipniā	-11.73575306	uil'ia	-9.966914177
ka'liuha	-11.73236275	ha'liaha	-11.0277195
maewupe	-11.73218632	maopopo	-9.799553871
'a'uhuni	-11.71700096	hakuhune	-10.27003098

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
kilikihi	-11.68970203	kuhikuhi	-9.802131653
nu‘oinu	-11.68958855	nu‘uanu	-10.19327259
iin‘aū	-11.685709	uil‘iā	-9.966914177
naokikena	-11.68250847	naikokene	-10.84278011
lomulomi	-11.68148708	lomilomi	-9.721998215
ailwua	-11.67650795	uil‘iā	-9.966914177
lakahia	-11.67630482	lakawia	-9.686953545
kipaupi	-11.67475414	mikouli	-10.00983238
putowaka	-11.66975689	puhipaka	-10.27510071
lāpekekā	-11.66600418	nāpekekā	-11.8182888
wīmuīa	-11.66495609	wūmuīa	-10.08550072
patāmika	-11.660182	kanāmika	-9.829124451
le‘okilei	-11.65471745	‘elikilea	-9.848243713
naolipa	-11.6542902	maopopo	-9.799553871
auw‘iā	-11.65161991	uil‘iā	-9.966914177
iil‘uē	-11.64705086	uil‘iā	-9.966914177
haehma	-11.63939381	luelmo	-11.46405697
‘uowno	-11.62978745	luelmo	-11.46405697
oilkaō	-11.61867332	uil‘iā	-9.966914177
maopupa	-11.60318089	maopopo	-9.799553871
pamiēki	-11.59987068	pamiāli	-10.2598877
wahipila	-11.58336258	wehimala	-10.07026577
puhipala	-11.58236599	puhipaka	-10.27510071
wenima‘e	-11.57572556	wehimala	-10.07026577
‘ākēmipa	-11.574687	kēkēmapa	-10.44022846
‘epenima	-11.57307529	‘epekema	-9.590374947
nealno	-11.56734657	luelmo	-11.46405697
nue‘lo	-11.56641483	luelmo	-11.46405697
ammek	-11.55874252	amman	-9.568943024
lielpo	-11.54730892	luelmo	-11.46405697
luelpe	-11.54334068	luelmo	-11.46405697
wimiāli	-11.53818226	pamiāli	-10.2598877
mīkoawa	-11.53735447	mīkiawa	-9.573453903
uil‘aā	-11.53547573	uil‘iā	-9.966914177
iun‘aā	-11.5340538	uil‘iā	-9.966914177
uīnakel	-11.53337479	uānahel	-11.2203064
‘amiēpu	-11.53095436	pamiāli	-10.2598877

Continued on next page

Appendix D. Experimental Stimuli: Experiment 3 Task 1

stimuli (nonword)	stimuli phonotactic probability	realword (match)	real phonotactic probability
‘ūhawa	-11.52799225	‘ōhi‘a	-9.550857544
pomaā‘i	-11.52477646	pamiāli	-10.2598877
muoho‘o	-11.52427387	maopopo	-9.799553871
elāmoku	-11.52405739	ilāmuku	-11.51108932
mukeunma	-11.52198887	mikio‘ma	-9.611008644
‘āhi‘a	-11.51460361	‘ōhi‘a	-9.550857544
ehuili	-11.51032639	eloelo	-9.767097473
awman	-11.50693321	amman	-9.568943024

Appendix E

Experimental Stimuli: Experiment 3 Task 2

Table E.1: Full possible stimuli list for the Word Identification Task, including the complete pool of real and nonreal word stimuli and their calculated phonotactic probabilities. Phonotactic probability values are reported on a log scale, where negative values reflect the logarithmic transformation of probabilities less than 1. For clarity, the stimuli presented in the experiment were selected as subsets from this full stimulus pool, with each experimental list containing a different subset of items. Items were semi-randomly assigned to lists to ensure comparable phonotactic probability means and distributions across lists, as well as an equal number of real and nonreal stimuli and other experimental considerations. Participants were randomly assigned to one of the experimental lists. See Chapter 5, Methods for further information.

Stimuli	Type	Phonotactic Probability
keo	nonreal	-3.035352707
ini	nonreal	-2.719213486
aki	nonreal	-2.901767254
kukomi	nonreal	-8.595476151
‘āpa	nonreal	-2.533417702
nua	nonreal	-2.330404282
potu	nonreal	-8.668377876
muti	nonreal	-8.674211502
kee	nonreal	-3.916077375
wākou	nonreal	-4.04815197
ani‘i	nonreal	-4.049718857
kenaka	nonreal	-4.079144478

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
kuoa	nonreal	-4.119190693
mono	nonreal	-3.916077375
woku	nonreal	-5.902549744
ino	nonreal	-2.487043858
‘ākou	nonreal	-2.368447542
kole	nonreal	-2.719213486
iaa	nonreal	-5.01328373
kiwo	nonreal	-7.275802612
akala	nonreal	-3.694455147
nowo	nonreal	-6.666563988
ma‘e	nonreal	-3.131410122
wia	nonreal	-3.916077375
‘okule	nonreal	-4.016505241
puno	nonreal	-4.698660851
poike	nonreal	-5.000532627
leku	nonreal	-5.024629116
wo‘o	nonreal	-5.536873341
lemo	nonreal	-6.141366005
akea	nonreal	-3.567395926
koku	nonreal	-3.631931305
nepi	nonreal	-6.793904781
monawa	nonreal	-6.880151272
liai	nonreal	-5.314245701
kēika	nonreal	-3.669445515
mailo	nonreal	-3.955245972
loaka	nonreal	-4.333293438
kumo	nonreal	-4.333293438
‘ile	nonreal	-4.561882496
kema	nonreal	-2.533417702
luaka	nonreal	-5.01328373
imolu	nonreal	-7.649828911
walimo	nonreal	-8.281145096
aika	nonreal	-3.727903128
owa	nonreal	-4.383863926
kīlā	nonreal	-5.004992962
maku	nonreal	-2.719213486

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
‘oikou	nonreal	-6.475445271
wu‘e	nonreal	-6.633051395
ema	nonreal	-3.4473598
womi	nonreal	-6.828740597
o‘awe	nonreal	-6.828740597
oeponi	nonreal	-7.154663563
maki	nonreal	-2.533417702
wu‘a	nonreal	-7.253604412
anoe	nonreal	-5.239580631
kelu	nonreal	-2.533417702
kūla	nonreal	-2.487043858
uko	nonreal	-5.311607361
punu	nonreal	-5.239580631
ai‘o	nonreal	-2.719213486
aii	nonreal	-4.325684071
makakuko	nonreal	-8.926109314
mako	nonreal	-3.153085947
pa‘o	nonreal	-3.167592049
manake	nonreal	-3.738444328
eka	nonreal	-4.104712963
mo‘olele	nonreal	-5.41158247
kānama	nonreal	-5.500247478
‘o‘epahi	nonreal	-14.75470638
‘omu	nonreal	-4.16805172
ikī	nonreal	-3.131410122
lāeo	nonreal	-5.355561256
leme	nonreal	-5.696377754
keki	nonreal	-2.92983675
kalu	nonreal	-2.92983675
pōle	nonreal	-2.92983675
mewa	nonreal	-5.591528416
momo	nonreal	-6.075169086
‘ame	nonreal	-3.035352707
anaika	nonreal	-5.779873848
lā‘ia	nonreal	-3.916077375
alā	nonreal	-3.340226173

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
wowe	nonreal	-6.633051395
mape	nonreal	-4.079144478
malema	nonreal	-5.000532627
kūkou	nonreal	-4.317144394
kaeime	nonreal	-10.10027981
laeo	nonreal	-5.355561256
meno	nonreal	-5.41158247
‘akou	nonreal	-3.796062469
pome	nonreal	-5.000532627
pama	nonreal	-3.916077375
we‘u	nonreal	-6.785011292
womu	nonreal	-7.32241106
ilī	nonreal	-3.484110832
wō‘ike	nonreal	-7.977184772
‘owu	nonreal	-7.977184772
mīape	nonreal	-9.953654289
pepu	nonreal	-4.317144394
wi‘e	nonreal	-4.333293438
neka	nonreal	-3.706899881
manuno	nonreal	-9.047803879
maimi	nonreal	-6.365170002
lumu	nonreal	-5.004992962
mee	nonreal	-4.113284588
wipo	nonreal	-5.939777851
wuma	nonreal	-5.902549744
malamu	nonreal	-5.902549744
‘onano	nonreal	-6.828740597
wānā	nonreal	-3.955245972
‘uma	nonreal	-4.514782906
napou	nonreal	-6.880151272
maikani	nonreal	-3.984095097
pume	nonreal	-4.191795349
waku	nonreal	-4.423651695
eme	nonreal	-5.004992962
mui‘e	nonreal	-7.458953381
mipope	nonreal	-7.977184772

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
wama	nonreal	-4.588213921
mawe	nonreal	-4.618588924
nāno	nonreal	-3.694455147
mamoe	nonreal	-7.85333252
maumoko	nonreal	-9.677916527
wuhei	nonreal	-9.677916527
kuo	nonreal	-5.540477276
wonoa	nonreal	-6.695352077
napi	nonreal	-5.992637634
pēpō	nonreal	-5.939777851
wopala	nonreal	-7.649828911
tōpō	nonreal	-8.516010284
lipi'ai	nonreal	-8.49919796
anawai	nonreal	-4.317144394
wōlū	nonreal	-8.283745766
kiwula	nonreal	-8.49919796
naleneke	nonreal	-8.165423393
pālouleu	nonreal	-11.58869553
wonu	nonreal	-6.475445271
ikia	nonreal	-3.707152128
ma'ale	nonreal	-3.796062469
'eikela	nonreal	-9.390766144
'ūpeku	nonreal	-9.390766144
luimoi	nonreal	-9.351485252
hi'omoli	nonreal	-12.37171841
tahika	nonreal	-12.38354874
'āwoli	nonreal	-9.530257225
nāwala	nonreal	-4.159717083
naniu	nonreal	-7.796487808
naleana	nonreal	-7.796487808
hikepu	nonreal	-11.26631451
nūelo	nonreal	-7.533736229
'ālēpē	nonreal	-8.165423393
waeloia	nonreal	-8.600200653
kō'aamai	nonreal	-10.8302145
pakiawe	nonreal	-9.21835804

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
niute	nonreal	-9.21835804
moupi	nonreal	-9.212049484
imoka	nonreal	-7.649828911
kukile	nonreal	-7.773289204
iopika	nonreal	-7.85333252
a'ualeli	nonreal	-9.424370766
kikuwā	nonreal	-9.434296608
makueli'o	nonreal	-9.673197746
kīwuhi	nonreal	-11.3669281
mapalō	nonreal	-8.689035416
mākipi	nonreal	-7.924990654
wowalaho	nonreal	-14.44317341
māmeka	nonreal	-5.004992962
woloma	nonreal	-8.29842186
elenū	nonreal	-8.302148819
kupuka	nonreal	-3.694455147
kūtali	nonreal	-9.269548416
lūlili	nonreal	-9.342009544
lelipuwe	nonreal	-10.30993938
pikūpakā	nonreal	-10.77358246
kū'ane	nonreal	-5.412136555
pepū	nonreal	-5.41158247
lewea	nonreal	-5.500247478
nō'ie	nonreal	-7.924990654
mi'u	nonreal	-6.793904781
akelene	nonreal	-6.793904781
wile'ia	nonreal	-8.516010284
wene'i	nonreal	-8.574795723
lihiu	nonreal	-8.574795723
emame	nonreal	-8.604872704
hunapule	nonreal	-12.59384727
konawaeni	nonreal	-9.342009544
wi'oleko	nonreal	-9.417358398
punako	nonreal	-4.698660851
nuaka	nonreal	-4.333293438
opelika	nonreal	-6.140192509

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
kowuki	nonreal	-8.689035416
‘eweloma	nonreal	-8.709891319
lāeni	nonreal	-8.071462631
kōmo‘o	nonreal	-8.165423393
kumukilu	nonreal	-8.223135948
hekia‘e	nonreal	-11.47149181
kūwīnahi	nonreal	-14.75470638
pūmomu	nonreal	-9.283093452
kamamena	nonreal	-8.574795723
‘ōpo‘o	nonreal	-5.500247478
kawuhā	nonreal	-12.20135403
‘aku	nonreal	-3.669445515
i‘iki‘i	nonreal	-6.142563343
wo‘ao	nonreal	-10.71238708
muea	nonreal	-4.333293438
hiwa‘ita	nonreal	-16.95898056
lototoe	nonreal	-13.55908585
māpua	nonreal	-6.695352077
pā‘ema	nonreal	-6.749982834
kaikena	nonreal	-6.785011292
lū‘ia	nonreal	-6.954136372
kepūkinō	nonreal	-11.80903625
kihipa	nonreal	-10.97105503
kūpea	nonreal	-6.156310081
kālipo	nonreal	-6.141366005
nome‘o	nonreal	-7.564875126
pana‘ane	nonreal	-6.487393856
wāpū	nonreal	-6.515633583
mele‘unau	nonreal	-14.44317341
kekopo	nonreal	-7.458953381
mūpō	nonreal	-7.909927845
melenela	nonreal	-7.773289204
‘ōmene	nonreal	-7.773247242
malawelo	nonreal	-7.533736229
lēkū	nonreal	-7.502708912
kīmiu	nonreal	-7.85333252

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
kaoiu	nonreal	-7.832990646
tīhui	nonreal	-11.80177975
mu‘iloli	nonreal	-11.26631451
wai‘ile	nonreal	-5.355561256
mume	nonreal	-5.311607361
puwoō	nonreal	-9.706343651
kumokau	nonreal	-9.677916527
pānola	nonreal	-8.217745781
mikulama	nonreal	-8.235926628
nawo	nonreal	-8.217745781
wetuni	nonreal	-9.710737228
mo‘akū	nonreal	-8.049596786
pu‘aiete	nonreal	-12.38354874
pōpūae	nonreal	-7.773247242
kī‘aku	nonreal	-7.856757164
tīlipu	nonreal	-12.71218586
koukunoī	nonreal	-13.80884361
hōkīkē	nonreal	-13.80884361
kopaliwa	nonreal	-9.047803879
hautipo	nonreal	-13.55908585
panakoō	nonreal	-9.087730408
mukokene	nonreal	-10.45592499
‘olā‘ulā	nonreal	-9.416716576
‘ā‘alo‘o	nonreal	-9.02824688
hepakili	nonreal	-12.71218586
unuhie	nonreal	-14.1661253
kuemuli	nonreal	-10.71238708
hililie	nonreal	-12.21024609
piwopī	nonreal	-12.20135403
lelopimau	nonreal	-14.75470638
muwi	nonreal	-6.954136372
taeke	nonreal	-8.223135948
nalepapa	nonreal	-8.600200653
kūkaehio	nonreal	-11.14692307
‘āmāliu	nonreal	-8.926109314
iwe	nonreal	-5.500247478

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
meleloi	nonreal	-6.828740597
hīhō	nonreal	-11.08396435
lūwūnī	nonreal	-13.23092937
mīwīlu	nonreal	-9.710737228
koekuoa	nonreal	-9.673197746
‘ōwomo	nonreal	-9.673197746
pamotami	nonreal	-14.75470638
piwo	nonreal	-7.932122231
maneno	nonreal	-7.924990654
nua‘inu	nonreal	-8.154133797
nī‘ono	nonreal	-7.856757164
munai	nonreal	-7.832990646
hiuhue	nonreal	-13.51743603
wowahe	nonreal	-13.18674946
kilewa‘a	nonreal	-9.707769394
ehopuno	nonreal	-12.71218586
‘oewo‘a	nonreal	-11.80903625
wahewea	nonreal	-11.66623497
wunikula	nonreal	-11.08396435
‘awukeki	nonreal	-10.8302145
kaiwine	nonreal	-7.773289204
o‘okino	nonreal	-6.828740597
pē‘ai	nonreal	-7.364573002
wikiliwa	nonreal	-8.165423393
mōhae	nonreal	-8.071462631
tima	nonreal	-7.932122231
woa‘e	nonreal	-8.154133797
kiōmana	nonreal	-8.154133797
‘āku‘o	nonreal	-6.197011948
akapana	nonreal	-4.514782906
uwu‘uli	nonreal	-10.5489254
kenea	nonreal	-4.317144394
owiaho	nonreal	-14.1661253
mā‘iu	nonreal	-5.536873341
‘oiai‘i	nonreal	-6.502534866
‘animō	nonreal	-10.14315796

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
‘ā‘epo	nonreal	-10.34864235
‘ehokui	nonreal	-11.08396435
wōku‘o	nonreal	-10.61230087
pu‘ipiu	nonreal	-10.30993938
manile‘u	nonreal	-10.30993938
pūake	nonreal	-6.633051395
‘akuluka	nonreal	-6.472995281
palione	nonreal	-5.939777851
kanale	nonreal	-3.962133169
papamuke	nonreal	-8.595476151
‘ēnapa	nonreal	-8.375267029
‘imilo	nonreal	-8.283745766
‘olomapa	nonreal	-8.283745766
kānamo	nonreal	-8.281145096
pukipiki	nonreal	-6.502534866
mekelia	nonreal	-6.472995281
pōpulu	nonreal	-5.992637634
lōloa	nonreal	-6.142563343
makeale	nonreal	-6.142563343
kamalia	nonreal	-5.72254324
tikelepa	nonreal	-12.406394
‘eneki‘i	nonreal	-7.773247242
pālai	nonreal	-4.423651695
wiha	nonreal	-8.595476151
paekumu	nonreal	-8.375267029
noloū	nonreal	-8.373103142
mea	real	-3.971785784
ana	real	-3.357239962
aku	real	-3.548609734
kekahi	real	-8.876512527
‘ana	real	-3.164143324
nui	real	-2.964545965
hana	real	-8.94432354
ho‘i	real	-9.009021759
loa	real	-4.613474846
lākou	real	-4.824007511

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
ali'i	real	-4.864592552
kanaka	real	-4.889188766
kēia	real	-5.039346218
pono	real	-4.628186226
hiki	real	-6.332588196
iho	real	-3.123483658
lakou	real	-3.078820705
hele	real	-3.341528893
oia	real	-5.647399902
wahi	real	-7.585589409
akula	real	-4.372840405
noho	real	-7.171102524
po'e	real	-3.742677212
mua	real	-4.603640556
'a'ole	real	-4.792137623
wale	real	-5.525684357
keiki	real	-5.584459782
luna	real	-5.656776905
hale	real	-6.118391991
paha	real	-6.6889081
akua	real	-4.157613277
loko	real	-4.204740524
kahi	real	-7.297968388
manawa	real	-7.361770153
kaua	real	-5.932384968
'āina	real	-4.275841713
maila	real	-4.645836353
loa'a	real	-5.285215378
kumu	real	-5.262200356
'ike	real	-5.418927193
'ole	real	-3.170024872
laila	real	-5.652090549
ihola	real	-7.86816597
wahine	real	-8.599855423
aina	real	-4.493168354
a'e	real	-5.320204735

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
kēlā	real	-5.603246689
ku‘u	real	-3.438390255
‘oukou	real	-6.972984314
holo	real	-7.101998329
ina	real	-4.043174267
hewa	real	-7.328465462
a‘ela	real	-7.339359283
aupuni	real	-7.522803783
lawe	real	-3.292544842
ko‘u	real	-7.567990303
inoa	real	-5.873892784
kela	real	-3.295050144
kāna	real	-3.146795034
ano	real	-5.894615173
‘ano	real	-5.864458084
ia‘u	real	-3.390598536
eia	real	-5.246250153
makahiki	real	-9.190203667
moku	real	-3.880402803
pa‘a	real	-3.896652937
mana‘o	real	-4.496020794
ola	real	-4.951825619
mo‘olelo	real	-6.02405262
kānaka	real	-6.077474117
ho‘okahi	real	-15.20251751
pili	real	-5.097811222
aka	real	-3.733172655
lāua	real	-5.997459412
lima	real	-6.217032433
kana	real	-3.648143768
kula	real	-3.650321007
kāne	real	-3.596664667
helu	real	-6.164613247
komo	real	-6.507868767
maka	real	-3.669803143
ikaika	real	-6.287825584

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
mākou	real	-4.625005722
akā	real	-3.962210178
lohe	real	-7.10639143
kapa	real	-4.914799213
maluna	real	-5.596417904
kākou	real	-5.194841862
kauoha	real	-10.39713001
laua	real	-5.997459412
kino	real	-6.025914192
makua	real	-4.522366524
pela	real	-5.597912788
puka	real	-4.619526386
haku	real	-7.249419689
hala	real	-7.603684902
inā	real	-4.082770348
hō‘ike	real	-8.224648476
huli	real	-8.222938538
hā‘awi	real	-10.25919724
lani	real	-5.199657917
na‘e	real	-5.258528709
mele	real	-4.408713341
kahiko	real	-9.315976143
maoli	real	-6.896925926
lalo	real	-5.621860027
moe	real	-4.9910326
kapu	real	-6.339345455
nona	real	-6.327584743
malama	real	-6.325918674
‘ohana	real	-7.343548775
nānā	real	-4.638807297
ka‘u	real	-5.390624523
na‘au	real	-7.366231918
maika‘i	real	-4.766602039
pule	real	-5.147541046
pu‘u	real	-5.335034847
iki	real	-5.612421036

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
waiho	real	-7.705848217
mahope	real	-8.243095398
wa'a	real	-5.448371887
pane	real	-5.465934277
nāna	real	-4.392981052
mamua	real	-8.051687241
maopopo	real	-10.06303978
pehea	real	-10.04835796
hoa	real	-6.158089638
honua	real	-7.197655201
nani	real	-6.411671162
pēlā	real	-6.33949852
kopala	real	-7.876888752
tōkō	real	-8.820620537
leli'oi	real	-8.784794807
alawai	real	-5.222242832
māpū	real	-8.614605904
kuwala	real	-8.787291527
ka'ane'e	real	-8.365372658
pālaulau	real	-11.77537537
'unu	real	-6.989454269
ania	real	-4.428977966
kakale	real	-4.543409348
'aikola	real	-9.69204998
kā'eko	real	-9.703625679
waihoa	real	-9.66727066
hi'ukole	real	-12.47425079
kahiku	real	-12.5977211
'ālo'i	real	-9.905451775
nāwali	real	-5.095379353
nanau	real	-7.994868755
kalaina	real	-7.989592552
hohopa	real	-11.59190559
nāuki	real	-7.762751579
pā'e'e	real	-8.378982544
'ae'oa	real	-8.909065247

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
kā‘ao‘ao	real	-11.26722336
makiawa	real	-9.544631004
kiupe	real	-9.544168472
niakala	real	-9.534957886
‘ākeka	real	-7.86526823
kukule	real	-7.946296215
aumiki	real	-8.07348156
a‘uane‘i	real	-9.810448647
kikowā	real	-9.851573944
makaui‘i	real	-10.00142097
kīnohi	real	-11.67489338
makalē	real	-9.023338318
māka‘o	real	-8.122993469
lohaloha	real	-14.85793495
pāpeka	real	-5.622373581
malohi	real	-8.663482666
‘elekū	real	-8.682164192
kekeka	real	-4.338783264
kīkoni	real	-9.562833786
‘ūlala	real	-9.630556107
kilipohe	real	-10.56433201
pakūpakū	real	-11.19849777
kīpuni	real	-6.06061697
pa‘ū	real	-6.012728691
lolea	real	-6.068639278
nōhie	real	-8.148867607
‘uki	real	-7.303767681
‘alelele	real	-7.297337055
‘ele‘io	real	-8.80600071
wana‘ā	real	-8.848265648
lēhai	real	-8.847119331
kāmano	real	-8.93231678
hanopilo	real	-12.87966442
konawaena	real	-9.627272606
ki‘oki‘o	real	-9.795951843
‘umiki	real	-5.553042889

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
paika	real	-5.289076328
‘awelika	real	-6.654603958
kahuki	real	-9.048099518
‘eweureka	real	-9.057476044
māunu	real	-8.307061195
kōkala	real	-8.379011154
kumukolu	real	-8.456173897
ha‘aa‘e	real	-11.72482491
kūnāhihi	real	-15.20399189
‘ūmalu	real	-9.606907845
kalamena	real	-8.844560623
‘āpo‘o	real	-6.082174301
kaluhā	real	-12.1334343
‘eki	real	-4.289984226
iwiko‘o	real	-6.722875595
no‘ao	real	-11.11488819
moea	real	-5.264684677
hi‘ahi‘a	real	-16.98842049
koho‘oi	real	-13.84484673
mākea	real	-7.193945885
pāpaka	real	-7.218648911
‘aikena	real	-7.281620502
kēlia	real	-7.44151783
kenākenā	real	-11.98026276
‘ohina	real	-11.3328476
kūhea	real	-6.736413002
kūlipo	real	-6.693984032
lohena	real	-7.792936802
panepane	real	-6.996452808
kūhā	real	-7.081188202
lelepinao	real	-14.78543377
kokuli	real	-7.693444729
mōhā	real	-8.104430199
malineli	real	-7.934333801
‘ōmilo	real	-7.919732571
walowalo	real	-7.763154507

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
lākō	real	-7.75103426
pōliu	real	-8.05141449
koaea	real	-8.024353027
hīlia	real	-11.90575504
ku‘ikoli	real	-11.5912962
wai‘ele	real	-6.000901222
ni‘a	real	-5.891427994
mikaō	real	-10.08660793
‘apakau	real	-10.05701923
kāwala	real	-8.422576904
mokupawa	real	-8.507600784
newe	real	-8.426215172
wewena	real	-10.14646244
ko‘akā	real	-8.290102959
pakaiele	real	-12.55930805
pāpāia	real	-7.916690826
kū‘alu	real	-8.084831238
‘ūlina	real	-13.00224686
kaukukui	real	-14.24735451
hōkīkī	real	-14.19800663
kopalaka	real	-9.30178833
hailipo	real	-13.88978195
panakiō	real	-9.344166756
mikokene	real	-10.68481731
‘olā‘olā	real	-9.775078773
pāmalo‘o	real	-9.287760735
mahapili	real	-12.98677635
unuhia	real	-14.57492065
kuekene	real	-11.13377094
‘ilihia	real	-12.24252892
pikolō	real	-12.15986347
lelepinau	real	-15.06742954
nonu	real	-7.443098545
kaela	real	-8.437355042
palapa‘a	real	-8.896215439
kūkaelio	real	-11.48919487

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
‘āmēnia	real	-9.194122314
one	real	-6.082973957
lele‘oi	real	-7.358806133
hūhā	real	-11.44244766
pānūnū	real	-13.47277832
māhelu	real	-10.15094662
kaekaea	real	-10.0011692
‘ōwelo	real	-10.00179482
hamuhamu	real	-15.1592865
kiti	real	-8.170611382
manino	real	-8.121929169
pua‘ina	real	-8.33502388
nonono	real	-8.079780579
kuwai	real	-8.008245468
hauhia	real	-13.62642193
lolohe	real	-13.34134197
kelewa‘a	real	-10.12086391
uhikino	real	-12.96764851
waele‘a	real	-11.92180538
kahakea	real	-11.80242252
wulekula	real	-11.45703411
‘awakeke	real	-11.23853207
kauwili	real	-7.953804016
olokino	real	-7.335605621
pū‘ai	real	-7.661560059
wanilina	real	-8.354373932
kāheu	real	-8.312042236
pulo	real	-8.171453476
poa‘e	real	-8.333660126
kuāmano	real	-8.334859848
nāku‘i	real	-6.770824432
ikakani	real	-5.395383358
ehe‘ula	real	-10.73754787
nonea	real	-5.195623398
oheohe	real	-14.64626026
mā‘oi	real	-6.147466183

Continued on next page

Table E.1: Caption for the table (continued).

Stimuli	Type	Phonotactic Probability
‘oiai‘o	real	-7.018208027
‘onipā	real	-10.43241501
‘ā‘apo	real	-10.59801292
‘ohikui	real	-11.42397308
wīkini	real	-10.84556389
pu‘upau	real	-10.51801682
manihe‘u	real	-10.52303791
pūa‘o	real	-7.110728741
‘ukelena	real	-6.960884094
‘ilioki	real	-6.333875656
pekale	real	-4.70127821
papa‘ula	real	-8.884222031
‘Ihepa	real	-8.714972496
numela	real	-8.624432564
‘olomana	real	-8.633859634
kāhi‘o	real	-8.597659111
pikipiki	real	-7.024395466
wekelia	real	-6.955413818
mōkila	real	-6.409376144
pānoa	real	-6.704796791
ma‘uele	real	-6.71477747
kamapia	real	-6.262169838
haku‘epa	real	-12.64373207
‘eleke‘i	real	-7.920758247
pāloa	real	-5.35800457
wahu	real	-8.882231712
paikumu	real	-8.717336655
kiloū	real	-8.691037178