Variability of Retroflex Perception and Production in Heritage Tamil Speakers

Kamala Muthukumarasamy

A Thesis submitted to the Faculty of Graduate Studies in
Partial Fulfillment of the Requirements for the Degree of
Master of Arts

Graduate Program in Linguistics
York University
Toronto, Ontario

August 2023

© [Kamala Muthukumarasamy, 2023]

Abstract

This thesis investigates how social factors, such as input and identity, affect heritage Tamil perception and production of the alveo-dental and retroflex liquid contrast ([l]-[ɭ]). Heritage Tamil speakers participated in perception discrimination tasks, minimal pair production tasks, heritage language questionnaires, and sociolinguistic interviews. Quantitative results showed a high degree of variation in productive salience, as some speakers clearly produce the contrast while others did not. There was evidence of incomplete heritage language acquisition, dominant language substitution, and incomplete category formation. Perceptual distance was also variable, with some participants clearly showing categorical discrimination while others did not. Qualitative results revealed that a concrete embedding of the heritage language within their culture and a strong linguistic identity can serve as a reliable indicator of accuracy in perception and production. This research addresses whether acoustically fragile contrasts are realized in heritage Tamil, and importantly, how identity and language motivations serve to maintain them.

Keywords: heritage Tamil; perception and production; retroflex consonants; linguistic identity; language ideologies.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Chandan Narayan, for his support, guidance, and advice over the past year. I am very lucky to say that I have gained not only a dependable mentor, but a friend for life. Having someone so educated and experienced who understands the heritage Tamil context just as well as I do has made this whole experience much less daunting. Thank you to my second reader, Dr. Liz Smeets, for her careful eye and consistent encouragement, and committee member Dr. Thanujeni Pathman for her commitment, time, and effort in this endeavour.

A huge heartfelt thank you to the participants involved in this study. I could not have done it without you, and your interest in my work has encouraged me immensely. The Tamil community is a very special one, and I feel eternally grateful to be working on research that means as much to others as it does to me. Finally, a thank you to my parents, for their unwavering confidence in me, my sister, for being my loudest cheerleader, and my partner, for all the moral support and housekeeping. They have helped me get through this past year in one piece. It takes a village to raise a thesis, and they have put in the work.

Thank you sincerely for reading.

Table of Contents

List of Tables

List of Figures

1. Introduction

A heritage language is a minority language typically used in a household or small community within a larger majority language environment (Scontras et al., 2015; Montrul, 2016; Polinsky, 2018a). A minority language can be described as a substratum language, or a language which has little to no context for usage, is not the community or working language, and does not have any type of social or political pull. In contrast, the majority language can be described as the superstratum language, which is found throughout the community (i.e., business, schooling, and signage) and has incredible social and political impact; in Canada, English is the main majority language, followed closely by French. This thesis focuses on Tamil, a minority language spoken by immigrants and who settled in Canada. The subsequent generations who came to Canada in infancy or who were born in Canada can be described as heritage Tamil speakers. The language abilities of heritage language speakers comprise an entire spectrum, from low receptive abilities to high expressive abilities, suggesting that heritage speakers of a language are not all at the same level of fluency (Lynch, 2003; Montrul, 2016; Brown & Bousquette, 2018; Genesee et al., 2021).

The majority of research in heritage language acquisition, development, and maintenance has focused on Korean, Mandarin, Arabic, or Spanish, all language groups who have established populations concentrated in urban centres, thereby allowing for large scale studies on individuals' language abilities and overall linguistic patterning (Scontras & Putnam, 2020). Identification and observation of understudied heritage languages, especially in minoritized contexts, could provide insightful observations into how other heritage language speakers acquire and retain their heritage language, as well as how younger generations learn, use, and

maintain heritage-specific linguistic properties. The focus of this thesis is Tamil, an understudied language in the heritage language context. Importantly, the phonological contrasts investigated in this thesis are unique to Tamil and have not been described in the heritage language literature.

Heritage Tamil in Context

Tamil heritage speakers are defined in this context as speakers using a minority language that has been learned and used almost exclusively in a home environment or in specific cultural communities (Scontras et al., 2015; Montrul, 2016; Brown & Bousquette, 2018). The acquisition of this language does not reflect typical L1 or L2 language acquisition or development (Lynch, 2003; O'Grady et al., 2011; Polinsky, 2018a; Kasstan et al., 2018; Cognola et al., 2019). While these speakers have varying histories of language learning, it is typical for heritage language speakers to be exposed to their heritage language from birth and to receive much of the majority language input as they grow, thereby shifting their first language (L1) and second language (L2) structure. Simultaneous heritage speakers may end up relying on classifying their languages as their heritage language (Tamil) and their dominant language (English) (Canagarajah, 2012). It is important to note that heritage speakers use language in a wide variety of domains that vary individually, creating unique receptive and expressive, or productive, language abilities (Shea, 2019; Li & Matthews, 2022).

That is to say, there may not be a "one-size-fits-all" description of heritage speakers' abilities, owing to the wide variety of socio-cultural environments in which the heritage language develops. The intricacies of language abilities differ in heritage speakers due to the uniqueness of their language input and exposure, which is why their performance in receptive and expressive language tasks may not exactly conform to a native or L2 speaker's performance and level of

accuracy. These discrepancies can occur at the syntactic, lexical, or phonological level (Lynch, 2003; Montrul, 2016; Polinsky, 2018a; Jegerski & Sekerina, 2021). In addition, the variable linguistic environments and socio-cultural situations leading to varied processing and productive performance in heritage language speakers can be influenced by personal social factors like identity and motivation (Rampton, 1995; Tse, 2000; Bailey, 2005; Kasstan, 2017).

Heritage Language Phonological Perception and Production

Previous research into how phonetic and phonological inventories are established and maintained within simultaneous and sequential bilinguals suggest that they distinguish and retain two distinct phonologies from the beginning of language development and draw from those inventories for perception and production of each target language (Polinsky, 2018b; Flege & Bohn, 2021). Access to and switching between these phonological inventories often results in what has been described as a *bilingual cognitive advantage* which allows heritage speakers to perform similar to their native peers in language perception and production tasks or tests (Polinsky, 2018b). However, this phonological access does not apply or develop equally in all heritage speakers, resulting in incomplete acquisition, weakened representation of phonological features, documented variation in proficiency, and variation in perception and production abilities (Chang et al., 2011; Kim & Repiso Puigdelliura, 2019; Kim 2020, Jegerski & Sekerina, 2021). This variation is found to be impacted by linguistic as well as extralinguistic factors (Tse, 2000; O'Grady et al., 2011; Scontras et al., 2015). In particular, this thesis focuses on heritage Tamil retroflex liquid perception and production, and the social factors that affect the accuracy of this heritage-specific contrast discrimination.

Phonetic and Phonological Properties of Tamil's [l] and [ɭ]

The Dravidian language Tamil is primarily spoken in Southern India and Sri Lanka, each with unique dialectal and regional features. Tamil-speaking immigrants have been moving to Canada since the 1960s, but the biggest waves took place between the 1980s and early 2000s with refugees from Sri Lanka escaping conditions of civil war and ethnic genocide, and Indian immigrants relocating for greater economic and academic opportunities (Das, 2008; Das, 2011). There are now around 250,000 people of Tamil-speaking heritage living in Canada, with a majority settling in the Greater Toronto Area (Statistics Canada, 2021). The generations following these settled immigrants learn Tamil mainly in the home and acquire a familial dialect. Consequently, the heritage Tamil population comes with many variational dialects, such as place of origin dialects, regional dialects, and dialects affected by external or majority language environments (Keane, 2004; Canagarajah, 2019).

The presence of retroflex consonants in contrast with alveolar and dental consonants can be seen as a distinct feature of Tamil phonology, regardless of dialect and are found in all of the Dravidian languages (Narayanan, 1996). The alveo-dental [l] and retroflex [ɭ] liquids are found in the Tamil consonant inventory. These consonants can be described as liquids or as lateral approximants. Table 1 from Keane (2004) shows the two consonants labelled as lateral approximants in the bottom row. They are phonologically contrastive in Tamil, but this contrast has a relatively low acoustic salience relative to the dental and retroflex plosive contrast ([d̪]-[ɖ]), suggesting that the cue to the lateral distinction is less robustly separated in their perceptually relevant acoustic space than in stops (Narayan, 2019).

Table 1

Tamil Consonant Inventory from Keane (2004)

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar
Plosive	p (b)		t̪ (d̪)			ʈ (ɖ)		k (g)
Affricate					tʃ (dʒ)			
Nasal	m			n		ɳ	(ɲ)	
Fricative				s				
Tap				ɾ				
Approximant		ʋ				ɻ	j	
Lateral approximant				l		ɭ		

The acoustic similarity of the liquid contrast is evident in the formant structure characterising the difference. Narayanan (1999) showed that the difference between these liquids is characterized by the patterning of the relationship between their second and third formant values. The second formant (F2) reflects tongue frontness/backness, while the third formant (F3) represents constriction characteristics. In retroflex [ɭ] production, F2 and F3 values are much closer together, with a difference of around 1200 Hz, than they are in alveolar [l] production, with a difference of around 340 Hz (Narayanan, 1999). This comparison is shown in Figure 1, adapted from Narayanan (1999), with the alveo-dental liquid on the left and the retroflex liquid on the right. Table 2 summarizes the similarities and differences between these consonants in a visually concise way, showing difference in place of articulation (POA) and manner of articulation (MOA) alongside acoustic information.

Figure 1

Formant Structure in Consonants [l] and [ɭ] (here "/L/") as a smoothed LPC trace overlaid on FFT from Narayanan (1999).

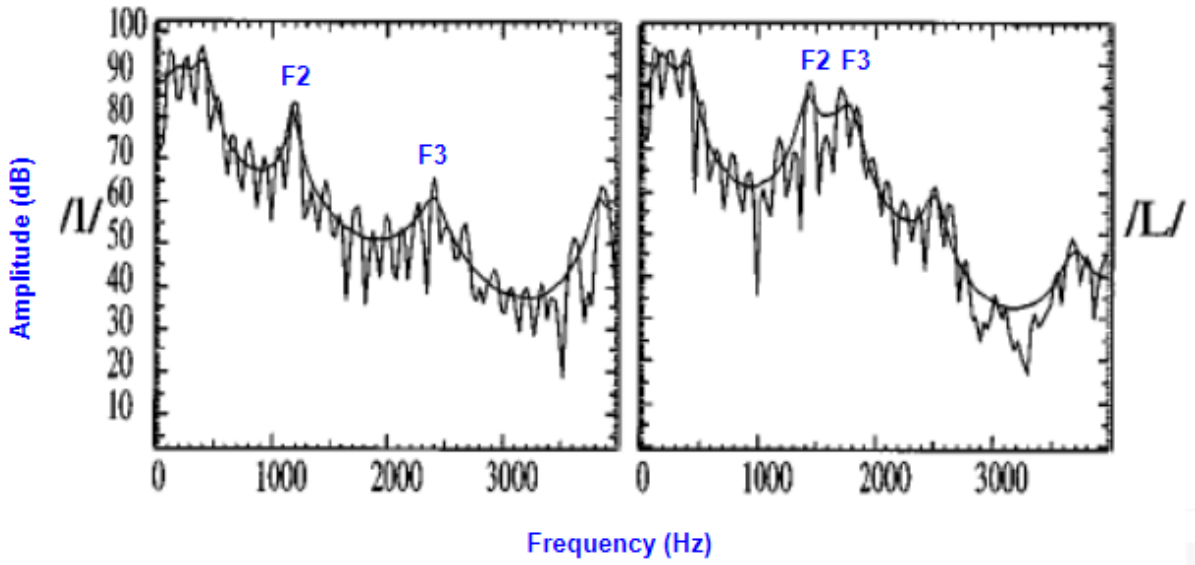


Table 2

Phonetic and Phonological Properties of [l] and [ɭ]

	POA	MOA	F2 (Hz)	F3 (Hz)	F3-F2 Window
[l]	alveo-dental	liquid	~1200	~2400	larger
[ɭ]	retroflex	liquid	~1460	~1800	smaller

Theoretical Framework

Since research into the development of heritage languages is quite new in the fields of bilingualism and language acquisition, many heritage language researchers rely on well-investigated L2 linguistic theory to inform and guide research. Flege's Speech Learning Model (SLM), recently revisited and updated to the Revised Speech Learning Model (SLM-r), is among these well-known second language acquisition models. The SLM and SLM-r have been well-

documented as effectively and productively guiding heritage language research (Chang et al., 2011; Chan, 2012; Fabiano-Smith & Bunta, 2012; Mazzaro, 2022; Rallo Fabra & Tyler, 2023; Yazawa et al., 2023).

Flege and Bohn (2021) posited the SLM-r, focusing on how L2 consonants and vowels are learned, retained, and organized within a speaker's phonetic and phonemic inventory. The SLM-r provides a realization of how a speaker's phonetic inventory forms, shifts, and evolves as they are continuously exposed to majority and target language input during naturalistic learning processes. Flege and Bohn (2021) theorize that if given sufficient and quality language input, a speaker can accurately perceive phonetic speech properties with accuracy. Flege and Bohn (2021) introduce the concept of *reorganization* of phonetic and phonemic categories when L2 input begins, and the reshuffling of these categories continues to change over time.

The SLM-r is based on three core principles:

1. That the phonetic and phonemic categories which are used in speech production are based on frequency distributions in input, which is a slow process.
2. That L2 learners use the same mechanisms and processes that children use when learning their L1.
3. That native and non-native specific differences in L2 production and perception are universal.

These principles allude to L2 acquisition mirroring L1 acquisition, resulting in similar levels of proficiency in each, which has not been the documented case. The contrasted outcomes of L1 and L2 acquired phonemic categories are explained through two internal phonetic operations, one, that L1 sounds initially substitute L2 sounds because the L2 sounds are

automatically linked and associated to sounds in the L1 phonetic inventory (substitution) and two, that pre-existing L1 phonetic categories interfere with and can potentially block the formation of new phonetic categories for L2 sounds (category blocking). In addition, one external environment factor is that the learning of L2 sounds is based on input that differs from the input that monolingual native speakers of the target L2 receive when learning the same sounds. At its core, the SLM-r aligns with similar speech learning models in its ideology that L2 phonetic and phonemic learning is heavily influenced and affected by L1 phonetic and phonemic inventories (Flege & Bohn, 2021).

One consequence of L1 influence on the creation of L2 phonetic and phonemic categories has been theorized as a perceived cross-language dissimilarity. This can be explained as how acoustically different or salient two phonetic properties between both languages are. If they have a greater perceived dissimilarity, the more likely a new phonetic and phonemic category will be created in the L2 inventory. If they have a low perceived dissimilarity, the category has a harder time forming, resulting in substitution or category blocking, which can hinder a speaker's perception and production of these L2 specific contrasts and sounds. The likelihood of a new phonetic or phonemic category being formed for these contrasts relies on three factors: One, the degree of perceived dissimilarity. Two, whether the closest L1 category has been defined precisely. Three, the quantity and quality of language input received. Variation in these factors is where differences are found between perceptive and productive abilities of these speakers between their L1 and L2 (Flege & Bohn, 2021).

However, the SLM-r has its shortcomings, as no theory can capture all the potential factors important to phonological acquisition and the creation of phonetic and phonemic categories. The first commentary on the SLM-r begins with theoretical issues raised by De Los

Santos and Alves (2021) surrounding the possibility of a relationship between language proficiency and category forming, the perception-production link, and the methodological impact of the model. They elaborate by stating that these blind spots can contribute to both theoretical and operational challenges in the field of bilingual development and psycholinguistic approaches to second language learning. To avoid these consequences, the authors suggest broadening research objectives to include these theoretical issues and keep bilingual cognition patterns in mind when analyzing data. They also emphasize the need to combine methods of data, *qualitative and quantitative*, to provide a more in-depth analysis. By utilizing these strategies, research informed by the SLM-r can contribute to an understanding greater than what the SLM-r has to offer for bilingual psycholinguistic research (De Los Santos & Alves, 2021). In addition to this, Gorba (2023) explored the weakness of one factor within the SLM-r in particular, the effectiveness of the Full-Time Equivalent (FTE) measure, which encompasses length of residence in an L2 environment and L2 language use. Though length of residence in an L2 environment has shown to have a significant effect on L2 and L1 production and perception, it does not have a significant effect when comparing populations or types of speakers. She tested this measure in a study with 19 L2 Spanish and 18 L2 English speakers performing a forced-choice two-alternative identification task including varying VOTs with bilabial and velar stops. She conducted correlation tests between FTE and task performance and found insignificant results, showing that FTE on its own is not a strong enough predictor of accuracy in perception tasks, and other factors need to be considered alongside FTE (Gorba, 2023). The responsibility that falls on the discussion and analysis of this thesis, being guided and informed by the SLM-r and using psycholinguistic methods, is to discuss possible factors *external* to the model that

could affect the quantitative results and highlight theoretical and methodological implications this perspective may have on psycholinguistic heritage language and second language research.

A theoretical framework that combines the SLM-r with qualitative analysis of individuals' extra-linguistic markers such as identity and motivation will inform the findings from the examination of alveolar-dental/retroflex liquid perception and production in heritage Tamil speakers. Though the SLM-r was developed to interpret phonological acquisition in L2 learners and speakers, the theory remains relevant to heritage speakers as even though their natural L1 is Tamil and natural L2 is English, this population learned both languages simultaneously, and consider English their dominant language. This unique permutation allows for the possibility of cross-language transfer, undergoing language attrition, language input dismissed as noise, and for incomplete or partially accessed phonetic and phonemic categories to exist (Gharibi & Boers, 2017; Polinsky, 2018b; Genesee et al., 2021; Li & Matthews, 2022)

2. Literature Review

Though research into language acquisition has broadly suggested that language input in infancy and childhood provide a speaker with developed accuracy in perceptive abilities (Chang, 2016; Polinsky, 2018b; Jegerski & Sekerina, 2021; Icardo Isasa, 2022), there are perceptual difficulties for infants and adults alike (Narayan, 2019). Research into how heritage speakers differ from native speakers and L2 learners has provided substantial insight into how these speakers acquire, perceive, produce, and retain language (Chang et al., 2011; Chang, 2016; Amengual, 2019; Kan & Schmid, 2019; Kim & Repiso Puigdelliura, 2019; Kim, 2020; Icardo Isasa, 2022; Seo et al., 2022). Additionally, research into how heritage speakers employ social factors to positively impact their heritage proficiency and abilities have also been well documented in multiple speaker populations (Comanaru & Noels, 2009; Alarcon, 2010; Wen, 2011; Shea, 2019; Canagarajah, 2019; Kim & Repiso Puigdelliura, 2019). Specific to heritage Tamil speakers, exploratory, preliminary, and introductory research has been collected on sociolinguistic structures surrounding their language environment, proficiencies, and strategies. Canagarajah has studied groups of heritage Tamil speakers in Canada, the USA, and the UK to find how social networks, family structures, and heritage language environments influence language use and proficiency in heritage Tamil speakers. He found that young heritage Tamil speakers stylize their language with code switching, using canned phrasing, semiotic resources, and hybrid methods of speaking to communicate (Canagarajah, 2012, 2013, 2019). These strategies, revolving around the dominant and majority language (English), had no influence or impact on individuals' cultural identity, as the ideology of the heritage Tamil itself was being shaped in a new modern perspective by heritage Tamil speaking youth. Since minimal psycholinguistic research into how these speakers perceive or produce heritage-specific

contrasts, there are many questions left unanswered in current literature. How do these social networks and ideologies surrounding proficiency and productive ability affect psycholinguistic processing of precise phonetic and phonological contrasts? How have heritage populations learned to perceive and produce heritage specific contrasts while navigating intricate social pressures in acquiring and maintaining a heritage language in a majority language environment?

Heritage Language Speech Perception

Research in heritage language speech perception has shown that heritage speakers hold perceptive abilities closer to native speakers than their L2 learning counterparts (Chang, 2016; Icardo Isasa, 2022; Seo et al., 2022). Findings also show that heritage speakers have a higher rate of success discriminating distinct sounds over similar sounds, as found in Kan & Schmid (2019). The following studies, on consonant, vowel, and suprasegmental perception in heritage-language speakers, highlight the perceptual differences between heritage speakers, native speakers, and L2 learners in various languages, as well as the patterning of heritage perception.

Chang (2016) explored bilingual perceptual benefits and deficits in a Korean heritage language context where the dominant language is American English. Participants included 28 native Korean speakers, 28 heritage Korean speakers, and 28 native English speakers. Chang hypothesized that since heritage speakers have the advantage of early onset to both languages and a unique language input experience, that they would mimic or outperform their native counterparts in perceptions of unreleased stops in Korean and English. Participants performed a Four-Alternative Forced Choice Identification Task. Participants heard a target word with a target stop, and then heard a second nonce word. Then they were given the option to choose which stop the word ended in /p/, /t/, /k/, or other. Three sets of stimuli were used: a set of

English minimal pairs, a set of English non-words, and a set of Korean non-words. Both English and Korean have these stops in their phonemic inventories, so the English non-words included /z/ and /ɹ/ to set them apart from the Korean non-words. Results supported the initial hypothesis, as heritage speakers performed just as well as native Korean speakers and outperformed native English speakers with all three sets of stimuli. These findings show that early acquisition of two languages can give bilingual speakers perceptual advantages in adulthood that monolingual speakers cannot achieve (Chang, 2016). The results of this study are relevant to the current study as it shows that heritage speakers have good perceptual abilities and a heightened sensitivity to phonological contrasts and identities due to their unique input and exposure experiences and effects.

Seo, Dmitrieva, and Cuza in 2022 examined the effects of crosslinguistic influence on contrastive stop perception and discrimination in Korean. Their study population included three key groups: 20 native Korean speakers, 20 heritage Korean speakers, and 20 L1 English-L2 Korean learners. A verbal narrative task was included to determine participants' proficiency in Korean, which included wordless pictures and a prompt to explain the story. Participants then performed an AX discrimination task using non words following a CV structure, using either lenis or aspirated stops [p, t, k, pʰ, tʰ, kʰ] with vowels [a, I, ɯ], resulting in 18 unique stimuli. Participants listened to contrastive or identical pairs and had to make judgements on whether the pairs were the same or different. Results showed that there was evidence of crosslinguistic influence and language dominance effects in the performance of L2 learners, but not in heritage speakers. Heritage speakers outperformed L2 learners with 85% accuracy in discrimination, compared to 65% accuracy; heritage speakers performed similarly to their native counterparts who scored 88% accuracy. Importantly, there was a positive correlation between heritage

proficiency and discrimination accuracy, suggesting that heritage speakers with higher expressive abilities also showed higher perceptive abilities (Seo et al., 2022). These findings are relevant to the current study as they show the relevant and impactful relationship between heritage language production and perception.

The literature on heritage language perception has also addressed vowels. Icardo Isasa (2022) studied the perceptual development of contrastive monophthong and diphthong vowels /e/ and /ei/ in 7 native Spanish speakers, 9 heritage Spanish speakers and 18 L2 Spanish learners, in an American English majority language environment. In order to measure the development of these speakers, Icardo Isasa implemented two types of instructions and included pre-test and post-test data collection periods to measure instructional impact. The types of instruction were high phonetic variability training (HPVT), which exposes learners to production of Spanish sounds from multiple speakers, sources, and contexts, and low variability of speech input (LPVT), where Spanish input comes from one source, the instructor. The pre-test and post-test measure was an ABX discrimination task, which included minimal pairs A and B spoken by one speaker, and a repeated A or B spoken by a second speaker. Participants were prompted to choose whether the repeated word X was word A or B. Results showed that heritage speakers perceived contrasts better than their L2 counterparts and do not differ significantly from their native counterparts. Effects of instruction showed that HPVT impacted heritage perception the greatest, showing the highest growth and produced the most similar performances to native speakers. These findings are important to the study at hand as they show that heritage speakers have perceptual abilities that differ from L2 learners and show more linguistic similarities to native speakers, as well as highlighting the importance of having naturalistic input from multiple speakers in multiple contexts.

Finally, a study conducted by Kan and Schmid in 2019 examined the tonal discrimination abilities in 64 L1 Cantonese speakers in Hong Kong and 67 heritage Cantonese speakers in America, aged 5-11. They aimed to test the credibility and validity of the Perceptual Assimilation Model for Suprasegmentals (PAM-S) and the L2 Intonation Learning Theory (LILT) at the intersection of child language heritage perception. An ABX discrimination task was administered using Cantonese tone pairs 1 (high level) to 4 (low-falling) and 2 (mid rising) to 5 (low rising) in a set of four minimal pairs in the words “tou” and “wai”. Results showed that heritage speakers were more accurate in distinguishing identical pairs over contrastive pairs, and that they struggled with the 2 to 5 tonal contrast more than the 1 to 4 tonal contrast. This is because the 1 to 4 tonal contrast has larger phonetic differences, making it more perceptually salient to discrimination in perception. Heritage speakers scored lower than their L1 counterparts in both sets of tone pairs. The factors that impacted the task performance the most were Chinese literacy abilities and frequency of tones 2 and 5 in language input. Lower literacy and lower frequency of tones 2 and 5 resulted in lower discrimination abilities (Kan & Schmid, 2019). These results are crucial to the current study as they show the impact of external linguistic factors on heritage language perception. Frequency and quality of language input affect perceptual abilities, as well as aptitudes in literacy, suggesting that linguistic competency in other areas of proficiency support heritage language perception. These findings also show the intricacy and importance of phonetic dissimilarity, as these speakers had more difficulty distinguishing tonal contours with low perceptual salience.

These studies in heritage language perception show that heritage speakers cannot be grouped with L2 learners solely based on expressive abilities and proficiency levels due to their differences in age of onset, or the age at which they were exposed to a target language, and input

qualities and quantities. Heritage speakers perform more closely to native speakers, but their acquisition differences result in task performance gaps that could be explained through language exposure and input experiences, as well as language domain differences and language use properties. These differences are directly correlated to perception and externally related to language proficiency as a whole affect heritage perception, which is why heritage perception should be studied in intersection with production, input, exposure, and social factors.

Regarding the gap between heritage perception and production, Kim (2020) investigated how heritage Spanish speakers recognized and used suprasegmental cues when applying lexical stress. The study population consisted of three key groups: 24 native Spanish speakers, 24 heritage Spanish speakers, and 20 L2 Spanish learners in an American English majority language environment. Stimuli consisted of 60 minimal pairs that differed in lexical stress position, mixed with 40 distractor words. This list of stimuli was used for the first perception experiment, as well as the second production experiment. To measure speech perception, a forced-choice identification task was conducted, where participants were to identify stress location in contexts with different suprasegmental cues. To measure speech production, 20 of the stimuli were chosen to be read aloud by participants, observing if they could use the proper cues and stress location. Results showed that heritage perception mimicked native perception, showing a clear perceptual advantage over L2 learners. Results also showed that heritage speakers and L2 speakers both differed from native distinguished correlate productions, by producing overlapped stress correlates lasting in longer duration. These findings conclude that early exposure to a language can support native-like perception but cannot guarantee native-like production (Kim, 2020). Crucially, they found that dominance in the heritage language, meaning it being the stronger language with the highest proficiency within a bilingual, has an impact on frequency of

standard linguistic production. The results are important to the current study as heritage speakers often have a larger receptive vocabulary and perceptual threshold than their expressive abilities and overall speech production, so this pattern could be observed in the population of heritage Tamil speakers as well. Kim (2020) evaluates perception as well as production, which the next section covers in more depth.

Heritage Language Speech Production

Previous research into heritage speech production has shown that heritage speakers have finely tuned abilities to produce heritage contrasts with higher accuracy than their L2 counterparts and dominant contrasts with higher accuracy than their native counterparts (Chang et al., 2011; Amengual, 2019; Kim & Repiso Puigdelliura, 2019). These abilities could be attributed to the early age of onset and exposure to both the heritage and dominant languages. The following studies highlight the differences in speech production between heritage speakers, native speakers, and L2 learners in various languages.

Kim and Repiso Puigdelliura (2020) examined heritage language dominance as it interacts with proficiency, use, and input. To test this interaction, they examined the production of the alveolar tap in 18 heritage Spanish speakers in an American English majority language environment. Rates of proficiency, input, and usage were collected through means of spontaneous speech, picture-naming tasks, self-reports, and a language background questionnaire. A drawing dictation task was administered to elicit Spanish rhotic consonants to ensure the phonological environment of the alveolar tap. Participants were placed in pairs and given two scenes, one with objects, and one without. One participant would dictate the objects and the other would draw the objects based on the description. Results showed that heritage

language dominance does not have an impact on frequency of accurate native-like productions of the alveolar tap. More importantly, increased language use, or elevated levels of heritage language speech production, and input were positively correlated with more accurate and more frequent production of Spanish alveolar taps in heritage speakers. These findings suggest that higher language use and input result in greater expressive abilities and higher accuracy in heritage Spanish production (Kim & Repiso Puigdelliura, 2020). The results are relevant to the current study as they show that language dominance does not affect heritage production, meaning that even heritage speakers with high English dominance (and consequently low Tamil use) could perform well on the targeted Tamil production task.

Chang, Yao, Haynes, and Rhodes (2011) studied the production of phonetic and phonological contrasts in Mandarin heritage speakers, alongside native Mandarin speakers and L2 Mandarin late learners in an American English majority language environment. They hypothesized that heritage speakers could outperform their L2 counterparts in language-internal phonological contrasts and cross-linguistic phonetic contrasts in both Mandarin and English. Three areas of production were targeted to test this hypothesis: the production of back vowels, Mandarin aspirated and English voiceless plosives, and Mandarin retroflex and English palato-alveolar fricatives. Across all three sets of productions, heritage speakers maintained the language-internal and cross-linguistic contrasts more accurately than their native and L2 counterparts, supporting the hypothesis that heritage speakers' dual exposure from a young age allows them to distinguish phonetic and phonological contrasts that are difficult for L2 learners of Mandarin and English. These findings can be attributed to phonetic and phonological properties being categorized and distinguished at an early age of onset can ensure accuracy in

heritage and dominant language productions. The results of this study are important for the current study as they show heritage production benefits in a bilingual context.

The effects of either simultaneous or sequential bilingualism on heritage language production were teased apart in a 2019 study by Amengual who investigated heritage Spanish speakers' production of an allophonic variant. The study included three key groups: 10 simultaneous heritage bilinguals, 10 sequential heritage bilinguals who learned Spanish first, and 10 L2 Spanish learners. The stimuli were embedded in the form targeted production task with 60 sentences including the target allophones, 30 with word-initial vocalic stops and 30 with word-medial intervocalic voiced stops, which participants recorded twice. Acoustic measurements were taken at the lowest intensity of the prevocalic target consonant and the intensity peak of the preceding vowel, resulting in a measurement of intensity difference. Results showed that sequential bilinguals outperformed their simultaneous and L2 counterparts in accuracy of allophonic variant productions. These results show that early onset to a heritage language, between the ages of 5 months and 5 years, carries through to articulatory and acoustic accuracy in heritage Spanish production in adulthood (Amengual, 2019). These findings are germane for the present study as adult heritage speakers all have varying exposure and age of onset backgrounds, and these results differentiating simultaneous bilinguals from sequential bilinguals could shed light on the production results.

The studies reviewed above show that language experience, input, type of bilingualism, and age of onset are all important factors to consider when examining heritage speech production abilities. Since the heritage experience is so diverse and varies due to multiple linguistic factors, it is crucial to consider the intersectionality of what goes into heritage production, as it is more complex to explain discrepancies and exceptionalities.

Effects of Language Input and Identity

When reviewing research into heritage language proficiency and its interaction with frequency of input, type of exposure, language attitudes, beliefs, and use, findings show that higher levels of language exposure and input yield higher linguistic capabilities (De Houwer, 2007; Kim & Repiso Puigdelliura, 2019; Shea, 2019; Daskalaki et al., 2020). The literature also importantly suggests that positive attitudes toward heritage language use contribute to higher linguistic output and abilities (Comanaru & Noels, 2009; Alarcon, 2010; Wen, 2011; Canagarajah, 2019).

Comanaru and Noels (2009) studied Chinese speakers to understand how motivations and beliefs surrounding language acquisition influence language learning and retention. Their objective was to specifically focus on heritage language learning and to examine how language attitudes differ based on type of learner. They assessed a group of 72 native speakers, 36 heritage speakers, and 33 L2 learners through survey methods in the form of a questionnaire, obtaining qualitative data on engagement in learning, personal motivation, aspirations, beliefs, community engagement, and reasons for language development and retention. Their general findings showed that if learners found heritage language learning meaningful, the more eager they were to develop their language abilities. Cultural connections to the language community and control over learning goals also showed more positive attitudes towards heritage language maintenance. Results showed that heritage speakers strongly indicated they were ameliorating their Chinese proficiency levels because it was an integral aspect to their self-concept and identity. They also indicated they felt a sense of obligation to retain their heritage language, and therefore had a responsibility to learn the language. These results suggest that motivation for heritage language learning is closely linked to personal sociolinguistic factors, which vary from speaker to speaker,

emphasizing the importance of *qualitative* data alongside quantitative data to provide an accurate portrayal of the variation in heritage speaker populations.

Shea (2019) examined how language experience, dominance, and proficiency affected 15 heritage Spanish speakers' bilingual vowel production. The measures used to quantify proficiency and dominance included a questionnaire on language use and picture-naming tasks, and expressive vocabulary scores. The measure used to investigate speech production was a targeted production task, in which participants read lists of three words followed by a fourth non-word, all with target vowels [i, ɪ, e, ɛ, æ, ɑ, ʌ, o, ʊ, u]. Results showed no significant patterns in speech production across the population of heritage Spanish speakers. The discussion surrounding these results supports the outcome completely. Shea explains that proficiency and language dominance can interact with results independently or together, but it is difficult to tease them apart. Also, that variability among heritage speakers is common, as they each have individual acquisition experiences and use their heritage language in different domains and contexts. Another finding of this study is the consideration to investigate more social factors attached to language proficiency and dominance, including ideologies of language attitudes, self-identity, and cultural affinity (Shea, 2019). The results of this study are crucial to the current study as the recommendations laid out in the discussion and conclusion are the very motivations to this thesis. Recognizing there are socio-phonological factors affecting heritage production is the first step to investigating what these factors are and how they exactly impact areas of speech production.

Finally, in 2011, Wen investigated the learning motivations of heritage Chinese speakers enrolled in Chinese university-level courses. His research objective was to study motivational constructs on a qualitative and quantitative level in order to compare learning motivations and

strategies among different populations. The sample population of 317 were divided into three groups, 118 bilinguals who have native speaking parents, or natural heritage speakers, 58 heritage-connected learners who have a cultural relevance to learning Chinese, or those with one native parent or a connection to Chinese ancestry, and 141 non-Chinese L2 learners. He presented a two-part questionnaire adapted from the Attitudes/Motivation Battery (Gardner, 1985, as cited in Wen, 2011, p.339), in which the first part was to evaluate the speakers' demographics, language attitudes, and motivations. The second portion was administered as a sociolinguistic interview to collect further information on motivational reasoning, social network, cultural interest, learning engagement, methods, and strategies. Generally, Wen found that positive attitudes and experience was the most predictive factor of motivational effort or magnitude. Heritage speakers were found to opt into further heritage language learning due to socio-cultural influences, forming a self-identity, or recovering parts of their cultural heritage. They also valued learning their heritage language for purposes of language usefulness, such as communicating with elders, family members, community, or using Chinese for academic or economic opportunity. The sociolinguistic interviews found that heritage speakers assigned cultural value and individual interest to their personal heritage language experience.

In sum, these studies regarding personal language attitudes, beliefs, and motivations show that these social factors have a consequential impact on heritage language development, ability, output, and use. When heritage speakers have control over their learning, the opportunity to create a cultural and linguistic identity, and the motivation to acquire and maintain their heritage language, the better their proficiency and abilities will be overall. Maintaining these factors can ensure higher accuracy in production and learning motivation.

3. Current Study

Though research in heritage perception and production has brought forth general findings, recurring patterns, and guiding methodologies, they focus on other heritage languages like Mandarin, Korean, and Spanish, as well as other phonological properties, such as stops, vowels, and tones (Chang, 2016; Amengual, 2019; Kan & Schmid, 2019; Scontras & Putnam, 2020). There has been no previous research on heritage Tamil perception or production, nor have there been studies to examine how sociolinguistic factors intersect with these psycholinguistic domains of heritage Tamil acquisition and retention. Previous sociolinguistic studies also focus on the metric of "ethnic orientation and ethnic identity", or how strongly one aligns themselves with their ethnicity, culture, and linguistic identity (Hoffman & Walker, 2010; Matsunaga et al., 2010; Nagy et al., 2014; Nagy, 2018). The current study aims to capture the perceptive and productive abilities of heritage Tamil speakers in a Canadian English majority language environment, and to investigate what factors might influence high perception and production abilities for heritage language retention and maintenance, as well as to tease apart the ideologies of cultural identity (CI) and linguistic identity (LI).

Research Questions and Hypotheses

Inspired by previous research into heritage language development, retention, and maintenance, as well as the objective to capture how sociolinguistic factors like identity and input affect heritage Tamil perception and production, the following research questions can be made:

1. Can Heritage Tamil speakers accurately distinguish retroflex liquid consonants from non-retroflex liquid consonants?

2. Can Heritage Tamil speakers accurately produce retroflex liquid consonants from non-retroflex liquid consonants?
3. What exposure and input factors affect accuracy in discrimination and production of these consonants?

Informed by previous research in the field of heritage language acquisition, perception, and production, as well as the sociolinguistic research into the impact of positive language attitudes, beliefs, expectations, and identity, the following hypotheses can be made:

A. Heritage Tamil speakers should be able to discriminate similarly to their native speaking peers, and to produce this contrast due to early age of onset and acquisition but will have more difficulty discriminating contrasts with low perceptual salience.
B. Increased language input and experience, as well as positive language attitudes and beliefs, should be positively correlated with accuracy in perception and production.

Participants

Recruitment occurred between February and April of 2023 through targeted community outreach, the snowball effect, and word of mouth. Study advertisements were spread via social media posts on Facebook and Instagram, by emailing points of contact, and study poster dissemination around York University's campus. Various Tamil cultural groups were contacted, including the York University Tamil Student Association, associations at several Toronto universities, Tamil Sangams (cultural organizations), and through Tamil communities of practice. At the end of the recruitment period, 18 heritage Tamil speakers and 2 native speakers participated in this study. To clarify, native speakers can be described as first-generation speakers, and heritage speakers can be described as second-generation speakers, however, this

thesis will refer to them as native and heritage to align with previous literature in the field of heritage language perception and production. Table 3 shows in-depth demographic information on all participants collected from an administered questionnaire (see sec. procedure). The mean heritage age was 23 years, and the mean native age was 55 years. The mean heritage age of arrival was 7 months, and the mean native age of arrival was 29. The sample population consisted of 13 Eelam Tamil speakers and 7 Indian Tamil speakers. All participants had an age of onset of 0, as they were exposed to Tamil from birth, and continuously at different rates throughout their upbringing. All participants received some amount of Tamil input and have circumstances in which they speak Tamil presently in their day-to-day adult life.

Table 3

Participant Demographic Information

Subject	Type of Speaker	Age (years)	Age of Arrival (years)	Self-Reported Tamil Proficiency (X/5)	Relative Input (%)	Relative Output (%)
N1	Native	57	31	5	0.25	0.75
N2	Native	54	28	5	0.25	0.5
H1	Heritage	21	0	4	0.75	0.5
H2	Heritage	21	0	3	0.38	0.3
H3	Heritage	24	2	3	0.33	0.33
H4	Heritage	21	1.5	5	0.63	0.7
H5	Heritage	23	0	3	0.42	0.4
H6	Heritage	25	0	2	0.38	0.2
H7	Heritage	18	12	5	0.63	0.56
H8	Heritage	21	0	4	0.63	0.56
H9	Heritage	19	3	4	0.43	0.4
H10	Heritage	30	0	2	0.17	0.1
H11	Heritage	18	0	4	0.5	0.33
H12	Heritage	22	3	4	0.58	0.53
H13	Heritage	27	0	5	0.63	0.56
H14	Heritage	23	0	3	0.5	0.46
H15	Heritage	28	2	3	0.5	0.33
H16	Heritage	21	5	2	0.46	0.33
H17	Heritage	19	0	2	0.33	0.33
H18	Heritage	26	0	1	0.16	0.1

*Note AOA 0 denotes the participant was born in Canada.

4. Methodology

Study Design

The study consists of three measures, conducted in the same sequence for every participant, in person at York University's Speech and Psycholinguistics Lab, over the course of a 45-minute data collection session. First, participants were engaged in a one-on-one language interview in English to extract data on social and personal language proficiencies, usage, and habits. Following this, participants judged contrastive and identical non-word pairs in an auditory discrimination task to assess the speech perception abilities of heritage Tamil speakers. Finally, participants orally produced Tamil minimal pairs in a targeted production task to examine the acoustics of their lateral liquid and retroflex liquid productions. A few weeks later, participants took part in a follow-up sociolinguistic interview in English, which took between 10 and 25 minutes to administer, to provide additional information on cultural and linguistic identity, as well as language attitudes, beliefs, and motivations. Participants were compensated for their time with a $15 gift card upon study completion. All procedures were consistent with York University's Office of Research Ethics protocols.

Procedure

Language Input & Exposure - Heritage Language Questionnaire

The heritage language questionnaire was comprised of thirty-seven questions adapted from three main sources, the Alberta Language and Development Questionnaire (Paradis et al., 2010), the Language History Questionnaire (Lee-Ellis, 2012), and the Heritage Signers: Language Profile Questionnaire (Isakson, 2016). The purpose of the questionnaire was to

provide a complete and descriptive background for each heritage speaker to cross analyze their individual language experiences, instances of input and exposure, social language factors, language habits, and attitudes with their task performance. Questions were outlined to retrieve information about age of onset, amount of input and exposure, family language habits, spoken, written, and reading proficiencies, and overall language usage. Each question had a rating score on a scale from 1 to 5, and the scoring template closely followed the Alberta Language and Development Questionnaire (Paradis et al., 2010), calculating relative input and proficiency. The heritage language questionnaire was conducted in an interview style with the participant by the principal investigator, an in-group member, to foster a comfortable environment and prime the participant to get in the heritage Tamil mindset. Since a university campus does not foster an environment that requires heritage Tamil, the participants need to be code-switched in a way in the lab environment for a more reliable and reflective performance in this task. The full set of interview questions can be found in Appendix A.

Speech Perception - AX Discrimination Task

The auditory discrimination task is constructed through an AX study design influenced and adapted from stimuli as seen in Narayan (2008), Bennett et al. (2018), and Seo et al. (2022). Target consonants [l, ɭ] and distractor consonants [d̪, ɖ, n, ɳ] were placed in a VCV non-word structure surrounded by one of four vowels [a, i, u, o]. This list of 24 non-words were recorded by an adult male native Tamil speaker in a soundproof booth three times, resulting in 72 sound files. Table 4 shows the average and standard deviation in the F3-F2 in the alveo-dental and retroflex stimuli. Consistent with Narayanan (1999), the acoustic difference between F3 and F2 is considerably smaller in the retroflex liquid than the alveo-dental liquid.

Table 4

Lateral Approximant Stimuli F3-F2 Acoustics

Target Liquid	Mean (Hz)	Standard Deviation (Hz)
Alveo-Dental F3-F2	1371.39	330.37
Retroflex F3-F2	491.68	196.42

Choosing non-words were essential to this speech perception task as heritage Tamil speakers on an individual basis receive immensely different variations of the language, amounts of Tamil input, and levels of exposure. By using non-words, any proficiency or vocabulary bias would be eliminated, ensuring that participants were to use their phonological processing abilities rather than their vocabulary knowledge to differentiate the stimuli. Administering an AX discrimination task has been well documented (Narayan, 2008; Chang, 2016; Icardo Isasa, 2022; Seo et al., 2022) to capture perceptive abilities of speakers, as it is relatively easy to design and produce for running in a lab, it requires minimal equipment, and it provides multiple types of quantitative data, such as perception rates of accuracy, reaction timing, and offline psycholinguistic processes.

These sound files, ranging from 47ms to 50ms, were then assembled in sequence PRAAT (Boersma & Weenick, 2022) with a 500ms interstimulus interval (ISI), resulting in 432 trials. The 500ms ISI was important for participants to rely on phonological properties and characteristics of the sounds instead of acoustic properties and characteristics (Narayan, 2008). This difference was reasoned to make it easier to observe the phonological perception of these heritage Tamil speakers. The AX task was then coded into PsychoPy (Peirce et al., 2019), a software designed for psycholinguistic studies. The AX discrimination task was coded to run through the stimulus, followed by a decision screen prompting participants to choose whether the

two sounds they heard were the same or different, denoted by two labelled stickers on the keyboard. Participants pressed [X] key if the pair was deemed “same” and [Y] key if they were deemed “different.” The full set of non-words can be found in Appendix B.

Speech Production - Targeted Production Task

The production task was administered through English to Tamil translation prompts, which included a list of twelve Tamil words organized in a set of six minimal pairs. A translation task was chosen over a reading task or a picture-naming task as some of the target words were not picturable, and not all heritage speakers had Tamil literacy skills. The set up for this task was adapted from multiple sources (Kim, 2019; Vijayakumar et al., 2021). Words were chosen by native and heritage Tamil speakers and deemed appropriate by both Indian Tamil and Sri Lankan Tamil speakers to ensure the list included words realistically known to heritage Tamil speakers. Each word either had the alveolar lateral liquid [l] or the target retroflex lateral liquid [ɭ] word-medially or word-finally to elicit a targeted production of each consonant. The experiment was set up with participants wearing a high-quality headset microphone running into a digital-audio workstation connected to a computer. The audio-recording software Audacity was used for recording alongside PowerPoint for prompting the production, in which one target word in English appeared onscreen at a time for the participant to translate into Tamil. This ensured one elicited token produced without list, prosody, or semantic bias. Each stimulus repeated randomly three times throughout the sequence, resulting in 36 audio recordings to analyze consistency and accuracy of the productions. Audio was recorded in the lab with a microphone and saved as a WAV file. The full set target words can be found in Table 5. Though the orthography is in Tamil in Table 5, no Tamil orthography was included in the procedure, as some heritage Tamil

speakers did not have Tamil literacy skills, prompting an English to Tamil translation task to keep it an attainable task for those with and without Tamil literacy skills.

Table 5

Minimal Pair List for Elicited Translation Task

Alveo-Dental Variant	IPA	Translation	Retroflex Variant	IPA	Translation
தோல்	θɔl	Skin	தோள்	θɔɭ	Shoulder
பாலம்	pɑːləm	Bridge	பழம்	pɑɻəm	Fruit
வலி	vəli	Pain	வழி	vəɻi	Way
புலி	puli	Tiger	புளி	puɻi	Tamarind
நீலம்	niləm	Blue	நீளம	niɻəm	Long
மலை	məleɪ	Mountain	மழை	məɻeɪ	Rain

Language Attitudes & Motivations - Sociolinguistic Interview

Finally, a casual follow-up sociolinguistic interview by the principal investigator, an in-group member, was conducted by asking six questions and prompts regarding the personal importance to acquiring and speaking heritage Tamil, the personal motivations to continue speaking the language, reasoning for continuing education and learning heritage Tamil, and how the language itself is embedded within Tamil culture. This interview took place over Zoom a few weeks after the lab session and took between 10 and 25 minutes to administer. This measure was taken post-testing as there was a concern that prompts about specific linguistic features could possibly prime the participants towards an explicit Tamil bias in pronunciation. Administering it post-test ensured no priming effect was observed. Participants were also asked to create a pie

chart of their Tamil identity, and to explain how large a piece their heritage language takes up on that chart to get a visual idea of how they value knowing heritage Tamil. Questions surrounding the division between cultural identity and linguistic identity were explored. Participants were encouraged to provide as much detail as they could and to be honest with their motivations and responses. The full list of questions and prompts can be found in Appendix C.

Analysis

Data was analyzed from all four measures. Statistical analysis was done through correlation tests in coding software R Studio to see specific correlation trends between heritage perception D-prime values, heritage production productive salience values, and factors of weighted heritage language input and self-rated proficiency.

Both qualitative and quantitative data were retrieved from the heritage language questionnaire. Relative language input and output were calculated with a weighted average to norm the data across individual language experiences. Relative weighted inputs were calculated self-reports of how often English or Tamil is used in a bilingual setting. So, if a bilingual setting could see 100% maximum of either Tamil or English, participants were asked what percentage of each language they hear every day, whether it be 50% Tamil and 50% English, or any other permutation. This number was assigned for hearing and speaking to all members of the family and calculated with a weighted average, showing how much relative Tamil input the participant hears overall. Age of onset, years in heritage language schooling, and self-rated proficiency levels were also recorded. Responses regarding social networks, media use, language beliefs, attitudes, and motivations were transcribed into a Word document and Excel spreadsheet for qualitative analysis of recurring patterns and language ideologies.

Heritage language perception was assessed from performance on the AX discrimination task. Reaction times, task scores, and overall accuracy were recorded by the software for data analysis. Using overall task accuracy scores, a D-prime analysis was conducted by first recording each amount of hits, or contrast identified correctly, misses, or contrast identified incorrectly, correct rejections, or identity identified correctly, false alarms, or identity identified incorrectly. D-prime is a reliable signal detection statistic that considers hits, false alarms, correct rejections, and misses, and provides a value for a participant's bias-free perception of identity and contrast pairs (Narayan, 2008). This was done by calculating the hit rate, or different stimuli are marked as different, and false alarm rate, or same stimuli are marked as different, and inputting them into a D-prime matrix to obtain a D-prime value, which evaluates how accurately a subject can perceive these contrast and identity pairs (Keating, 2005). The lower the D-prime value, the less perceptible the alveolar and retroflex liquids are to the participant, and the higher the D-prime value, the more the liquids are perceptually different to the participant.

Quantitative data on productive performance was retrieved from the elicited production task. Each token was isolated for analysis within PRAAT. First, the total duration of the alveolar or retroflex liquid was measured, then the middle of said duration was selected to measure the F2 and F3 values. Once all F2 and F3 values were obtained, each F2 value was subtracted from its F3 value to create a window value, measuring the salience of the liquid production. Then, alveolar windows and retroflex windows were averaged, and the retroflex window average was subtracted from the alveolar window to obtain a "salience" measure. The purpose of this analysis is to show that the larger the gap between the alveolar and retroflex windows, the speaker's production is more acoustically salient or perceptible. Sample tokens from participant H6 producing the alveo-dental liquid in the word /məleɪ/ is shown in Figure 2 and producing the

retroflex liquid in the word /məlɛɪ/ in Figure 3. Each figure shows the acoustic wavelengths, with the spectrogram underneath, and a spectral slice at the consonant onset to the right. The FFT spectra is taken at the onset of the lateral consonant with an LPC envelope shown in red. The duration of each liquid has been highlighted in red on the left. F2 and F3 measurements are labelled in blue on the right.

Figure 2

Waveform and Spectrogram Representations Alongside Formant Measurements for an Alveo-Dental Sample Token

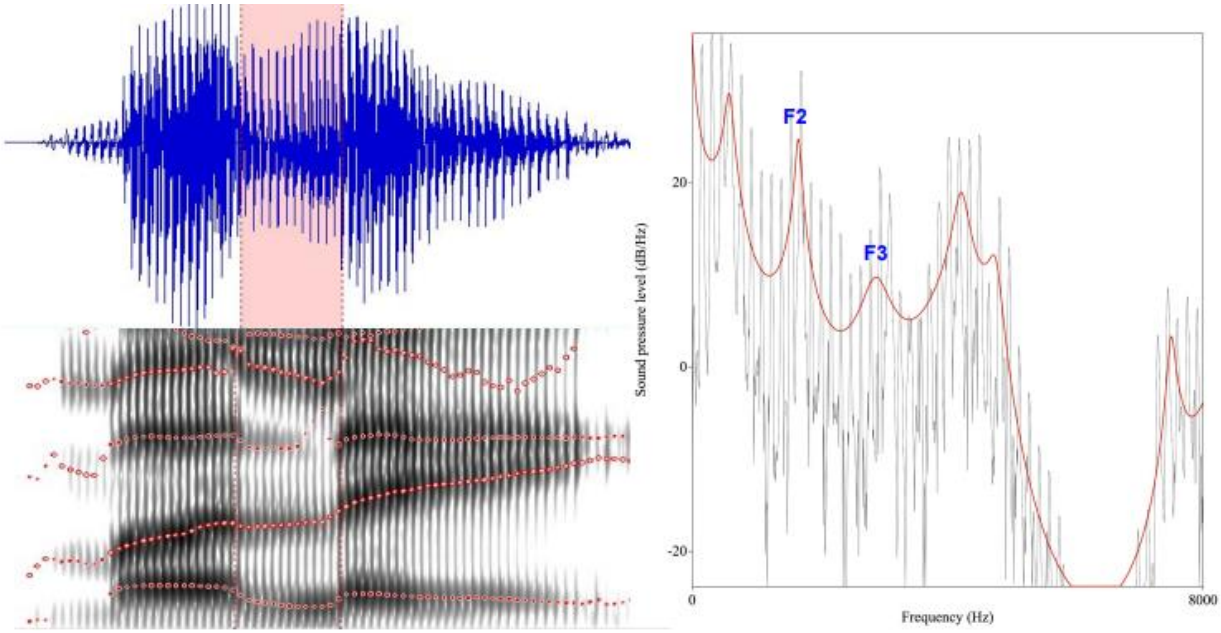
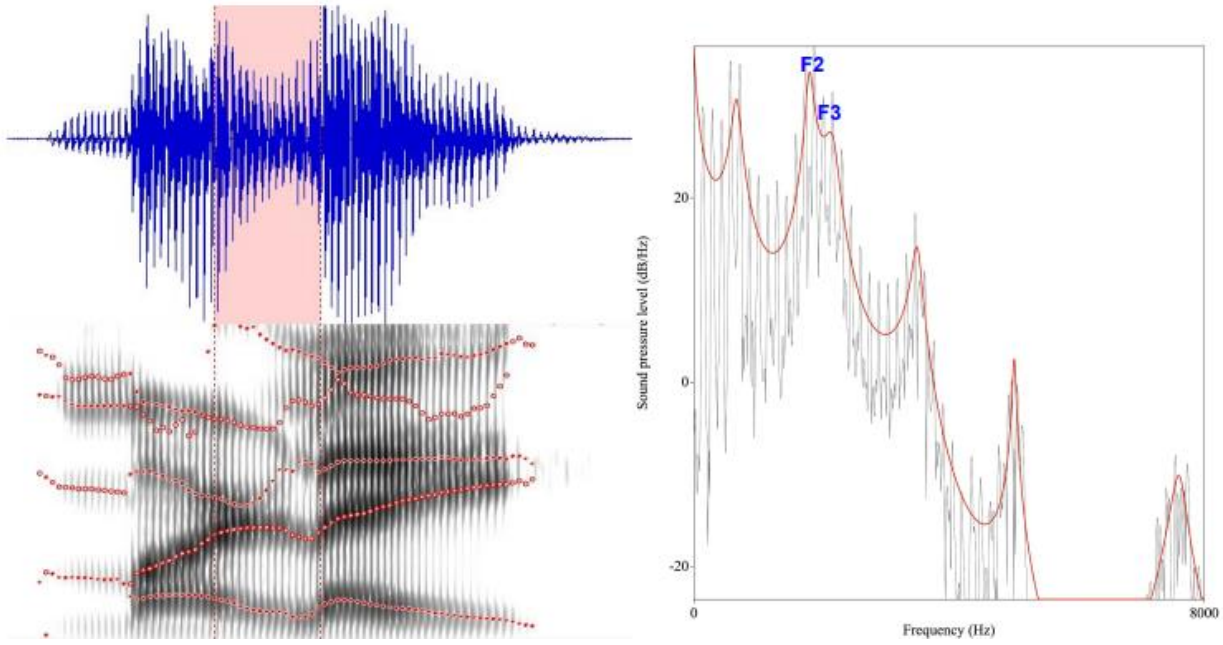


Figure 3

Waveform and Spectrogram Representations Alongside Formant Measurements for a Retroflex Sample Token



Qualitative data was retrieved from the follow-up sociolinguistic interviews. Responses were recorded on Zoom and relevant answers were transcribed for anonymized quotation. Individual responses were tallied for similar ideologies like speaking the language for reasons of communication, family, self-identity, or obligation. Spoken information regarding cultural identity, linguistic identity, proficiency targets, learning opportunities, language attitudes, and motivations were collected and transcribed for discussion of results alongside quantitative data pulled from the other tasks. Repeating patterns of ideologies or themes were recorded in an Excel spreadsheet for anonymized quotation.

5. Results

General quantitative task results are summarized in Table 6. Results are divided into two sections, first being quantitative results analyzed from the questionnaire, perception, and production tasks, followed by qualitative results taken from the questionnaire and the sociolinguistic interviews. Quantitative data from the two native speakers were included in comparison for analysis between heritage speaker perception and production data as they provided control data, being one Indian male and one Sri Lankan female who learned Tamil as an L1 in an immersed language environment.

Questionnaire Results

Results from the heritage language questionnaire showed that all participants heard and spoke some amount of Tamil every day. All participants had lower rates of Tamil output, or language usage, ranging from 10% to 70% of the time, than Tamil input, or how much Tamil they heard through people or media, ranging from 16% to 70% of the time (this is a relative frequency calculation, measuring only the time Tamil could be used in specific bilingual settings, not including overall input in English throughout the day). A total of 9 participants reported levels of literacy and 7 participants reported levels of orthographic knowledge. A total of 11 participants received between 4 and 14 years of heritage language education during their childhood, but no participants had attended any Tamil lessons as adults. A total of 4 participants stated that their time in heritage language education did not contribute to their current level of Tamil productive or spoken proficiency. A total of 17 participants responded that they wanted to ameliorate or develop their Tamil skills further, and that they were unsatisfied with their proficiency levels. A total of 14 participants stated they had some kind of Tamil friend group, but

only 5 participants stated they used Tamil when socializing with them. Frequency of meeting times with Tamil friend groups ranged from seeing each other every day to once every six months. Those who met every day had instances of speaking Tamil with each other, but those who met infrequently relied on communicating in English mostly. Only 3 participants felt a non-existent to low sense of Tamil identity; 7 stated they felt a sense of Tamil identity moderately; 8 stated they felt a sense of Tamil identity strongly. This general self-identity rating did not correlate significantly ($p>0.05$) with perception ($r = 0.18$) or production ($r = -0.25$) levels.

Perception and Production Results

Table 6

General Perception and Production Results

Task	Mean	Standard Deviation
Heritage Overall Discrimination Accuracy	91%	0.08%
Heritage Liquid Discrimination Accuracy	69.56%	9.93%
Heritage Alveo-Dental F3-F2 Production	1389.59 Hz	388.84 Hz
Heritage Retroflex F3-F2 Production	1088.77 Hz	468.19 Hz

In the AX task overall, heritage speakers scored 91%, with a range of 67% to 97%. Native speakers scored 94.5% on the AX task overall. Regarding retroflex liquid perception, heritage speakers were accurate 69.56% of the time with a range of 50.69% to 92.36%. Native speakers had 89.58% accuracy with a range of 86.80% to 92.36%. Heritage speakers made an average of 29 contrast errors or marking two different sounds as the same. Participants made an average of 12 identity errors or marking two same sounds as different. This suggests that participants had a harder time differentiating different sounds than the same sounds. Also, participants had a harder time discriminating the [l] and [ɭ] liquid contrast than the [d̪] and [ɖ]

stop contrast. The average number of errors for the liquid contrast was 40, with an accuracy of 72%, while the average number of stop errors was 14, with an accuracy of 90%. This means that the stop contrast with a higher perceptual salience was easier to discriminate compared to the liquid contrast with lower perceptual salience. There was no apparent difference between average reaction times between contrast and identity pairs, as the averages were 2.59s (*SD*=0.34s) and 2.56s (*SD*=0.35s) respectively, suggesting it took around the same time to process the stimuli and decide on whether the sounds were the different or the same, regardless of the type of phonemic contrast.

As for alveo-dental and retroflex liquid production, results showed no apparent patterns in liquid production or formant measurements, as there was immense variation in levels of productive salience. Some speakers had a distinguished alveo-dental and retroflex contrast in their productions, and some speakers produced both liquids with an alveo-dental F3-F2 window. Overall, averaging all heritage speakers and their productions of all alveo-dental and retroflex productions, retroflex productions showed a lower F3-F2 difference window than alveo-dental productions in Figure 4. On average, the F3-F2 difference was 1389 Hz (*SD*=388.84 Hz) for retroflexes and 1088.77 Hz (*SD*=468.19 Hz) for alveo-dentals. Though consistently lower, the ranges are very similar, showing the low perceptual salience of this contrast.

Figure 4

Productive Salience Between Alveo-Dental and Retroflex Liquid Productions

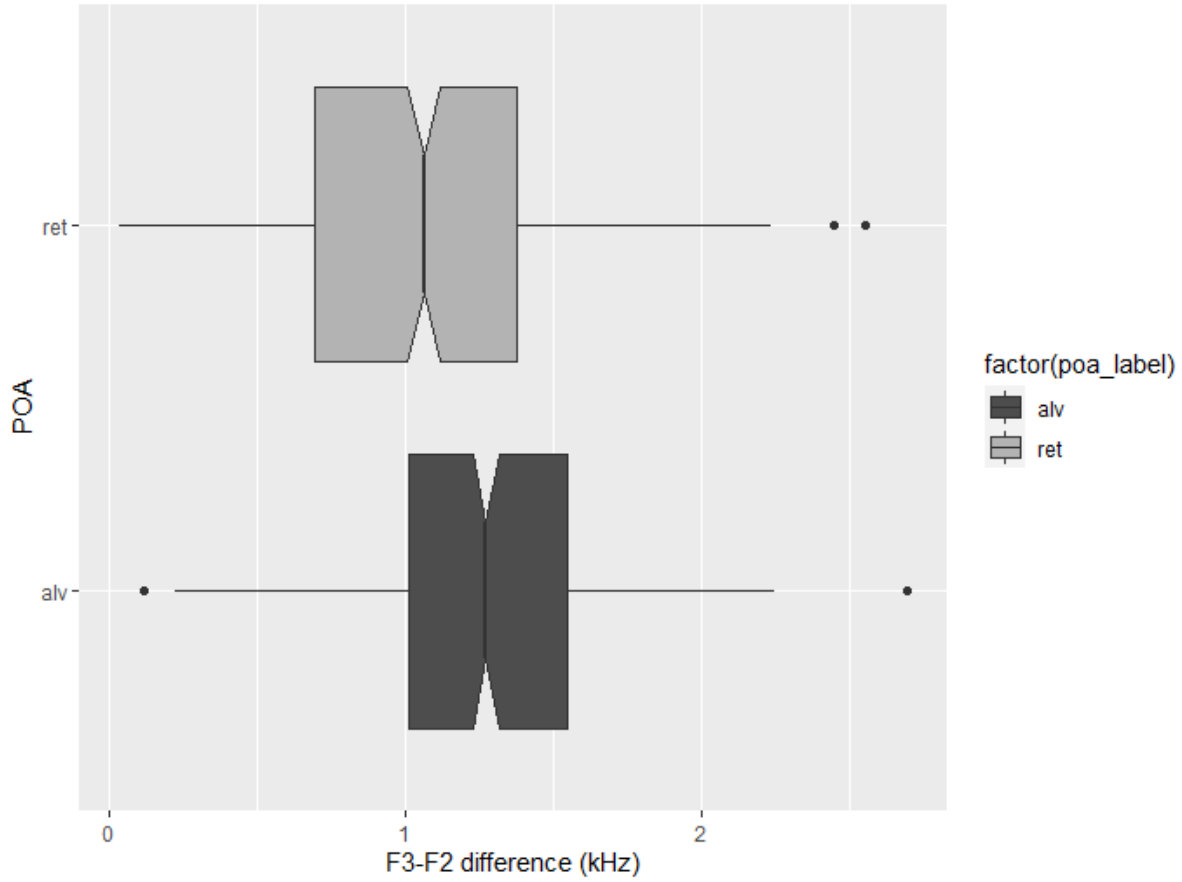
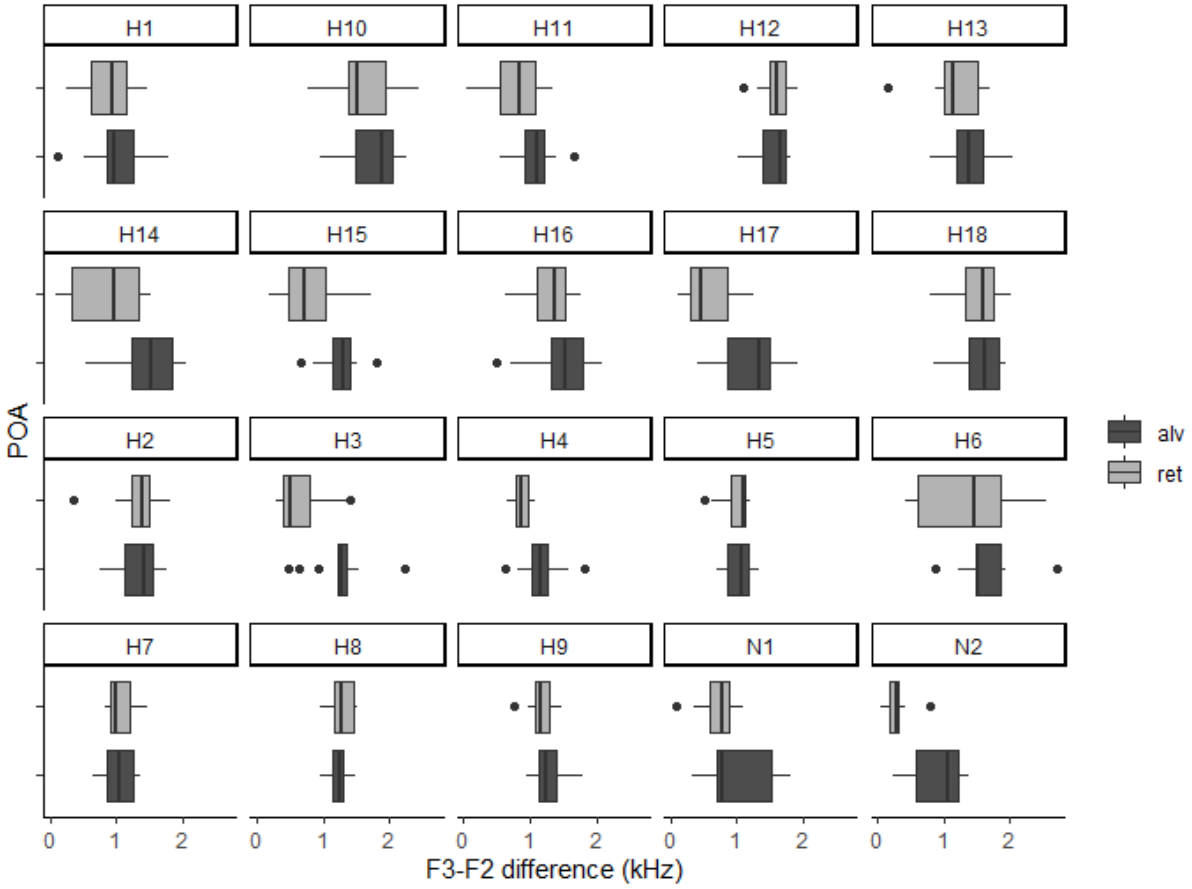


Figure 5 shows each participant's production of the alveo-dental and retroflex liquids in comparison. Native speakers had an alveo-dental productive salience of 952.13 Hz, with a range between 893.14 Hz and 1011.11 Hz. They had a retroflex productive salience of 493.45 Hz, with a range between 272.17 Hz and 714.73 Hz. Native production demonstrated steady larger F3-F2 windows for the alveo-dental productions and smaller F3-F2 windows for retroflex productions. A total of 8 heritage participants showed consistent patterning of larger F3-F2 windows in alveo-dental production and smaller F3-F2 windows in retroflex production, similar to the 2 native speakers shown in the bottom right corner of Figure 5, labelled N1 and N2. These productive salience values ranged from 204.54 Hz to 623.59 Hz. Some examples of these speakers are H3,

H14, H15, and H17. A total of 10 heritage speakers had smaller productive salience windows, showing their productions for both liquids had similar F3-F2 windows. Their productive salience values ranged from -48.59 Hz to 171.89 Hz. This means both the alveo-dental and retroflex liquids were pronounced closer to an alveo-dental liquid. Some examples of these speakers are H2, H5, H7, and H8. Some speakers showed an amount of intra-speaker variation, or perceptible differences within their own productions, such as speakers H14 and H6.

Figure 5

Alveo-dental vs. Retroflex Liquid Production by Participant

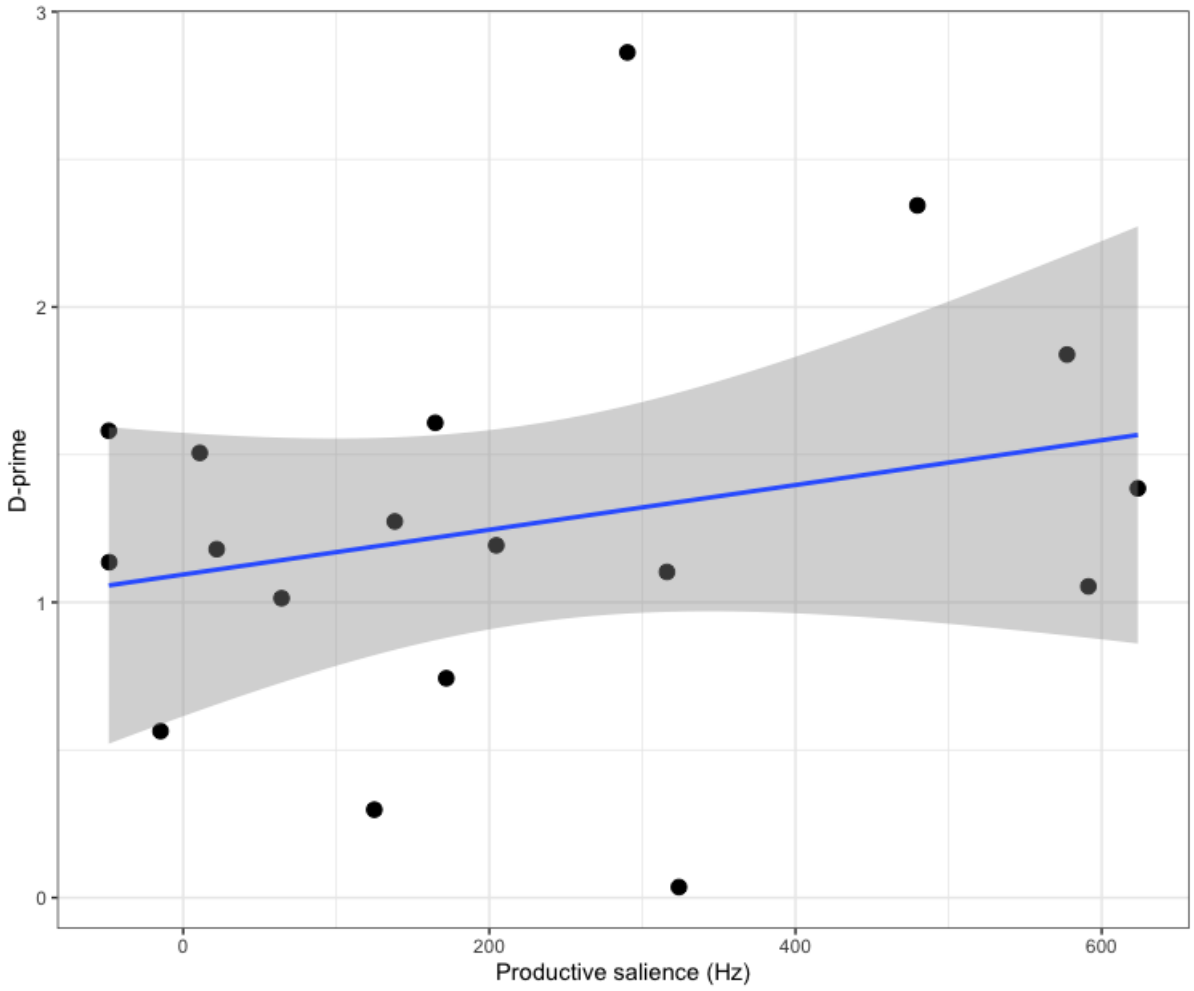


Production vs. Perception

Figure 6 shows the relationship between productive salience, or the difference between alveo-dentals and retroflexes in F3-F2 space, and D-prime (described above). The figure shows a general positive relationship between production and perception: as accuracy in perception levels increased, accuracy in production also increased. This is an overall trend as the correlation test between heritage perception and heritage production did not achieve significance (r=0.01, p > 0.5). This non-significant correlation may be due to the relatively low n in the sample. Despite this shortcoming, the data trends towards an interpretation that participants who could hear the contrast more accurately also produced the contrast with more accuracy. Figure 6 shows the relationship between accuracy in perception and production. Even those who produced the contrast successfully and efficiently have similar ranges with little difference for alveo-dental and retroflex production, showing the low perceptual salience of this heritage specific contrast.

Figure 6

Accuracy in Retroflex Perception (D-Prime) vs. Accuracy in Retroflex Production (Productive Salience (Hz))



The quantitative results drawn from the heritage language questionnaire, the AX discrimination task, and the elicited production task yield no apparent patterning in perception or production among this group of heritage Tamil speakers. Correlation tests between perception, production, and identity yielded insignificant results. This could be due to a low *n* in the sample, as well as differences among individual results in perception and production. These results show immense variation among speakers and these quantitative factors have no real effect on

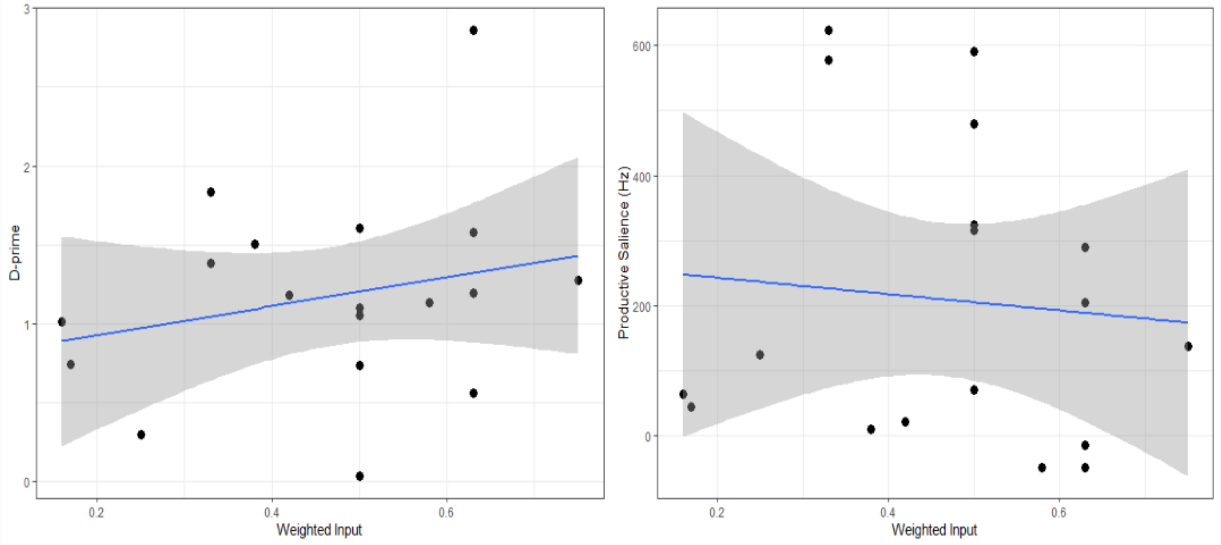
perception and production. This leads us to look further into qualitative results and their role in heritage Tamil perception and production, outlined in further detail in the next sections.

Input and Proficiency Results

General statistical analysis through correlation tests found no significant correlations or trends between language input or exposure with accuracy in perception or production, as they all yielded a p value larger than 0.5. This finding was the same for age of arrival as well, as only 12 participants were born in Canada, and the 6 others were born in India, Sri Lanka, France, the UK, amongst other places, and came between the ages of 1.5 and 12 years. So, there may have been simply too much variation in the age of arrival information to produce any sort of patterning with perception and production. Figure 7 shows how the weighted input measurements do not act as a reliable factor impacting accuracy in retroflex production. This could be attributed to the measure being too individual-specific or not sensitive enough to capture the accurate amount of Tamil language input compared to majority English input.

Figure 7

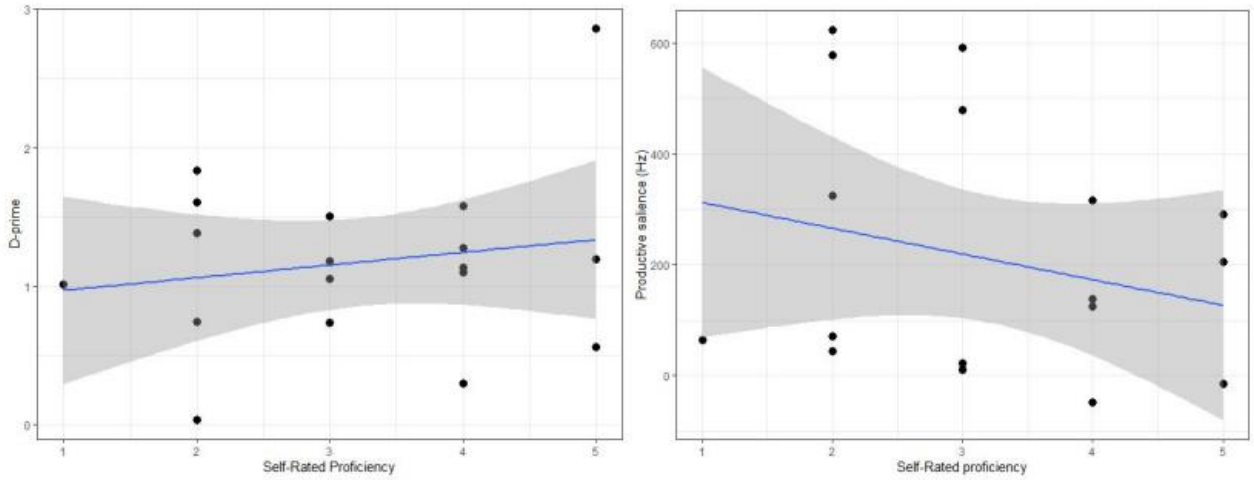
Effect of Weighted Input on Accuracy in Perception (D-Prime) and Production (Productive Salience (Hz))



There was also a non-significant ($r = 0.17$, $p > 0.5$) trend whereby increases in self-rated levels of proficiency were positively correlated with levels of accuracy in perception among heritage speakers. Figure 8 below shows the variation at each level of self-rated proficiency, meaning proficiency itself is not a reliable indicator for accuracy in retroflex perception or production. There is also variation in the data across proficiency levels, as a self-rating scale can cause speakers to rate themselves against what they personally believe to be fully fluent.

Figure 8

Accuracy in Perception (D-Prime) and Production (Productive Salience (Hz)) Across Self-Rated Proficiency Levels



Interview Results

Qualitative data retrieved from the heritage language questionnaire and sociolinguistic interviews found five main ideologies surrounding heritage language proficiency, language beliefs, and learning motivations. These five ideologies can be broadly identified as:

1. Connection to self
2. Differentiating a cultural identity (CI) and linguistic identity (LI)
3. Language as an obligation
4. "Good enough language policy" for proficiency
5. Majority Language Dominance

Many participants felt speaking Tamil acted as a connection to their personal identity, that it provided them an avenue to truly be themselves and to learn more about their personal history, as example (1) outlines below. Participants expressed a difference between their cultural

identity and their linguistic identity as a Venn diagram, in which some people had a larger overlap of each circle than others. Some participants stated that their proficiency in Tamil did not have an effect on how strong their cultural identity was, as they relied on other cultural factors to make up their cultural identity such as food, family traditions, and religion. Example (2) outlines how language makes up only 10% of their Tamil identity, and how other factors take precedence and priority in expressing their culture. Whereas example (3) expresses the importance of a linguistic identity alongside other cultural factors.

> (1) “It feels hard to choose the most important thing, because I feel like learning and speaking Tamil for me is personal and also political, of course, but I think ultimately the most important thing for me is that personal aspect.” (Participant H14)
>
> (2) “I feel like language is only 10%... so religion and then family, clothing, food is a big part of Tamil culture for sure. [Also being] family oriented, you know, spending time with them. Valu[ing] the holidays we like to celebrate.” (Participant H15)
>
> (3) “It's good to especially know your mother tongue… It's so important. As someone who's proud of where I'm from… A lot of my identity, I think it's ingrained into the language.” (Participant H13)

Regarding language as an obligation, many participants felt as if the act of learning and speaking Tamil was an obligation to older and future generations. They expressed the importance of communication with parents, grandparents, community elders, and extended family as an obligation to maintain relationships and relate to cultural and familial experiences. They also mentioned the importance of teaching children and future generations heritage Tamil as a way of passing on the language and the obligation of knowing the language. Examples (4) and (5) from participants express these two types of language learning as an obligation.

> (4) “I think what's most important is being able to communicate with my elders. Communication is important and as well, going back to India and being able to communicate with other relatives.” (Participant H10)

> (5) “A lot of the time when you are like a first or second generation immigrant, the language kind of dies a little bit with you and I think the motivation is that hopefully it doesn't and that like our kids and our future generations, would also want to continue [speaking Tamil].” (Participant H17)

Another observed ideology is best described as a “good enough language policy,” in which participants expressed that their Tamil proficiency was relatively low, but that they could communicate their general ideas to those closest to them. The motivation to further proficiency did not manifest in a way to acquire the language further through formal language courses or online resources, but rather in a way to passively maintain language acquisition that had transpired in the past. Though the general feeling among participants was that their Tamil could be ameliorated, there were no concrete steps to further acquire or develop their abilities, as shown in examples (6) to (8), outlining their personal proficiency motivations and reliance on the “good enough” language policy.

> (6) “I would say it’s okay. So, someone who I'm speaking to can understand and comprehend, but I don't think it's at a really good level. Personally, I think it's just good enough to get by.” (Participant H2)

> (7) “My Tamil skills could definitely be improved. Although I can fully understand what others are saying, it can be challenging to respond back at times. I am also limited in my vocabulary. I am decently comfortable with my current skills, but there is ample room for improvement.” (Participant H16)

> (8) “It's probably mediocre at best, but I think a lot of people would appreciate the conscious effort more than if it's actually correct or not.” (Participant H5)

One ideology that was tied to heritage language proficiency that multiple participants expressed was majority language dominance. These speakers are simultaneous bilinguals, exposed to Tamil and English from childhood, and living in a Canadian English majority language environment has consequently fostered their English to become their dominant language. This dichotomy of needing to use English everywhere and not needing Tamil often has resulted in a type of heritage language attrition among speakers. Examples (9) to (14) below

show how language dominance in Canadian English affects their heritage fluency. Meanwhile, other speakers expressed how natural it was to speak in Tamil, as examples (15) and (16) show below:

(9) “English comes as a reflex.” (Participant H1)

(10) “I am more comfortable using English, as I sometimes mess up Tamil words.” (Participant H11)

(11) “English is the default language that comes to mind.” (Participant H16)

(12) “It's easier and quicker. I can communicate more in English, as I have difficulty communicating clearly and quickly in Tamil.” (Participant H18)

(13) “I only use English to get my thoughts out more comfortably” (Participant H12)

(14) “I opt for English when I can’t think of the word in Tamil.” (Participant H8)

(15) “I only opt for English if my peers have a low fluency in Tamil, if not, I prefer speaking in Tamil.” (Participant H13)

(16) “[Speaking in Tamil] is comfortable. Speaking in English loses that personalization and that sense of self.” (Participant H12)

A total of 9 participants stated that their first language was English, as all speakers were simultaneous bilinguals, acquiring Tamil and English at the same time from birth. This means 50% of speakers equate their dominant language with their first language, even if they learned two simultaneously. Speakers also spoke of family members using less and less Tamil over time and opting to communicate in English more often, as well as extended family members abroad learning English to communicate with them. These factors contribute to language dominance in English in a heritage context, as well as heritage language attrition, and influence the way these heritage Tamil speakers choose to speak in a bilingual context.

Overall, results showed a high degree of variation in productive salience, i.e., some speakers with clear differentiation between alveolar and retroflex, and others very little.

Likewise, perceptual distance was also variable, with some participants clearly showing categorical discrimination while others have very little perceptual space for both liquids. Interestingly, a simple language exposure metric in the form of weighted relative input did not predict productive or perceptual scores. Follow-up sociolinguistic interviews, however, revealed that personal connections to the heritage language, speaking Tamil as an obligation, and having a strong cultural identity with a distinct linguistic identity were important and prominent language ideologies within this speaker population, which may contribute to accuracy in perception and production.

Cultural and Linguistic Identity

Another critical finding from the qualitative data taken from the heritage language questionnaire and sociolinguistic interviews was the ideology that cultural identity and linguistic identity were two separate entities within the organization of an individual's Tamil identity. Some participants self-identified through means of clothing, jewellery, political activism, family values, religion, or traditions more so than through heritage language proficiency and ability. Other participants attributed knowing the language to being more in tune with their Tamil culture. Examples (15) to (17) outline a general cultural identity with tangible aspects while examples (18) to (20) outline more of a linguistic identity surrounding proficiency and language as an embedded aspect of individual cultural aspect:

> (15) "[Tamil keeps me] connected to my background and culture" (Participant H16)
>
> (16) "For me, it's important to speak your language because then you lose your sense of identity and you lose where you come from, like your ancestors and all that, so you lose that connection to you like you would probably never have before if you are not able to speak, or at least understand the language." (Participant H2)

> (17) I would say having an understanding of the language would be helpful in knowing our culture, but it's not a necessity in my opinion. I think anyone who wants to immerse in our identity, can do so regardless of if you know Tamil or not. If you do understand or speak it, it kind of opens like another layer of depth to the, but it's not something that's necessary. I think to me, a big part of my identity just comes from taking part in traditions and activities and stuff, rather than knowing the language" (Participant H5)
>
> (18) "[Living in a predominantly white community], I think language was the main thing that we were able to really immerse ourselves in within the community because we would have to drive out in order to go to a temple or… to see other friends who were Indian. I think language [is] kind of like the main tie that we could have within the house to our culture and identity." (Participant H17)
>
> (19) "I think communication also goes into language and everything because you need to know the language in order to communicate. I love the fact that I can show my love and communicate and be able to talk to them. My [heritage] language is very important to me." (Participant H13)
>
> (20) "I think that it celebrates a part of me. It connects me to other people. I want to say it's something almost as primal as an instinct. It just doesn't feel right any other way and it's just kind of like natural [to me]" (Participant H4)

General findings showed that participants who valued a cultural identity stronger than a linguistic identity showed variation in their perception and production; whereas participants who valued a strong linguistic identity embedded within their cultural identity had an advantage in their retroflex production. Participants in examples (15) to (17) above had lower productive salience than participants in examples (18) to (20), meaning the latter group produced the liquid contrast with more accuracy. By attributing conscious attention and importance to a linguistic identity, this social factor might give them an advantage in producing the heritage specific retroflex contrast. However, this same pattern of advantage was not observed for their perception, which might point to language input as a helpful indicator for heritage perception, and a metric of realized linguistic identity for heritage production.

Individual Case Studies

To fully recognize the variation in these results, it is important to highlight the particular circumstances of participants. Take participant H3 for example, who is 25, was born in Delhi, India, and moved to Canada at 2 years old. H3 values a strong cultural identity through family values, religious holidays, traditional clothing, and Tamil food. H3 also values a strong linguistic identity, as "language is a large part of [their] identity", and they believe the way they understand and speak Tamil allows them to engage with themselves and their community in a meaningful way. Over time, H3 "noticed the tremendous influence that language has on the ability to share my culture with my peers", and the importance of heritage language knowledge alongside cultural knowledge. H3 receives most of their generally low Tamil input from their parents, as Tamil is spoken 38% of the time, and they have one sibling they speak to in only English. They listen to Tamil music and watch Tamil movies occasionally. Their overall self-rated Tamil proficiency is low, as they say they "tend to mix in English words" and code switch often. They also have very low literacy and writing skills, retained from heritage language schooling in childhood. Results found that H3 has high accuracy in retroflex liquid perception, scoring 78.47% with a D-prime value of 1.84, as well as retroflex liquid production, having distinct alveo-dental and retroflex productions with a productive salience level of 577.33 Hz. Even though H3 struggles with overall heritage fluency and low proficiency, their perception and production abilities are higher than their peers in this study. H3 has a strong cultural identity, a strong linguistic identity, low overall proficiency, and high perceptive and productive abilities.

Another participant, H9, is 19, was born in France, and moved to Canada at 3 years old. H9 values a strong cultural identity, as they are active in the traditional and political aspects of Tamil culture, take part in Tamil organizations, and spend time engaging in Tamil media and pop

culture. They focus on cultural aspects that are important to them, such as traditions, values, quality time, and community organization. H9 also has a strong Tamil-speaking social network with many in-group peer members who meet regularly and speak to each other in both Tamil and English. H9 has a high self-rated Tamil proficiency as they can "speak to other Tamil people with ease" and "very strongly identifies" with their Tamil identity. They have intermediate literacy skills but low writing skills, retained from heritage language schooling in childhood. Most of H9's Tamil input comes from their parents, as Tamil is spoken 43% of the time, and they speak with their two siblings in mostly English and French. Though H9 has a higher rate of Tamil output at around 40% of the time, their perception in discriminating alveo-dental liquids from retroflex liquids was quite low, accurately discriminating 54.86% of the time, with a D-prime value of 0.29. They also had low accuracy in producing this contrast, with a productive salience level of 124.87 Hz, meaning the two liquids were produced more similarly than contrastively. H9 speaks Tamil frequently with parents and peers, has a high level of overall proficiency, but they had difficulty and low accuracy in both perception and production. H9 can speak the heritage language with ease and has a very strong cultural identity, but this overall proficiency and connection to culturally specific aspects do not positively impact accuracy in their perception and production, which could be due to the "good enough language policy", as they value effective communication over accuracy in communication.

Participant H14, is 23 and was born in Canada. H14 values a strong cultural identity alongside a linguistic identity, as they practice speaking their heritage language as an act of "political resistance and resilience" but also because it "makes [them] feel more connected to [their] family". They are involved in community organization, education, and awareness. They also value linguistic knowledge and speaking their heritage language as a big portion of their

overall Tamil identity. H14 has a moderate overall self-rated proficiency in Tamil, as they can "hold simple conversations in Tamil", with literacy and writing skills retained from heritage language schooling in childhood. They get most of their Tamil input from their parents, hearing Tamil about 50% of the time, and speaks with their sibling in English. Participation in this study showed that H14 had performed moderately in contrast discrimination with 69.44% accuracy, a D-prime value of 1.054, and had high accuracy in contrast production, with a productive salience of 591.31 Hz. H14 has a strong cultural identity, a strong linguistic identity, with moderate proficiency, resulting in moderate perceptive abilities, and high accuracy in retroflex production.

Participant H2, is 19 and was born in Canada. H2 values a strong cultural identity, as they participate in Tamil organizations, community events, cultural groups, and engage in political discussions for cultural freedom, education, and awareness. They "very strongly" associate with their Tamil identity and hold strong cultural values. H2 self-rated their proficiency moderately, as they say their abilities are "just okay", and they can "hold simple conversations". They do not have any literacy or writing skills, and they did not attend any heritage language schooling. H2 has a motivation to learn these skills and has a plan in place during summer vacation to start learning how to read and write in Tamil. H2 receives most of their Tamil input from their parents, hearing Tamil around 38% of the time, and speaks to their siblings in English. H2 performed moderately well in the perception task, scoring 75.69% with a D-prime of 1.51. After performing in the production task, H2 said they "know the difference between the two L sounds but [they] cannot pronounce the difference", which is realized with a productive salience level of 10.86 Hz, meaning they were producing the same sound for each contrast pair. H2 has a strong cultural identity, with moderately good perceptive abilities, but is self aware of their repeated inaccuracy in retroflex production.

Participant H18 is 26 years old and was born in Canada. In the past year, H18 has been on a journey to reclaim their Tamil heritage language. They classify themselves as not proficient in Tamil at all, as they "only know a few words", and have only been trying to speak it more and more during the past year. In childhood, they heard a lot of Tamil input from their parents and could understand it fluently, but never got around to speaking it. Over time, familial Tamil input decreased, and H18's parents and sibling only speak to them in English. H18 did not attend heritage language schooling and has no Tamil literacy or writing skills. H18 has been active in Tamil organizations for over 5 years, taking part in community events and cultural traditions. Only recently did H18 realize the importance of a linguistic identity alongside their cultural efforts. After joining Tamil groups and learning about the importance of a heritage language did H18 get involved in revitalizing their own abilities. Their new sparked interest in learning the language led them to participate in this study as they were curious to know what heritage language research is about. H18 scored 64.6% accuracy on liquid discrimination, with a D-prime value of 1.01. In the production task, H18 had a low productive salience of 64.27 Hz. H18 has a relatively new cultural identity, a new interest in their linguistic identity, but has low perceptive and productive abilities due to this delayed commitment to their heritage language.

Most of the participants in this study valued a strong cultural identity, and a few valued a strong linguistic identity as well, but this was not the case for all participants. H6 is 25 and was born in Canada. They do not identify strongly as a Tamil Canadian, as they prefer being broadly described as Indo-Canadian. H6 has low Tamil proficiency, as they know "some basic phrases" in Tamil, but they "don't prefer speaking in Tamil". They find using English to be "easier, more universal, more applicable, and accessible". H6 does not spend time consuming Tamil media or immersing themselves in Tamil pop culture. They did not attend heritage language schooling,

and do not have Tamil literacy or orthographic skills. They are an only child and receive their Tamil input from their parents, who speak it around 38% of the time. H6 scored the lowest on the discrimination task of all participants, with a score of 50.7% on liquid discrimination, with the lowest D-prime value at 0.04. On the production task, H6 struggled with pronunciation and production of the minimal pairs and had an overall low productive salience of 138.25 Hz. This is an example of a heritage Tamil speaker with a weaker cultural identity, a weaker linguistic identity, as well as low perceptive and productive abilities due to reliance on the dominant language and a general disinterest in the heritage language.

These case studies highlight the variation found in this sample population of heritage Tamil speakers. They vary in proficiency level, sense of cultural and linguistic identity, place of origin, as well as heritage language experience, input, and schooling. This variation shows in their perceptive and productive abilities, as some speakers have high accuracy or low accuracy in both perception and production, while others have moderate levels of accuracy in either perception or production. For the most part, speakers' self identification around a Tamil identity and self awareness around their own perception and production limitations was extracted from the qualitative data.

6. Discussion

Regarding the AX discrimination task, participants had a harder time discriminating the liquid contrast ([l]-[ɭ]) with low perceptual salience than the stop contrast ([d̪]-[ɖ]) with higher perceptual salience. This could be due to the perceived dissimilarity being greater in the latter contrast than the former, which reflects recent findings by Kan & Schmid (2019). They found that heritage Cantonese speakers had more difficulty discriminating between acoustically similar tones than acoustically different tones, comparable to the present study where heritage Tamil speakers showed difficulty discriminating between the acoustically similar liquid contrasts than acoustically different stop contrasts. These parallel findings show that heritage speakers have a harder time discriminating contrasts that are acoustically more similar. It could be that heritage speakers are not performing at the native ceiling due to low perceptual salience of this heritage specific [l] and [ɭ] contrast. Where these studies diverge is in the perception of identity versus contrast pairing. The present study showed heritage Tamil speakers were less accurate in discriminating contrast pairs over identity pairs, the opposite of what Kan & Schmid (2019) found with Cantonese tones. This could be attributed to the fact that suprasegmentals are simply processed differently than consonants. Comparing these two languages and two vastly different phonological properties cannot be done equally, so different outcomes in these perception tasks might be expected.

Effects of Input and Proficiency

Results showed that input and proficiency effects did not correlate significantly with perception and production data. As Gorba (2023) stated, using simplistic measurements of language experience on their own is not a reliable metric to analyze language ability within

bilingual speakers. It is important to consider another factor that many participants lived in multilingual households where other languages like French, Telugu, and Hindi were also spoken, or they lived away from home and contacted family through regular visits, phone calls, and video calls. Meanwhile, most heritage language studies investigating input as a factor involve populations of children, where parental input is more consistent as children reside at home with their families (Saravanan, 2001, Park & Sarkar, 2007; Daskalaki et al., 2020). These diverging findings could be due to the age difference of the population studied. However, in retroflex perception, there was an overall positive trend suggesting that as the amount of language input increased, the perceptibility of the contrast also increased, indexed by an increasing D-prime. It could be that the more Tamil input these speakers here, the more attuned they are to hearing this heritage specific contrast. This effect has been documented in other heritage speaking populations as well (Kan & Schmid, 2019; Kim & Repiso Puigdelliura, 2019; Kim, 2020)

As found in Seo et al., (2022), speakers with higher productive or expressive abilities also have higher perceptive abilities. This trend was observed in the data comparing perception abilities with productive abilities. While level of overall proficiency serves as a measure for more quantifiable aspects of language knowledge, such as grammatical ability and fluency, as reported in Shea (2019) and Kim & Repiso Puigdelliura (2019), it has difficulty measuring sensitivity in perception and production of low acoustically salient heritage-specific contrasts in this context.

Assessing Cultural Identity and Linguistic Identity

Sociolinguistic research has previously found that how a person self-identifies their cultural and linguistic identity can result in intra-group variation (Tse, 2000; Bailey, 2005;

Eckert, 2008; Kim & Chao, 2009; Canagarajah 2012). The findings of the current study are parallel to a study conducted by Kim & Chao (2009), who evaluated the embedding of heritage language fluency within cultural and ethnic identity in 207 Chinese and 354 Spanish heritage language speakers. They found that Spanish speakers felt that their heritage language proficiency was a big portion of their ethnic identity, and this fluency was a descriptor of their linguistic and cultural identity, while Chinese speakers did not feel as strongly that heritage language proficiency impacted their cultural and ethnic identity. These findings reinforce the ideology that heritage language speakers are not a monolith, and that social factors like linguistic and cultural identity affect proficiency and motivation, but not all populations equally. The qualitative results from these heritage Tamil speakers show the same pattern within a sub-group of heritage speakers who share the same ethnicity, as some speakers value tangible cultural qualities while others value linguistic qualities to self-identify as a heritage Tamil speaker. There is variation on how these speakers recognize and build their cultural and linguistic identities, which in turn can affect their overall proficiency, perceptive, and productive abilities (Kim & Chao, 2009; Canagarajah, 2012; Hwa Lee, 2014).

Language Dominance and Phonological Category Transfer

Some heritage Tamil speakers performed at the native ceiling level in the perception and production tasks, while others exhibited variation with their [l]-[ɭ] contrast. As Flege and Bohn (2021) theorized, if the perceived dissimilarity is low between a specific contrast, the likelihood of a new category forming is also low. If heritage Tamil speakers who consider their first language to be English have this majority language influence on their heritage Tamil, it would consequently be difficult for them to form a retroflex liquid category to perceive and produce. Since this alveo-dental and retroflex liquid contrast has low perceptual salience, these results

could be the result of incomplete category formation or category-blocking influenced by dominant language transfer of the alveo-dental liquid, which is found in English and Tamil. Results reflect this theory, as speakers with low productive salience substituted the alveo-dental liquid in place of the retroflex liquid, resulting in low productive salience ranges between -48.59 Hz to 171.89 Hz. Since the retroflex liquid is only present in Tamil, and these speakers are hearing less and less Tamil input as they get older, it could be possible that these phonological categories were weak to begin with due to low perceived dissimilarity and are diminishing over time due to incomplete acquisition in childhood, heritage language attrition in adulthood, and majority language dominance. Research also shows that English language input from elder siblings has a significant impact on target language development in younger siblings (Sorenson Duncan & Paradis, 2020). As these speakers mainly speak to their siblings in English, and have been hearing this English input since birth, this could also affect language dominance and incomplete heritage language acquisition. Language dominance has been shown to produce variation in other heritage speakers, and it could also be the case for heritage Tamil as well (Park-Johnson, 2017; Kim & Repiso Puigdelliura, 2019; Shea, 2019; Soo & Monahan, 2021).

Flege and Bohn (2021) also highlight the importance of naturalistic, continuous, and consistent input of the target language to ensure the creation of concrete phonemic categories. This is an area where heritage language acquisition and development can differ, as this heritage Tamil population varies regarding their language experience, history of language use, and journey of heritage language development. Though all heritage Tamil speakers have heard Tamil input from birth, not all of them have heard an equal amount. Some speakers have attended Tamil classes outside the home, and some have not. Some speakers have been speaking Tamil their whole lives, some have stopped speaking Tamil after a certain age, and some have just

begun speaking more as they have grown older. Hwa Lee (2014) found that heritage Korean speakers' productions showed variation due to incomplete exposure to the heritage language, which could also be the case for these heritage Tamil speakers. Their perception and production could be affected by incomplete and decreasing Tamil input from sources that were already established as minimal or negligible sources. It could be theorized that since these speakers have no access to naturalistic, continuous, and consistent target language input, they would struggle to maintain the division between these phonemic categories and this acoustically fragile contrast from a purely psycholinguistic perspective.

Exploring Functional Load and Idiolectal Variation

Another linguistic feature that may potentially affect this data is the functional load of the alveo-dental and retroflex lateral contrast. Functional load measures a phonemic contrast on a scale of low to high, measuring the contrast's usefulness within the phonemic inventory for the sake of communication (Liu et al., 2017; Bennet et al., 2018). For example, while the /f/-/θ/ is very much part of the English phonological contrast inventory, it might be considered low in terms of functional load given the relatively small number of lexical items relying on the contrast (e.g., with-whiff, think-fink, etc.) (King, 1967). Phonological factors like low acoustic salience and low perceived dissimilarity can possibly affect the functional load of a phonemic contrast. If a phonological feature or property can be omitted without changing the semantics or pragmatics of an utterance, the feature is then deemed to have a low functional load. Tsui (2012) found that Cantonese tones with a low functional load are variably perceived and show feature merging, or similar productive features for both phonemes in the contrast, in production. Bennet et al. (2018) also found that functional load interacts with the perception of phonemic contrasts. If the liquid contrast had a universal low functional load, it would have reflected in the production data with

many speakers using alveo-dental substitution or by producing a retroflex with features similar to the alveo-dental liquid, but this was not the generally observed case, as 10 speakers has evidence of substitution or merged features, and 8 had distinct production of each liquid in the contrast.

This is where idiolectal variation might be affecting the data alongside functional load. Each heritage speaker has their own idiolect, or a dialect specific to an individual, shaped by their language experience, development, and use, as well as impacted by extralinguistic factors such as social network, language environment, and context. As heritage speakers have unique circumstances surrounding input, experience, and development, it could be probable that they have individualized dialects specific to their heritage contexts in the home and their cultural communities, resulting in some speakers using the liquid contrast with high accuracy to differentiate meaning and assigning it a high functional load within their idiolect. Others, however, can get by and get their point across without having to produce the retroflex variant, assigning a low functional load in their idiolect. Looking at the data from the perception and production tasks, it could be that functional load and idiolect play a role in how heritage Tamil speakers perceive and produce this contrast, as generally, those who had difficulty perceiving the contrast also had difficulty in accurately producing the contrast as well. These findings parallel several studies examining idiolect and variation in heritage speakers, as Özsoy & Blum (2023) found individual variation that could not be defined generally among Turkish speakers' discourse markers, and Shin (2022) found individual variation in the representation of heritage Spanish speakers' grammars. It could be possible that an idiolect can be comprised of specific phonological features as well, and that some speakers adopt a more accurate heritage-specific contrast than other speakers, and this could also be driven by characteristics of functional load, perceptual salience, and perceived dissimilarity. This could also be interacting with the "good

enough language policy" observed by Jegerski & Sekerina (2021) as well as in the data of heritage Tamil speakers' liquid contrast production, and in their qualitative data surrounding heritage proficiency. Some speakers may think it is "good enough" to use the same liquid for both phonemes, as context and semantic meaning convey the message without the need of distinguishing the liquids. While other speakers value preserving this contrast in their own speech as they believe it is useful to utilize it for the sake of communication and a properly realized linguistic identity. This also ties into motivation as a social factor, as previous research into heritage speaker motivation has found that high motivation and enthusiastic attitudes positively correlate with language abilities (Comanaru & Noels, 2009; Alarcon, 2010; Wen, 2011; Shea, 2019). Speakers who are more motivated or have more positive attitudes would have a higher standard for their "good-enough language policy" which could encompass finite contrast production and pronunciation.

Study Limitations

It is important to highlight the limitations of the current study to inform future research in the field of heritage Tamil perception and production. The first limitation would be the lack of tested and proven proficiency scores for speakers. While the language questionnaire, AX discrimination task, targeted production task, and follow-up sociolinguistic interview produced reliable data, they were constrained by time restrictions and participant recruitment for a highly specialized population. The measures that should be considered for future research would be a Tamil proficiency test to accurately capture a fluency metric other than a self-rated proficiency scale, as well as extending the follow-up sociolinguistic interview to include more in-depth questions on individual language identity and attitudes. Adding these measures to the study

design would allow for a more in-depth analysis of heritage Tamil perceptual and productive abilities as they interact with social factors like identity and motivation.

Another limitation of the current study would be its narrow scope of focusing on the retroflex and non-retroflex lateral approximant consonants. This liquid contrast carries a relatively low functional load, i.e., it may not be as frequent in the lexicon and conversational speech as the dental and retroflex stop contrast, and as a result is only a narrow slice of heritage-Tamil speakers' repertoire. Further research should include a greater number of heritage-specific phonemic contrasts found in Tamil and not in English, studied on a sociophonetic level, in order to provide a more holistic representation of heritage Tamil speakers' perceptions and productions. Of course, the sample size of the current study is small and therefore difficult to definitively draw concrete conclusions. Extending this study to a larger population will further solidify or perhaps adjust these findings. It will be important, however, to couple an enriched sample size of heritage speakers with comprehensive qualitative sociolinguistic methods in order to gain insight into the variety of heritage-Tamil speaker types.

7. Conclusion

This study explored how different social factors like language input, identity, and personal motivations impact heritage Tamil alveo-dental and retroflex liquid perception and production. Some heritage Tamil speakers fell within the range of native-like perception, but not all speakers patterned like their native counterparts in alveo-dental and retroflex liquid discrimination or production. Some heritage Tamil speakers showed instances of incomplete category formation by substituting the alveo-dental variant in retroflex positions in the elicited production task. Language input and exposure measurements were not significant factors in contribution to accuracy in perception and production, but the metric of language input showed a minimal effect on accuracy in contrast perception. Positive language beliefs, personal motivations, a concrete embedding of the heritage language within their individual culture, as well as a strong linguistic identity may serve as an indicator of accuracy in perception and production. Results show that heritage Tamil speakers can retain and maintain this heritage specific acoustically fragile contrast by valuing a strong linguistic identity and by increasing instances of Tamil input and language use through variability in types of quality input.

Ultimately, these findings are pertinent to the field of heritage language sociolinguistics and psycholinguistics, as they provide data from an understudied heritage language. They contribute to established previous findings in the field in heritage perception and production among other heritage languages. This research provides an introductory literature on heritage Tamil perception and production, as well as furthers sociophonetic research at the intersection of sociolinguistics and psycholinguistics. These findings also speak to the ideology that heritage speakers are not a homogenous group of speakers, as their language input and experience varies so immensely and may not pattern like L1 or L2 speakers. The discussion brings forth an

important contribution, that heritage speakers have immense variation, but that some speakers pattern similarly due to various social and input factors. This analysis also highlights the importance of considering qualitative data alongside quantitative data, as some quantitative statistics may not capture individual and personal social factors as accurately or consistently. The concluding findings from this thesis brings forth questions regarding the realization of cultural identity and linguistic identity within a speaker– how do they support heritage proficiency, and how can we measure these intangible factors to see how they affect heritage language perception and production?

References

Alarcon, I. (2010). Advanced Heritage Learners of Spanish: A Sociolinguistic Profile for Pedagogical Purposes. *Foreign Language Annals, 43*(2), 269–288. https://doi.org/10.1111/j.19449720.2010.01078.x

Amengual, M. (2019). Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 23*(5), 954–970. https://doi.org/10.1177/1367006917741364

Bailey, B. (2005). *The language of multiple identities among Dominican Americans*. In Scott Kiesling & Christina Bratt Paulson, eds. Intercultural Discourse and Communication, pp. 255-276. Oxford: Blackwell.

Bennett, R., Tang, K., & Ajsivinac Sian, J. (2018). Statistical and acoustic effects on the perception of stop consonants in Kaqchikel (Mayan). *Laboratory Phonology, 9*(1), 1-42. https://doi.org/10.5334/labphon.100

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.10, retrieved 2022 from http://www.praat.org/

Brown, J. R., & Bousquette, J. (2018). Heritage Languages in North America: Sociolinguistic Approaches. *Journal of Language Contact*, *11*(2), 201-207. https://doi.org/10.1163/19552629-01102002

Canagarajah, S. (2012). Styling One's Own in the Sri Lankan Tamil Diaspora: Implications for Language and Ethnicity. *Journal of Language, Identity, and Education, 11*(2), 124–135. https://doi.org/10.1080/15348458.2012.667309

Canagarajah, S. (2013). Reconstructing heritage language: resolving dilemmas in language maintenance for Sri Lankan Tamil migrants. *International Journal of the Sociology of Language, 222*, 131–155. https://doi.org/10.1515/ijsl-2013-0035

Canagarajah, S. (2019). Changing Orientations to Heritage Language: The Practice-Based Ideology of Sri Lankan Tamil Diaspora Families. *International Journal of the Sociology of Language*, 255, 9-44.

Chan, A. Y. W. (2012). Cantonese English as a Second Language Learners' Perceived Relations Between "Similar" L1 and L2 Speech Sounds: A Test of the Speech Learning Model. *The Modern Language Journal (Boulder, Colorado.), 96*(1), 1–19. https://doi.org/10.1111/j.1540-4781.2012.01291.x

Chang, C. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition, 19(4)*, 791-809. doi:10.1017/S1366728914000261

Chang, C., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America, 129*(6), 3964–3980. https://doi.org/10.1121/1.3569736

Cognola, F., Baronchelli, I., & Molinari, E. (2019) Inter- vs. Intra-Speaker Variation in Mixed Heritage Syntax: A Statistical Analysis. *Frontiers in Psychology, 10*(1528). doi: 10.3389/fpsyg.2019.01528

Comanaru, R. & Noels, K. A. (2009). Self-Determination, Motivation, and the Learning of Chinese as a Heritage Language. *Canadian Modern Language Review, 66*(1), 131–158. https://doi.org/10.3138/cmlr.66.1.131

Das, S. (2008). Between Convergence and Divergence: Reformatting Language Purism in the Montreal Tamil Diasporas. *Journal of Linguistic Anthropology, 18*(1), 1–23. https://doi.org/10.1111/j.1548-1395.2008.00001.x

Das, S. (2011). Rewriting the past and reimagining the future: The social life of a Tamil heritage language industry. *American Ethnologist, 38*(4), 774–789. https://doi.org/10.1111/j.1548-1425.2011.01336.x

Daskalaki, E., Blom, E., Chondrogianni, V., & Paradis, J. (2020). Effects of Parental Input Quality in Child Heritage Language Acquisition. *Journal of Child Language*, 47(4), 709-736.

De Houwer, A. (2007). Parental Language Input Patterns and Children's Bilingual Use. *Applied Psycholinguistics*, 28(3), 411-424.

De Los Santos, B. & Alves, U. K. (2021). Desenvolvimento fonético-fonológico bi/multilíngue e atenção – Questões teóricas abordadas no Speech Learning Model (SLM/SLM-r) e desafios futuros para a pesquisa em segunda língua (L2). *Revista X (Curitiba, Brazil), 16*(5), 1203–1230. https://doi.org/10.5380/rvx.v16i5.81523

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics*, *26*(2), 148–163. https://doi.org/10.3109/02699206.2011.595526

Flege, J. & Bohn, O. (2021). *The Revised Speech Learning Model (SLM-r*). In Second Language Speech Learning (pp. 3–83). https://doi.org/10.1017/9781108886901.002

Genesee, F., Paradis, J., & Crago, M. B. (2021). Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning. *Paul H. Brookes Pub. Co*.

Gharibi, K. & Boers, F. (2017). Influential factors in incomplete acquisition and attrition of heritage speakers' vocabulary knowledge. *Language Acquisition, 24*(1), 52–69. https://doi.org/10.1080/10489223.2016.1187613

Gorba, C. (2023). Is Full-Time Equivalent an Appropriate Measure to Assess L1 and L2 Perception of L2 Speakers with Limited L2 Experience? *Languages (Basel), 8*(1), 56–. https://doi.org/10.3390/languages8010056

Hoffman, M. & Walker, J. A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change, 22*(1), 37–67. https://doi.org/10.1017/S0954394509990238

Hwa Lee, E. (2014). *Speech Production and Perception of Heritage Speakers of Korean*. (UMI: 1563474). [Master's Thesis, UCLA]. ProQuest.

Icardo Isasa, A. (2022). Preliminary Results on the Development of the Perception of Spanish /e/ and /ei/ by Heritage Learners vs. L2 Learners of Spanish in the Classroom. *Languages (Basel), 7*(1), 7, 1-21. https://doi.org/10.3390/languages7010007

Isakson, S. (2018). The Case for Heritage ASL Instruction for Hearing Heritage Signers. *Sign Language Studies, 18*(3), 385–411. https://doi.org/10.1353/sls.2018.0008

Jegerski, J. & Sekerina, I. (2021). *The Psycholinguistics of Heritage Languages*. In The Cambridge Handbook of Heritage Languages and Linguistics (pp. 449–470). https://doi.org/10.1017/9781108766340.021

Kan, R., & Schmid, M. (2019). Development of tonal discrimination in young heritage speakers of Cantonese, *Journal of Phonetics, 73*, 40-54. https://doi.org/10.1016/j.wocn.2018.12.004.

Kasstan, J. (2017). New speakers: Challenges and opportunities for variationist sociolinguistics. *Language and Linguistic Compass 11*(8), 1-16.

Kasstan, J., Auer, A., & Salmons, J. (2018). Heritage-language speakers: Theoretical and empirical challenges on sociolinguistic attitudes and prestige. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 22*(4), 387–394. https://doi.org/10.1177/1367006918762151

Keane, E. (2004). Tamil. *Journal of the International Phonetic Association, 34*(1), 111–116. https://doi.org/10.1017/S0025100304001549

Keating, P. (2005). *D-prime (signal detection) analysis*. Phonetics: Linguistics at UCLA. http://phonetics.linguistics.ucla.edu/facilities/statistics/dprime.htm

Kim, J. Y. (2020). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition 23*, 233–250. https://doi.org/10.1017/S1366728918001220

Kim, J. & Repiso Puigdelliura, G. (2020). Deconstructing Heritage Language Dominance: Effects of Proficiency, Use, and Input on Heritage Speakers' Production of the Spanish Alveolar Tap. *Phonetica*, *77*(1), 55–80. https://doi.org/10.1159/000501188

Kim, S.Y. & Chao, R. K. (2009). Heritage Language Fluency, Ethnic Identity, and School Effort of Immigrant Chinese and Mexican Adolescents. *Cultural Diversity & Ethnic Minority Psychology, 15*(1), 27–37. https://doi.org/10.1037/a0013052

King, R. (1967). Functional Load and Sound Change. *Language (Baltimore), 43*(4), 831–852. https://doi.org/10.2307/411969

Lee-Ellis, S. (2012). *Looking into bilingualism through the heritage speaker's mind*. [Dissertation, University of Maryland]. ProQuest Dissertations Publishing.

Li, Y.A., & Matthews, S. (2022). Turning the tide: Reversing heritage language attrition and shift in a Chinese American family. *Journal of Chinese Linguistics, 50*(2), 291-332. 322. doi:10.1353/jcl.2022.0012.

Liu, L., Kong, C., & Wang, W. (2017). The Functional Load of Chinese Tones and the Tonal Evolution. *Journal of Chinese Linguistics, 51*(1), 73-100.

Lynch, A. (2003). The Relationship between Second and Heritage Language Acquisition: Notes on Research and Theory Building, *Heritage Language Journal*, *1*(1), 26-43. doi: https://doi.org/10.46538/hlj.1.1.2

Matsunaga, M., Hecht, M. L., Elek, E., & Ndiaye, K. (2010). Ethnic Identity Development and Acculturation: A Longitudinal Analysis of Mexican-Heritage Youth in the Southwest

United States. *Journal of Cross-Cultural Psychology, 41*(3), 410–427. https://doi.org/10.1177/0022022109359689

Mazzaro, N. (2022). The Effect of Language Contact on /tʃ/ Deaffrication in Spanish from US–Mexico Borderland. *Languages (Basel), 7*(2). https://doi.org/10.3390/languages7020101

Montrul, S. (2016). The Acquisition of Heritage Languages. Cambridge University Press.

Nagy, N. (2018). Linguistic attitudes and contact effects in Toronto's heritage languages: A variationist sociolinguistic investigation. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 22*(4), 429–446. https://doi.org/10.1177/1367006918762160

Nagy, N., Chociej, J., & Hoffman, M. F. (2014). Analyzing Ethnic Orientation in the quantitative sociolinguistic paradigm. *Language & Communication, 35*(1), 9–26. https://doi.org/10.1016/j.langcom.2013.11.002

Narayan, C. (2008). The acoustic–perceptual salience of nasal place contrasts. *Journal of Phonetics, 36*(1), 191–217. https://doi.org/10.1016/j.wocn.2007.10.001

Narayan, C. (2019). An acoustic perspective on 45 years of infant speech perception, Part 1: Consonants. *Language and Linguistics Compass*, *13*(10). https://doi.org/10.1111/lnc3.12352

Narayanan, S., Kaun, A., Byra, D., Ladefoged, P., & Alwan, A. (1996). *Liquids in Tamil*. Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96, 2, 797–800 vol.2. https://doi.org/10.1109/ICSLP.1996.607721

Narayanan, S., Byrd, D., & Kaun, A. (1999). Geometry, Kinematics and Acoustics of Tamil Liquid Consonants. *The Journal of the Acoustical Society of America, 106*(4), 1993–2007. https://doi.org/10.1121/1.427946

O'Grady, W., Lee, O., & Lee, J. (2011). Practical and Theoretical Issues in the Study of Heritage Language Acquisition. *Heritage Language Journal, 8*(3), 315-332. doi: https://doi.org/10.46538/hlj.8.3.2

Özsoy, O. & Blum, F. (2023). Exploring individual variation in Turkish heritage speakers' complex linguistic productions: Evidence from discourse markers. *Applied Psycholinguistics, 1*(31). https://doi.org/10.1017/S0142716423000267

Paradis, J., & Grüter, T. (2014). *Input and experience in bilingual development*. In J. Paradis & T. Grüter (Eds.), Trends in Language Acquisition Research: Input and Experience in Bilingual Development. John Benjamins Publishing Company. 1-14.

Park, S. M. & Sarkar, M. (2007). Parents' Attitudes Toward Heritage Language Maintenance for Their Children and Their Efforts to Help Their Children Maintain the Heritage Language: A Case Study of Korean-Canadian Immigrants. *Language, Culture and Curriculum, 20*(3), 223–235. https://doi.org/10.2167/lcc337.0

Park-Johnson, S. (2017). Code Mixing as a Window into Language Dominance: Evidence from Korean Heritage Speakers. *Heritage Language Journal, 14*(1), 49–69. https://doi.org/10.46538//hlj.14.1.3

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* 10.3758/s13428-018-01193-y

Polinsky, M. (2018a). *Phonetics and Phonology*. In Heritage Languages and their Speakers (Cambridge Studies in Linguistics, pp. 76-113). Cambridge: Cambridge University Press. doi:10.1017/9781107252349.004

Polinsky, M. (2018b). *How to Study Heritage Speakers: Observations on Methodologies and Approaches.* In Heritage Languages and their Speakers (Cambridge Studies in Linguistics, pp. 114-163). Cambridge: Cambridge University Press. doi:10.1017/9781107252349.004

Rallo Fabra, L. & Tyler, M. D. (2023). Predicting discrimination difficulty of Californian English vowel contrasts from L2-to-L1 categorization. Ampersand (Oxford, UK), 10, 100109. https://doi.org/10.1016/j.amper.2023.100109

Rampton, B. (1995). Language crossing and the problematicization of ethnicity and socialization. *Pragmatics 5*(4), 485-513.

Saravanan, V. (2001). The Significance of Bilingual Chinese, Malay, and Tamil Children's English Network Patterns on Community Language Use Patterns. *Early Child Development and Care*, *166*(1), 81–91.

Scontras, G., Fuchs, Z., & Polinsky, M. (2015) Heritage language and linguistic theory. *Frontiers of Psychology*. 6:1545. doi: 10.3389/fpsyg.2015.01545

Scontras, G. & Putnam, M. T. (2020). Lesser-studied Heritage Languages: An Appeal to the Dyad. *Heritage Language Journal, 17*(2), 152–155. https://doi.org/10.46538/hlj.17.2.2

Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners. *Languages*, *7*(1):6. https://doi.org/10.3390/languages7010006

Shea, C. (2019). Dominance, Proficiency, and Spanish Heritage Speakers' Production of English and Spanish Vowels. *Studies in Second Language Acquisition*, *41*(1), 123–149. https://doi.org/10.1017/S0272263117000328

Shin, N. (2022). Structured variation in child heritage speakers' grammars. *Language and Linguistics Compass, 16*(12). https://doi.org/10.1111/lnc3.12480

Soo, R. & Monahan, P. J. (2021). Language dominance affects auditory translation priming in heritage speakers. *The Journal of the Acoustical Society of America, 150*(4), A44. https://doi.org/10.1121/10.0007572

Sorenson Duncan, T. & Paradis, J. (2020). Home language environment and children's second language acquisition: the special status of input from older siblings. *Journal of Child Language, 47*(5), 982–1005. https://doi.org/10.1017/S0305000919000977

Statistics Canada. (2021). *Census of Population.* [Data visualization tool]. https://www12.statcan.gc.ca/census-recensement/index-eng.cfm

Tse, L. (2000). The Effects of Ethnic Identity Formation on Bilingual Maintenance and Development: An Analysis of Asian American Narratives. *International Journal of Bilingual Education and Bilingualism, 3*(3), 185–200. https://doi.org/10.1080/13670050008667706

Tsui, T. (2012). Tonal variation in Hong Kong Cantonese Acoustic distance & functional load. *Proceedings from the Annual Meeting of the Chicago Linguistic Society, 48*(1), 579–588.

Vijayakumar, K., Gunalan, S., & Rajeshwaran, R. (2021). Development of Minimal Pair Test in Tamil (MPT-T). *Journal of Clinical and Diagnostic Research. 15*(9). 1-4. 10.7860/JCDR/2021/46807.15357

Wen, X. (2011). Chinese Language Learning Motivation: A Comparative Study of Heritage and Non-heritage Learners. *Heritage Language Journal, 8*(3), 333–358. https://doi.org/10.46538/hlj.8.3.3

Yazawa, K., Konishi, T., Whang, J., Escudero, P., & Kondo, M. (2023). Spectral and temporal implementation of Japanese speakers' English vowel categories: A corpus-based study. *Laboratory Phonology, 24*(1). https://doi.org/10.16995/labphon.6427

Appendix A: Heritage Language Questionnaire

Part 1: Individual Questions

1. How old are you?
2. Were you born in Canada? If not, what was your age upon arrival?
2. What is your first language?
3. What is your second language?
4. What additional other languages do you speak? (Even if you aren't fluent)
5. How long have you been speaking Tamil?
7. How fluent are you in speaking Tamil?

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all topics and to strangers

8. Compared to others in your situation, do you believe you speak Tamil less or more fluently than your peers? (This is your friend group)

Less	More

9. How often do you speak Tamil and English on average? (this is in a bilingual setting)

1 – 100% English and 0% Tamil	2 – 75% English and 25% Tamil	3 – 50% English and 50% Tamil	4 – 75% Tamil and 25% English	5 – 100% Tamil and 0% English

10. Compared to other 1st or 2nd generation Tamil speakers, how well would you rate your Tamil? (This is all Tamil kids in Canada)

Worse	Better

11. Can you read Tamil?

Yes	No

12. Can you write in Tamil?

Yes	No

13. What activities do you partake in Tamil? (Temple, TSA, Tamil Sangam, Dance)
14. Did you attend Tamil School?

Yes	No

a. How old were you and for how long?

15. Have you ever taken Tamil classes as an adult?

Yes	No

16. Do you have a Tamil friend group?

Yes	No

a. How often do you hang out with them?

b. How many hours a week do you hang out with them?

b. Do you speak to each other in Tamil?

Yes	No

17. How many hours a week do you spend consuming Tamil media (movies, TV shows, music, podcasts, magazines, books)?

Part 2: Family Questions

19. Were your parents born in Canada? If not, when did they come to Canada? How old were they?

20. What languages are spoken at home?

21. What dialect of Tamil does your family speak? (Indian, Chennai, Eelam, Jaffna)

22. How fluent are your parents in Tamil? (If you live with grandparents add their scale too)

a. Mom Scale 1-5

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all topics and to strangers

b. Dad Scale 1-5

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all topics and to strangers

23. How fluent are your parents in English? (Same for grandparents)

a. Mom Scale 1-5

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all

				topics and to strangers

b. Dad Scale 1-5

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all topics and to strangers

24. What languages do your parents speak to you in?

1 – 100% English and 0% Tamil	2 – 75% English and 25% Tamil	3 – 50% English and 50% Tamil	4 – 75% Tamil and 25% English	5 – 100% Tamil and 0% English

What languages do you speak to your parents in?

1 – 100% English and 0% Tamil	2 – 75% English and 25% Tamil	3 – 50% English and 50% Tamil	4 – 75% Tamil and 25% English	5 – 100% Tamil and 0% English

25. Do you have any siblings?

Yes	No

How old are they?

26. How fluent are your siblings in Tamil?

1 – only know a couple words	2 – know some phrases	3 – can hold simple conversations	4 – can speak with other Tamil people with ease	5 – can communicate effectively in all topics and to strangers

27. What languages do your siblings speak to you in?

1 – 100% English and 0% Tamil	2 – 75% English and 25% Tamil	3 – 50% English and 50% Tamil	4 – 75% Tamil and 25% English	5 – 100% Tamil and 0% English

What languages do you speak to your siblings in?

1 – 100% English and 0% Tamil	2 – 75% English and 25% Tamil	3 – 50% English and 50% Tamil	4 – 75% Tamil and 25% English	5 – 100% Tamil and 0% English

28. Can they read Tamil?

Yes	No

29. Can they write in Tamil?

Yes	No

30. How many hours a week do they spend consuming Tamil media (movies, TV shows, music, podcasts, magazines, books)

a. Parents

b. Siblings

Part 3: Language Questions

31. Do you like speaking Tamil?

Yes	No

32. Do you wish you could speak more fluently in Tamil?

Yes	No

33. Do you speak Tamil with your family in public?

Yes	No

34. Do you prefer speaking English over Tamil? (In a bilingual setting)

Yes	No

a. If yes, why? (Easier, quicker, can communicate more in English, shy to speak Tamil)

35. How strongly do you associate your identity with being Tamil?

1 – not at all	2 – maybe I do	3 – yes I do	4 – Strongly	5 – very strongly

36. Do you identify as Tamil, Tamil-Canadian, Indo-Canadian, or Canadian? Are you proud of this identity?

Yes	No

37. Please tell me about your experience speaking Tamil.

Appendix B: AX Discrimination Non-Word List & Permutations

1.	ala	10.	aɖa	19.	aɳa
2.	aɭa	11.	iɖi	20.	ini
3.	ili	12.	iɖi	21.	iɳi
4.	iɭi	13.	uɖu	22.	unu
5.	ulu	14.	uɖu	23.	uɳu
6.	uɭu	15.	oɖo	24.	ono
7.	olo	16.	oɖo	25.	oɳo
8.	oɭo	17.	aɖa		
9.	aɖa	18.	ana		

1 Set = 2 Phones
3 Tokens per Phone = 6 Tokens
A1, A2, A3, B1, B2, B3

Contrast		Identity	
A1B1	B1A1	A1A1	B1B1
A1B2	B1A2	A1A2	B1B2
A1B3	B1A3	A1A3	B1B3
A2B1	B2A1	A2A1	B2B1
A2B2	B2A2	A2A2	B2B2
A2B3	B2A3	A2A3	B2B3
A3B1	B3A1	A3A1	B3B1
A3B2	B3A2	A3A2	B3B2
A3B3	B3A3	A3A3	B3B3

Appendix C: Sociolinguistic Interview Prompts

1. What is most important to you about speaking Tamil?
2. What motivates you to continue speaking Tamil?
3. How do you feel about your skills in Tamil as of right now? (pronunciation, grammar, conversational skills, vocabulary, etc.)
4. Are you currently taking steps to acquire it further? Are you comfortable with your skills now?
5. Can you explain the connection between culture and language in a way that aligns with your identity?
6. If you were to make a pie chart of all the parts of your Tamil identity (language, clothing, food, family values, traditions, jewellery, etc.), how BIG would the language chunk of this pie chart be? You can use fractions or percentages.