

JournalLing is based in Montreal, which is situated on the traditional unceded territory of the Kanien'kehá:ka, a place which has long served as a site of meeting and exchange amongst nations. JournalLing recognizes and respects the Kanien'kehá:ka as the traditional custodians of the lands and waters on which our journal operates.

The 2023-2024 JournalLing Team

Acting Editor-in-Chief

Jackson Corfield

Layout Design

Jackson Corfield

Graphic Design

Natalia Ventura Carrero

Editors

Chris Samojedny

Isabella Santiago

Alli McFarlane

Allegra Mammoli

Chiara Spino

Anneka Delilkan-Russell

Frances MacInnis

Table of Contents

Foreword	4
Hawai'ian and Te Reo Māori: School as Hearth in Language Restoration	5
Natalia Feu	
How Should We Formulate Quantity?	16
William Zumchak	
The Computation of Scalar Implicatures: Pragmatic, Lexical or Grammatical?	34
Jordan Webster	
The Construction and Motivations for Immobile Verbs in German	44
Simone Brown & Rachel Mattie	
Effects of Community Involvement and Age on Nativization: A Sociolinguistic Analysis o	f
Jewish Communities	56
Ella Brown	
Measuring the Voice Onset Time of Navajo and Tlingit Stops	72
Simone Brown	
The Gay Way to Say: A Review of Sociolinguistic Research on Gay Speech	92
Kaikoura Gutteridge	
Sounding Dyke-ish: Demystifying the Conventions of Lesbian Speech1	01
Natalia Ventura Carrero	

Foreword

Language is one of the most mysterious, personal, and fascinating aspects of human interaction. Linguistics is a relatively young field, and we have only just begun to uncover the mystery that is natural language. Equally special as language itself are those who study it, the wonderful Linguistics undergraduate students here at McGill University. The novelty of this field of study means that there are relatively few opportunities for students to have their hard work published, and *JournalLing* is an opportunity to do just this. Within these pages are eight articles written by hard-working, passionate undergraduate students who love everything about the fascinating phenomenon that is human language and deserve to have their work seen by all those who share this same passion for language as we do.

Over the past year, this journal has become something of a passion project for me. I came into the role of VP Journals and Editor-in-Chief of this edition in May of last year (2024), when at a council meeting for the Linguistics Undergraduate Association at McGill (LingUA) it was brought up that *JournalLing* was at risk of going defunct. Eager to get more involved in council, I volunteered to take on the project with a desire to prove myself and little idea of how to manage an academic journal. Since that moment, *JournalLing* has become so much more to me than a desire to prove myself. I have learned so much from *JournalLing*, and I am tremendously proud of what has become of it. However, I would not have been able to do it without the wonderful *JournalLing* team who helped and learned with me throughout the process. I would like to extend an extra special thank you to Alli McFarlane and Sophia Flaim, the Editors-in-Chief of the 2024-2025 edition of *JournalLing* releasing concurrently with the delayed release of this 2023-2024 issue, as well as the rest of the 2024-2025 team. I truly would not have been able to do it without you. I would also like to thank the authors and editors of this edition of *JournalLing* for helping me along in preparing your hard work for release, and for your patience throughout the process.

Finally, I would like to give a special thank you to Robin Desmeules, Jennifer Innes, and Deepak Chauhan at McGill Libraries for guiding me through the process of onboarding onto OJS, providing us with a website, archiving past editions, and helping us establish the policies that will allow *JournalLing* to thrive for years to come. Once again, I would not have been able to do it without you, and I am deeply grateful for your guidance. It is truly fantastic to know that the work and care that I have given to *JournalLing* over the past year will persist in the years after I am gone, and I owe that to you.

I hope you enjoy the knowledge within these pages and find it as special as I do.

Jackson Corfield Acting Editor-in-Chief March 2025 Hawai'ian and Te Reo Māori: School as Hearth in Language Restoration © 2024 by Natalia Feu is licensed

under <u>CC BY-ND 4.0</u>

Hawai'ian and Te Reo Māori: School as Hearth in Language Restoration

Natalia Feu

Department of Linguistics, McGill University LING 320: Sociolinguistics 1 Professor Charles Boberg

Abstract

In the broader global moment of Indigenous political and cultural resurgence, the revitalization and reclamation of traditional languages has emerged as an important facet of these decolonial efforts. While there are many ongoing language restoration projects around the world, Māori and Hawai'ian feature prominently in this field as two success stories of seemingly commensurate standing. However, closer scrutiny reveals not only nuances in the language restoration targets of each, but differing success rates for those respective goals. Despite developing parallel to each other, the Hawai'ian language movement has had greater success in its language revival efforts due to its prioritization of schools as centers for strong culturally diffused language networks, while the competing biculturalist and neotraditionalist ideological frameworks of the Māori language movement are paired in such a way that undermines the effectiveness of its linguistic revitalization efforts.

1 Introduction

The restoration of Indigenous languages has been a topic of significant international interest in the last half century as advocates push to decolonize the social and political landscapes in which they reside. It often follows a broader renaissance movement to reclaim Indigenous political rights and cultural heritage. Māori and Hawaiian- often cited in the same breath- are widely regarded as success stories in Indigenous language restoration, though for slightly different motives (Pine and Turin 2017, 7). Māori, the Indigenous language of Aotearoa/New Zealand, is seen as a successful case of language revitalization, the robust expansion in the domains and the number of speakers of a language in decline. Meanwhile, Hawai'ian, the Indigenous language of Hawai'i, is viewed as an almost miraculous example of language revival, the regaining of native speakers after all domains of use were lost and intergenerational transmission of language was disrupted.

2 Discussion

Situated in similar contexts, located in Polynesian island nations dominated by an Englishspeaking majority, Hawai'ian and Māori are fitting cases to contrast. Fittingly, the Hawai'ian language movement based their language planning on a model for preschool "language nests" first conceptualized and realized in Aotearoa/New Zealand. The movements, both initiated in the early 1980s, essentially developed in parallel to each other (Kawai'ae'a, Houseman, and Alencastre 2007). Despite these shared features and widespread acclamations of success, the instances of language planning have produced drastically different outcomes and varying levels of effectiveness in their respective goals. Due to differing prioritizations in school-centered language planning strategies, Hawai'ian has been more overtly successful in its effort of language revival, whereas Māori's vision of language revitalization, while not an explicit failure, has fallen short of expectations.

Language shift is a phenomenon that occurs when the domains in which a certain language is utilized start to change or are eliminated in favor of another, as a result of contact with that external language. It can occur due to a wide variety of factors, yet it is an all-too-prevalent phenomenon, particularly in colonial contexts in which the traditional language of a community is systematically marginalized by government policies of language assimilation. Maori and Hawai'ian are only two of the thousands of languages whose vitality has been harmed by these policies (Horowitz 2021). Fishman, deemed the preeminent authority in the field of language shift reversal, offers a basic framework for addressing this issue in his 1990 piece, "What is Reversing Language Shift (RLS) and how can it succeed?" The article delineates an 8-point Graded Intergenerational Disruption Scale (GIDS) for identifying the extent of language shift and thus stages for its reversal. He stresses the importance of stage 6, family-, neighborhood-, and community reinforcement, as foundational for the success of the entire enterprise and argues: "If this stage is not satisfied, all else can amount to little more than biding time" (Fishman 1990, 21). He warns against the temptation to rely too heavily on the educational system for reversing language shift, which he considers a policy often over-indulged. If this education does not meaningfully strengthen stage 6 language bonds, which he believes can only occur if the education is accompanied by higher government-administered cultural autonomy, then, he asserts, "there is absolutely no reason to assume that schooling...is either a guarantee of or even a prop for successful RLS" (Fishman 1990, 23).

Fishman's RLS theory is revisited in his 2001 book *Can Threatened Languages be Saved?*, a follow-up to his seminal 1991 compilation of work on reversing language shift. The former revisits case studies around the globe in detail—Māori included—to see where they fall on his GIDS in comparison to

a decade earlier. The chapter on Māori, by Nena Benton and Richard Benton, analyzes the discrepancy between the superficial success of Māori revitalization expressed in the sudden increase in Māori competency on state-wide census data, and the on-the-ground reality. The 8% jump in conversational ability in Māori from 1995 to 1996, most notably in younger speakers, owes credit to the educational institutions that, for many children were "the only places they ever hear[d] or sp[oke] Māori" (Fishman 2001, 423). However, the authors go on to state that due primarily to thinly spread resources, these institutions, particularly Māori-immersion schools, often lack "teachers who have sufficient knowledge of Māori to teach effectively through the language" (Fishman 2001, 436). If these centers act as the sole language input for children, then the quality of that input will either make or break children's competence in the language. The census data, in its phrasing of conversational ability rather than fluency, masks the true figure of speakers proficient enough to incorporate the language into their home lives. The staffing issue in immersion schools hinders children's ability to adequately acquire the language and decreases the likelihood that they will pass the language on to their own children.

"Pu'a I ka'Olelo, Ola ka'Ohana: Three Generations of Hawaiian Language Revitalization," coauthored by Kawai'ae'a, Houseman, and Alencastre, contains testimony from the pioneers of the Hawai'ian language revival campaign, including the first new generation of L1 speakers. It offers an intimate perspective on the development of this movement, the status of the language prior to this, and the factors needed to ensure its success. The article makes clear that by the time of its revival, Hawai'ian was experiencing the most advanced stages of language shift. By the time Aha Punana Leo, a non-profit dedicated to language revival, was established in 1983, knowledge of the language was no longer within living memory for any generation of Hawai'ians, apart from the isolated speakers of Ni'ihau island (Kawai'ae'a, Houseman, and Alencastre 2007). With no fluent parents, or even grandparents, to aid in the transmission of the language, these determined families relied entirely on the learning environment provided by their children's Hawai'ian-immersion schools for language acquisition. The core tenets of this program involved an emphasis on parent engagement, providing high-quality culturally based education, and a commitment to "speak Hawaiian everywhere and at all times" (214). As preparation for this ambitious project, native speakers and teachers lived together at a site for 2-3 weeks during the summer, functioning to strengthen social networks, develop curriculum, engage in cultural activities, and converse with native speakers.

The result of such steadfast dedication to language revival is explored in Brenzinger and Heinrich's (2013) "The return of Hawaiian: Language networks of the revival movement," in which they discuss how the creation of strong dense Hawai'ian language networks built around the educational system were the key to the revival and maintenance of the language. The greatest challenge for Hawai'ian

7

language revival was two-fold: it had to establish a way to achieve native fluency in children without the family-, neighborhood-, community nexus, since this process of intergenerational transmission was unavailable for use, as well as regain domains for the language that had been lost completely. The authors demonstrate how close-knit language networks act as a solution to both issues by inculcating its members in an all-encompassing environment. Hawai'ian serves as the primary language for all domains, rather than pursuing an attempt to sequentially recuperate one domain after another. The bonds developed in this community of practice help fill the language transmission role normally carried out by the home and family and, by putting children in close contact with the language in more than one domain, the language is further reinforced. The article further discusses that these language networks are effective precisely because they do not confine themselves to a particular ethnicity or nationality. Rather, membership is extended based on language attitude and competence.

Cowell (2012) delves deeper into how these issues of identity play into the success of language revitalization movements in his paper, "The Hawaiian model of language revitalization: problems of extension to mainland native America," in which he argues that the "dispersed cultural" nature of Hawai'ian identity, rather than a strictly ethnic one, is what has allowed the language revival movement to gain such public support and, especially, attracted so many young participants. He references the collective memory, still shared by Hawai'ians today, of "a literate, self-governing" nation which carried out its functions almost entirely through the Hawai'ian language up until the end of the 19th century, when the monarchy was overthrown, and a subsequent policy of English-only education was implemented (178). This is a shared past that crosses ethnic lines and invests the entire population of the state in the recuperation of such a rich history, allowing a "local" Hawai'ian identity and its associated cultural and linguistic qualities to become salient and take on prestige over external identifiers. Cowell contrasts this notably with the recent trend towards linking language and ethnicity occurring in the Māori context—in essence racializing such language learning— as a possible explanation for the latter's underperformance regarding language revitalization.

Tracing the trajectory of Māori language planning grants crucial insights into how its development from one ideological policy framework to another has influenced the efforts of language revitalization. Albury (2015) analyzes Aotearoa/New Zealand's shift from a biculturalist policy to a more neotraditionalist approach in his paper, "Your language or ours? Inclusion and exclusion of non-indigenous majorities in Māori and Sámi language revitalization policy." When language revitalization efforts began in earnest in the 1980s, the government adopted a biculturalist policy agenda that saw the Māori language as part of a shared cultural heritage for all citizens, Indigenous and nonindigenous alike. Lately, however, there's been a shift towards viewing Indigenous language, and thus language

revitalization efforts, as reserved for Indigenous peoples, a neotraditionalist approach that aims to return the language to its people "as a matter of post-colonial healing" (Albury, 2015). The biculturalist agenda has rendered the Māori language incredibly salient in the broader social landscape, and elevated its status on a structural level, fostering the conception that the language is thriving. This has generated a sense of complacency about the language among the public whereby "Māori and non-Māori appear generally positive about language revitalization but do not act on this through language acquisition" (Albury 2015). This complacency is reinforced, particularly for non-Māori, by a neotraditionalist attitude that has shifted the responsibility for actually acquiring the language onto ethnic Māori, further reducing the demographic for which it can become a language for everyday use.

The language revitalization and language revival efforts for Maori and Hawai'ian, respectively, must be situated in their proper contexts if their impact is to be adequately understood. By the time genuine attempts at reversing language shift were made in both Aotearoa/New Zealand and Hawai'i, intergenerational transmission was on its last legs: fewer than 100 children were fluent Māori speakers, whereas for Hawai'ian, this figure had dropped to less than 50 (Fishman 2001, 425; Kawai'ae'a, Houseman, and Alencastre 2007, 183). However, Maori was still very much a living language for the older demographic, with 18% of the geographically dispersed Maori population-mostly 55 and olderfluent in the language (Hardman 2018, 17). Thus, the Māori language of the early 1980s, with an active and extended elderly population, can be associated with stage 7 of Fishman's language shift GIDS. This was not the case for Hawai'ian, whose native speakers were restricted to the older segment of the roughly 200 inhabitants of the privately-owned Ni'ihau island (State of Hawaii 2011, 5). This geographic isolation meant that fluent speakers weren't widely available, so that Hawai'ian in the 1980s appeared to correspond more to the 8th and final stage of language shift. Yet Hawai'ian has now resurged to become the 5th most spoken language in the home—besides English—comprising 5.7% of the population of Hawai'i (Detailed Languages Spoken at Home in the State of Hawaii 2016). This ranking is higher among youth, for which Hawai'ian, along with Japanese, "is the non-English language most commonly spoken" (The Surprising Revival of the Hawai'ian Language 2019). In Aotearoa/New Zealand, Māori is the most spoken language besides English, though its speakers make up a smaller percentage overall, with just 3.2% of the country's entire population (Top 25 Languages in New Zealand 2022).

In an identical stretch of time, Hawai'ian has seemingly made greater strides than Māori, though it started at a sharper disadvantage. The root of this success can be found in the Hawai'ian education program, whose popularity and reach has only continued to grow in recent years. Children can now start their language immersion path from as young as six weeks up until they graduate their doctorate program. Rather than hindering potential success later in life, as many parents feared would occur if they veered from mainstream English schooling, Hawai'ian immersion education has enhanced academic performance in its students. Graduation and college attendance rates are significantly above the state average, and at the Nāwahī immersion school, students who make up "less than 2% of the Hilo High School Senior class [account] for 16% of its summa cum laude graduates" (Benzinger and Heinrich 2013). As Cowell (2012) mentions, the success of this model could be informed partly by the precedence set by the highly literate Hawai'ian language education during the monarchy still embedded in the cultural memory of Hawai'ians. The Māori people of Aotearoa/New Zealand, though highly literate in the 19th century as well, have never had this history of self-government or robust literary work associated with their language, leaving people with less faith in the potential of the Māori language as a medium for education (Hardman 2018). This lack of faith is evident in the declining role of Māori immersion programs in children's education. After a peak in 1993, enrollment rates in the very preschool immersion models that inspired the Hawai'ian movement fell from almost 50% of Māori children to less than a quarter, and this was coupled by a drop in the number of centers, from over 800 "language nests" at their peak to less than 300 in 2015 (Albury 2015; Hardman 2018, 23).

As Indigenous languages in island states with low linguistic diversity, Maori and Hawai'ian are particularly primed for efforts at language restoration since funding and resources can be directed at one language revitalization effort. This is in opposition to other Indigenous contexts where many threatened languages are frequently vying for resource allocation. Despite this, Māori-language educational efforts have been continuously plagued by scarce resources at a far greater level than their Hawai'ian counterpart. As Nena and Richard Benton note, Māori immersion primary schools are often "ill-equipped, lacking stable staffing," and "unable to recruit trained teachers" (Fishman 2001, 436). This ensures that Māori immersion schools lag behind English-medium schools in terms of quality of education, disincentivizing enrollment. Though immersion schools are better able to impart language competency over bilingual or Māori-as-a foreign-language models, when that is the only other option, many Māori families will instead enroll their children in English-medium schools. For the children who do attend these immersion schools, early language acquisition does seem to occur successfully, however, "many first language speakers have not been able to maintain proficiency in Maori into adulthood" (Lane 2020, 351). This drop-off in proficiency cuts off the chance to reestablish a self-sustaining cycle of intergenerational transmission, making language revitalization efforts entirely reliant on an already pressured educational system.

The biculturalist policy strategy adopted by the government may also be to blame for the strain on human resources in schools. The push to create a bilingual nation and increase all citizens' exposure to Māori has led the government to explicitly voice plans for "more than 20 percent of the country's

10

population to speak basic Māori by 2040," by pledging to "provide Māori lessons in all New Zealand schools by 2025" (Graham McLay 2018). In a country already experiencing a shortage of Māori speaking educators, expanding the scope of the language revitalization project prematurely will only undermine the effectiveness of existing efforts. The government is essentially advancing a "quantity over quality" approach, expressed in its vision of "basic" Māori proficiency as opposed to a smaller proportion of truly fluent speakers. This is an issue as it doesn't create a core speaker base for which the language can be comfortably used in daily life, and more importantly, it prioritizes outsiders who don't have a stake in the language. Fishman (1990) affirms that outside support through language learning is often "situational, temporary, idiosyncratic and even reversible," and not conducive to intergenerational transmission as the threatened language is not normally adopted as the language of the home (465).

The core motive of the biculturalist approach in Aotearoa/New Zealand, in its attempt to widen the Māori-speaking demographic, however effective this has been, seeks to address the very real concern of bringing the language back into the public domain. Several sources have pointed to the importance of generating a critical mass of speakers for the language to hold its own in a social fabric dominated by English. Bauer (2008) speaks of a "dilution" effect in which, if Māori speakers are not in enough concentration in a community, their odds of being able to use the language in daily conversation, regardless of their knowledge of it, are greatly diminished. She argues that "for a language to survive, what matters is not who *can* speak it, but who *does* speak it" (63). This tipping point, to avoid codeswitching to the dominant language, is 80%, Hardman (2018) finds. That is: for the status of the language to be considered stable, 80% of the community must be able to speak Māori. It is here that the geographical diffusion of the ethnic Māori population throughout Aotearoa/New Zealand, especially in urban centers filled with non-Māori, presents a challenge for creating a cohesive community of Māori speakers. The neotraditionalist attitudes in the country prevent a serious uptake in non-Māori language learners, making this figure of 80% difficult to achieve for interethnic communities.

The Hawai'ian language revival, in contrast, has been able to overcome this issue by deemphasizing an ethnic association with language. Rather than centering communities on ethnic kinship bonds as the Māori do in Aotearoa/New Zealand, communities are established around the Hawai'ian language itself. Hawai'ian speakers don't have to worry about geographical distribution because community can be created anywhere there is interest in the language, indiscriminate of ethnic boundaries. These small but dense language networks identified by Brenzinger and Heinrich evade the effects of diglossia or code-switching by maintaining a Hawai'ian-only environment in which Hawai'ian is the socially acceptable language and members don't feel pressure to default to English. In this social space, non-speakers of Hawai'ian are "not linguistically accommodated," and in fact the language network is quite exceptional in its implementation of "strict language behavioral norms" (Brenzinger and Heinrich 2013). These norms are enforced through Hawai'ian educational institutions, which require parents to take Hawai'ian language courses along with their children—to promote the language at home—in order to participate in the program. This is one strategy that uses the educational system to directly target language acquisition in the family sphere, with quite effective results. Idiomatic expressions and speech styles, which deviate from the standard Hawai'ian they have been taught, demonstrate that the new native speakers of Hawai'ian have taken command of the language and are already leading it forward towards change, the unequivocal sign of a living language (Brenzinger and Heinrich 2013). This contradicts Fishman's (1990) assertion that educational institutions are incapable of replacing the role of the family and community in reversing language shift. When done right, educational programs can be vital strongholds of language revitalization.

3 Conclusion

Language policies that target the family and community, the cornerstone of language survival, are difficult to implement, which is why most strategies attempt to indirectly influence this sphere by beginning at the level of the school. Hawai'ian immersion schools avoid the pitfalls of other revitalization efforts centered around schooling by transcending their role as educational institutions to become pillars of community. They achieve this by building strong social networks around the language that impel parent engagement and incentivize membership though high quality educational standards. Māori language revitalization strategies, however, due to the waning influence of their language immersion programs, have likely only slowed down the decline of the language. Emphasis on a bicultural language strategy has depleted educational resources and imbued the movement with a sense of complacency while neotraditionalist attitudes have prioritized passive comprehension over fluency among the majority population which is insufficient for intergenerational transmission. This is not to say that the Māori language revitalization cannot be successful through these policies, but rather, a better balance between the two ideological approaches must be achieved if the movement wishes to maintain viable Māori language communities long-term.

The entire field dedicated to the study of reversing language shift is relatively new, leaving many potential avenues for future research. None of the sources I referenced addressed the presence of Hawai'ian pidgin in relation to language revival efforts in Hawai'i. Future research could analyze the existence of the pidgin-Standard English continuum in Hawai'i compared to its absence in Aotearoa/New Zealand to see if and how this affects language restoration efforts. Whether the Hawai'ian language

12

revival effort carries its momentum forward remains to be seen, and follow-up research on the families that first started this process a generation or two from now and whether they have maintained intergenerational language transmission would be of significant interest. In the absence of strong ethnic ties to the language, it would be intriguing to investigate factors in the Hawai'ian revival movement that influence who decides to pass on the language to their children. Undoubtedly, language restoration is an ongoing process, so it would be disingenuous to call either movement a success or a failure, but rather, the Hawai'ian language revival is succeeding whereas the Māori language revitalization has yet to succeed.

References

Albury, Nathan John. 2015. "Your language or ours? Inclusion and exclusion of

nonindigenous majorities in Māori and Sámi language revitalization policy." *Current Issues in Language Planning* 16.3: 315-334.

Bauer, Winifred. 2008. "Is the health of te reo Maori improving?." Reo, Te 51: 33-73.

Brenzinger, Matthias, and Heinrich, Patrick. 2013. "The return of Hawaiian: Language

networks of the revival movement." Current Issues in Language Planning 14.2: 300-316.

Cowell, Andrew. 2012. "The Hawaiian model of language revitalization: problems of

extension to mainland native America." *International Journal of the Sociology of Language*: 167-193.

Fishman, Joshua A. 1990. "What is reversing language shift (RLS) and how can it

succeed?", Journal of Multilingual & Multicultural Development, 11:1-2, 5-36, DOI: 10.1080/01434632.1990.9994399.

Fishman, Joshua A. 2001. Can Threatened Languages be Saved?, Bristol, Blue Ridge

Summit: Multilingual Matters, https://doi.org/10.21832/9781853597060

Graham-McLay, Charlotte. 2018. "Maori Language, Once Shunned, Is Having a

Renaissance in New Zealand." *The New York Times*. https://www.nytimes.com/2018/09/16/world/asia/new-zealand-maorilanguage.html.

Hardman, Anne. 2018. "Literature Review: Perceptions of the Health of the Māori Language."

Ministry of Māori Development.

- Horowitz, Teddy. 2021. "Revitalizing Endangered Languages." *The International Affairs Review*. <u>https://www.iargwu.org/blog/vsba8c5mqrhvufzl4gjfmqz39e20x0.</u>
- Kawai'ae'a, Keiki KC, Alohalani Kaluhiokalani Housman, and Makalapua Alencastre. 2007. "Pu'a i

ka'Olelo, Ola ka'Ohana: Three Generations of Hawaiian Language Revitalization." *Online Submission* 4, no. 1: 183-237.

Pine, Aidan, and Turin, Mark. 2017. "Language revitalization." Oxford Research Encyclopedia of Linguistics.

2019. "The Surprising Revival of the Hawai'ian Language." The Economist.

https://www.economist.com/united-states/2019/02/21/the-surprising-revivalof-the-hawaiian-language.

2022. "Top 25 Languages in New Zealand." Ministry for Ethnic Communities.

https://www.ethniccommunities.govt.nz/resources-2/our-languages-o-tatoureo/new-registry-page/.

State of Hawaii. 2011. "Island Population and Housing Units, State of Hawaii: 2010." Department of

Business, Economic Development & Tourism Research and Economic Analysis Division. https://files.hawaii.gov/dbedt/census/Census 2010/PL94171/Island Report Final.pdf.

US Census Bureau. 2016. "Detailed Languages Spoken at Home in the State of Hawaii."

https://files.hawaii.gov/dbedt/census/acs/Report/Detailed Language March20 16.pdf.

How Should We Formulate Quantity?

William Zumchak

Department of Linguistics, McGill University LING 489: Independent Study 2 Professor Bernhard Schwarz

Abstract

Neo-Gricean accounts of how exhaustivity is derived typically rely on 3 maxims: Quality and Relevance, which form the "base conditions" and set the standards for what *can* be said, and Quantity, which uses similar conditions to those found in the first two maxims to determine what *should* be said. Typical formulations of Quantity make direct reference to the base conditions, allowing alterations to either Quality or Relevance to be reflected in Quantity, as well. However, some formulations forgo this reference in favor of restating the base conditions directly in Quantity. We examine a novel scenario that alters the conditions of Quality alone, and, by looking at the exhaustivity inferences generated in this scenario, find that Quantity does indeed reference, and not restate, the base conditions. In order to determine the impact of this finding, we examine the case of Westera's (2022) *Attentional Pragmatics*, which uses a non-referential, "independent" Quantity. Ultimately, we find that if one assumes a referential, "dependent" Quantity, as would be suggested by our findings, the proposal is unable to generate the predictions necessary for one of its main conclusions.

1 Introduction

The neo-Gricean system of deriving exhaustivity inferences typically relies on three conversational maxims: Quality, Relevance, and Quantity. In a normal, cooperative conversation, it is assumed that all three of these maxims are being followed by a speaker, allowing one to draw inferences about what the speaker means to say beyond the semantic content of a sentence. The maxims are normally phrased in reference to an "assertion," a statement that has a direct informational content. The term "direct" is important here, as it distinguishes statements like "he plays a sport" from questions like "which sport does he play?"- although the implication of the question is the same as the content of the statement, only the statement directly "asserts" that information and therefore can be called an assertion.

The first two of these maxims, Quality and Relevance, provide the "base conditions" necessary for something to be meet the minimum threshold for acceptability in a standard conversation (Katzir, 2007). Typically, they are formulated as follows, for a given proposition *p* and a speaker *s*:

(1) The Base Conditions

a. Quality: $Qual(p, s) \leftrightarrow bel_s(p)$ b. Relevance: $Rel(p, s) \leftrightarrow bel_s(rel(p))^1$

One could easily imagine why these are considered base conditions for a cooperative conversation. If speakers are not required to obey Quality, then nothing they say can be trusted, and if speakers could ignore relevance, then a conversation could never be moved in a satisfactory direction.

However, whereas the base conditions provide a standard for what should not be asserted, the third maxim, Quantity, does the opposite, instead describing what must be asserted. In further contrast to the base conditions, Quantity takes the form of a conditional, stating that everything that fulfils the antecedent must be included in the assertion. In other words, whereas the base conditions give us the bare minimum requirements for an assertion to be acceptable, Quantity gives us the gold standard, moving us past what a cooperative speaker *could* say to what they *should* say.

The formulation of this antecedent raises an interesting question. In essence, Quantity states that "If you consider something true and relevant, then include it in the assertion." One might quickly notice, though, that the antecedent's conditions sound strikingly familiar. In fact, they appear to be identical to the base conditions described above. Grice's initial presentation of the maxims does not include this "true and relevant" clause; rather, Grice's Quantity simply requires than an assertion be "as informative as possible." This simplified presentation, however, can lead to issues when maxims "clash," or have conflicting demands for the assertion. Most commonly, this clash occurs between Quality and Quantity, when a more informative assertion theoretically exists, but is not believed to be true by the speaker. (Grice, 1975) Grice implies a system of hierarchy between the maxims, where certain conditions take priority over others, but later neo-Gricean systems tend to cite the base conditions in Quantity (Katzir, 2007) (Gamut, 1991) to resolve clash within the maxims themselves. The following is a possible formulation of Quantity in this style, for a given proposition *p*:

(2) Dependent Quantity

 $Quan(p, s) \leftrightarrow \forall q((Qual(p, s) \land Rel(p, s)) \rightarrow p \subseteq q)$

However, Quantity could be re-written in a way that avoids referencing the base conditions directly:

¹ When written formally, Relevance sometimes includes reference to a set of propositions under discussion in a given conversation (the "question under discussion," or QUD) (Westera, 2022) as the condition for what is relevant and what is not. As the QUD does not play a role in this discussion, a simplified version has been adopted instead.

(3) Independent Quantity

 $Quan(p, s) \leftrightarrow \forall q(bel_s(p) \land bel_s(rel(p)) \rightarrow p \subseteq q)$

This alternative, "independent" Quantity is equivalent to the more common, "dependent" Quantity- the same conditions that fulfil the antecedent of the dependent fulfill those of independent, and vice-versa. In a "typical" conversation, where the maxims are assumed to act exactly as described, there is no difference between the two formulations

This equivalence importantly assumes that the maxims in (1) are invariable. However, let's assume momentarily an alternative version of the maxims in (1), (1'), with variable conditions:

(1') Variable Base Conditions

a. Quality': $Qual(p, s) \leftrightarrow \alpha(p, s)$ b. Relevance': $Rel(p, s) \leftrightarrow \beta(p, s)$

Under this version of the base conditions, equivalence can no longer be so easily assumed. Independent Quantity will always make reference to relation and belief in its antecedent, whereas dependent Quantity would use whatever α and β refer to as its conditions.

This would be a non-issue if α and β were themselves equivalent to the conditions in (1). However, it has been observed that maxims are sometimes altered or outright cancelled; the game show problem, for example, imagines a scenario where the maxim of Quantity is completely ignored (Fox, 2014), and some proposals to address this problem have actually suggested further cancellations (Westera, 2022). A dependent Quantity would suggest that any alterations made to the base conditions would have to be reflected in Quantity's antecedent. If Quantity were truly independent, however, we should observe it acting identically no matter the other maxims' status.

Although both formulations lead to the same inferences in a typical conversation, is there one that better reflects the reality of how we speak? In order to figure this out we've split this paper into two sections.

In the first section, we will focus on a scenario in which Quality alone is altered, but Quantity remains unchanged. This will allow us to gauge whether those changes are reflected in Quantity, and by extension, whether Quantity makes reference to the base conditions or stands as an independent maxim. However, maxims are impossible to observe directly. While it might be possible to come to conclusions about the maxims' true forms by assessing how speakers speak, we instead focus on the inferences a *listener* generates from various utterances in a given scenario. Similarly, while it may be theoretically

possible to find situations that alter Relevance as well, Relevance is extremely resistant to alteration, as listeners tend to interpret the relevance of a given utterance as generously as possible. (Gunlogson, 2011) This squib examines the case of a game show where participants are unsure who is a liar and who is telling the truth. The liars operate under what is essentially an altered Quality, one where they say what they expect someone else to believe rather than act on their own beliefs. As participants are aware of this condition, we can directly observe how it affects the implicatures they generate and if it better aligns with what is predicted by independent or dependent Quantity. Importantly, this paper assumes a fully pragmatic account of exhaustification and implicature, although alternatives, especially those that assume a grammatical exhaustivity operator (Fox, 2007), do exist. Ultimately, we find that an independent Quantity has no issue doing so.

The question of Quantity's dependency is not merely a bureaucratic exercise. In the second section, we use a case study to examine the consequences assuming one version of Quantity over the other can have on a proposal. In particular, we examine a piece in the literature (Westera, 2022) that cannot function without an independent Quantity, and examine how assuming dependency results in a proposal that is unable to generate the predictions it claims to be able to.

2 Altering Quality

In this section, we propose a novel scenario for altering Quality exclusively and examine the predictions generated by the two versions of Quantity to determine which version of Quantity better accounts for the scenario. We also examine a potential alternative scenario and explain why it would not function for our purposes. Section 2.1 introduces our scenario, the game show *To Tell the Truth*. Section 2.2 provides the main predictions generated by the two proposals, as well as their derivations. In section 2.3, we compare these predictions to those actually expected in the scenario. Finally, in section 2.4 we consider a possible alternative method of altering Quality, the case of a liar, and explain why it was not chosen over the more complicated game show scenario.

2.1 The Scenario

The case we are examining is that of the American game show *To Tell the Truth*. This case, as will soon become apparent, is rather complex, as it requires listeners to juggle two potential Qualities (one altered and one unaltered) simultaneously. However, it carries the distinct advantage of being a real-world example (or at least as "real world" as the rules imposed by a game show can be considered) and

therefore allows us to avoid pure hypotheticals. Additionally, as will be discussed later, other, seemingly less complicated options come with their own set of issues that make them unfit for the task at hand.

In a typical episode, a panel of celebrities are presented with a group of three "challengers" who all claim to be the same person. For instance, the challengers may all claim to be the trainer for a celebrity dog named Baxton. Of course, only one of these challengers can actually be the trainer (the "central character", or CC). The celebrities are tasked with posing questions to individual challengers to help them determine who is the central character and who are mere pretenders. When the central character is asked a question, they must respond as truthfully as possible. However, the pretenders are permitted to say anything at all in response. It follows, then, that although the central character follows the Quality described above, the pretenders have their own version, where *s* is the speaker and *c* is the central character:

(4) Alternate Qualities

a. Pretender's Quality $Qual(p, s) \leftrightarrow bel_s(\diamond bel_c(p))$

b. Central Character's Quality $Qual(p, s) \leftrightarrow bel_s(p)$

Although the pretenders are permitted to say anything they'd like, the format of the game encourages them to say things that would seem reasonable for the person they are imitating to say. This is also why we're able to argue that Quantity remains unaltered in this scenario. In attempting to sound like a reasonable alternative to the central character, the pretender adopts similar conditions to the central character. This includes the requirement to be maximally informative. Of course, since the pretender cannot be sure what the central character actually believes, this attempt to seem informative is still based off of their assumptions about the CC's beliefs, and not their own beliefs. The question of Independent vs. Dependent is ultimately a question of when the pretender attempts to embody the central character. Dependent Quality would mean that the pretender asks the question of "would the listener believe the central character would say this" to each individual proposition included in their final assertion, as well as the assertion itself. Independent Quality, on the other hand, only requires that the final, overall assertion be one the listener might believe the CC to say, and not the individual propositions, which are still subject to the beliefs of the pretender.

A panelist, of course, never knows whether they are speaking to a pretender or the central character. Thus, a panelist always needs to consider two cases when interpreting an assertion: CASE P,

where a pretender is speaking and (4b) is used, and CASE C, where the central character is speaking and (4a) is used.

2.2 The Predictions

So, with these conditions in mind, what can we observe about Quantity? Let's consider the dog trainer example presented above, and a question-answer interaction that might arise in such a scenario:

- (5) a. Panelist: "What did Baxton eat for breakfast this morning?"
 - b. Challenger: "Baxton had turkey and kibble for breakfast."

A panelist's thought process might follow along these lines:

CASE P: If we assume pretender's Quality is operative, the speaker does not necessarily believe (5). Rather, we can only conclude that they consider it plausible that the true central character would believe something like (5). Since the central character must not be the one speaking if we are assuming (4b) is operative, our final Quality inference for CASE P is that "the speaker is not the CC and the speaker considers it plausible that the CC could believe that Baxton had turkey and kibble for breakfast."

CASE C: Since (4a) is identical to standard Quality, we can conclude exactly what we would if this were a standard conversation: that "the speaker believes Baxton had turkey and kibble for breakfast," and, since we are assuming the CC is speaking, that "the speaker is the CC".

Of course, since the listener can't know exactly who is speaking, they must consider both equally, arriving at (6) for their final Quality inferences (where c is the central character, s is the speaker, and b is Baxton). Note that, for legibility's sake, the inferences from the two cases are presented independently; however, given that a listener must consider both cases possible, they are better considered as the disjuncts of a larger, unified inference.

(6) Quality Inferences

CASE P: $s \neq c \land bel_s(\diamond bel_c(bf t_b(turkey \land kibble)))$

"The speaker is not the CC and considers it possible that the CC believes Baxton had turkey and kibble for break fast."

CASE C: $s = c \land bel_s(bf t_b(turkey \land kibble))$

"The speaker is the CC and believes that Baxton had turkey and kibble for breakfast."

Now that we have the Quality inference, we can begin to consider the two Quantity inferences, beginning with that of Dependent Quantity. As with the Quality inference, the listener must consider the possibility both that the speaker is the pretender, and that the speaker is the true central character.

CASE P: We'll start by assuming the pretender is speaking. Since they assert (5), it must be the case that it fulfils their version of Quantity. By the definition of Quantity, (5) is the strongest statement that fulfills both Relevance and the pretender's Quality. Thus, the fact that they don't assert the stronger statement "Baxton had turkey, kibble, and fish for breakfast" implies that it fails the conditions of either Relevance or Quality. Since the stronger statement is clearly relevant to the question, it must be that it does not fulfil Quality and therefore "the speaker must be uncertain whether they consider it plausible that the central character believes Baxton had fish for breakfast."

CASE C: We'll now assume the central character is speaking. As before, in order to assert (5) it must obey the CC's Quantity. Thus, any stronger statement than (5) is either irrelevant or does not obey Quality. Since the stronger statement "Baxton had turkey, kibble, and fish" is clearly relevant, it must not be the case that the CC believes it to be true. Thus, we reach the inference that "the central character is uncertain whether Baxton ate fish for breakfast."²

By combining the conclusions from the two cases and the Quality inferences from before, we can create a final set of "complete" inferences for (5) using dependent Quantity:

² Typically, an additional, "epistemic," step is taken after reaching this inference, where the listener assumes that, since the speaker has good reason to know what Baxton had for breakfast, the reason they are uncertain is because they don't believe Baxton had fish at all. However, since this step is irrelevant to the arguments made in this squib, it is ignored here and in all further locations it could be included for the sake of simplicity.

(7) Final Inferences (Dependent Quantity)

CASE P: $s \neq c \land bel_s(\diamond bel_c(bf t_b(turkey \land kibble))) \land \neg bel_s(\diamond bel_c(bf t_b(fish)))$

"The speaker is not the CC and is uncertain whether the CC believes Baxton had fish for breakfast."

CASE C: $s = c \land bels(bf t_b(turkey \land kibble)) \land \neg bel_s(bf t_b(fish))$ "The speaker is the CC and is uncertain whether Baxton had fish for breakfast."

Independent Quantity, as expected, behaves quite differently. Most notably, although in theory the listener must consider two cases when deriving any Quantity inferences, since dependent Quantity never varies, those two cases proceed identically.

CASE P/C: In order to assert (5), it must be the case that it obeys Quantity. Thus, any stronger statements, such as "Baxton had turkey, kibble, and fish for breakfast," must either not be relevant, or not be believed by the speaker. Since the stronger statement is clearly relevant, the speaker must not be certain it is true. Thus, it can be concluded that "the speaker is not certain whether Baxton had fish for breakfast."

Like above, we can combine this inference with the Quality inference to reach these final inferences:

(8) Final Inferences (Independent Quantity)

CASE P: $s \neq c \land bel_s(\diamond bel_c(bf t_b(turkey \land kibble))) \land \neg bel_s(bf t_b(fish))$

"The speaker is not the CC, considers it possible that the CC believes Baxton had turkey and kibble for breakfast, and is uncertain whether Baxton had fish for breakfast"

CASE C: $s = c \land bel_s(bf t_b(turkey \land kibble)) \land \sim bel_s(bf t_b(fish))$

"The speaker is the CC and believes that Baxton had turkey and kibble, but not fish, for breakfast."

2.3 Assessing the Predictions

Now that we have our predictions, let's take another look at (5):

(5) a. Panelist: "What did Baxton eat for breakfast this morning?"b. Challenger: "Baxton had turkey and kibble for breakfast."

Putting ourselves in the shoes of a panelist, it makes sense that CASE C is identical no matter what Quantity is operative. The listener assumes that the central character is honest and cooperative, and thus that they follow a version of the base conditions identical to those in (1). (1) is, as discussed, the version of the base conditions that render the two Quantities equivalent, and thus it makes sense that the inferences under these circumstances would be similarly equivalent. CASE P, on the other hand, demonstrates a noticeable divergence between the two Quantities. Independent Quantity would have us believe that a listener would conclude that a pretender would be uncertain whether the CC believes Baxton to have had fish for breakfast. At first this seems to be a reasonable conclusion to come to; after all, why should a pretender have any knowledge of Baxton's breakfast? However, a pretender, according to their altered Quality, is assumed to only say things that they believe the CC could believe. Why, then, is the pretender's personal belief about fish for breakfast relevant? It seems more reasonable to assume, as a listener, that the reason for fish's omission is not because of personal belief about Baxton's breakfast, but rather uncertainty about what the CC might believe Baxton's breakfast to be, as is stated in (7). With this in mind, it appears as though assuming Dependent Quantity more accurately reflects the thought process a panelist might have when interpreting an interaction like (5).

2.4 Why Not Something More Simple?

Before we move on to our case study, let's take a moment to discuss why the *To Tell the Truth* example was chosen. It would be entirely reasonable to question why it is we resort to such a contrived scenario to alter Quality. A potentially simpler option may come to mind, namely the case of a liar. It appears cut and dry. After all, what better way is there to change the maxim which says "tell the truth" to simply "do not tell the truth?" Let's imagine our listener has a friend, someone who famously never tells the truth. More specifically, they only say something if they believe it is not true. Like in the *To Tell the Truth* example, this fact is known to the listener, so they're able to make judgements about what the liar is saying. We can even imagine a possible formulation of this liar's Quality looking something like this:

(9) Liar's Quality

 $Qual(p, s) \leftrightarrow bel_s(\sim p)$

Like before, we can take this new Quality and see what the predictions would look like for our Dependent and Independent Quantities. Let's use a simplified version of (5) to see what sorts of inferences are drawn.

(10) a. Listener: "What did Baxton eat for breakfast this morning?"b. Liar: "Baxton had turkey for breakfast."

As in the *To Tell the Truth* example, the Quality inference is unaffected by which version of Quantity is operative. Thus, the inference process for Quality would look something like this: We know the liar only says things they believe not to be true. Thus, the fact that they said that Baxton had turkey means that they actually believe Baxton did not have turkey.

(11) Quality Inference (Liar's Quality)

 $bel_s(\sim bf t_b(turkey))$

We can then consider the inferences that would be predicted by our two versions of Quantity:

DEPENDENT: The liar did not assert "Baxton had turkey and fish for breakfast." Since the conditions of Quantity dictate that the liar include everything relevant they believe is false in their assertion, it must be the case that "Baxton had fish" is either irrelevant or not believed to be untrue by the liar. Since "Baxton had fish" is clearly relevant, however, we can conclude that "the liar is not certain that Baxton did not have fish for breakfast."

(12) Final Inference (Dependent Quantity)

$bel_s(\sim bf t_b(turkey)) \land \sim bel_s(\sim bf t_b(fish))$

"The speaker believes Baxton did not have turkey and is not certain Baxton didn't have fish."

INDEPENDENT: The liar did not assert "Baxton had turkey and fish for breakfast." Since the conditions of Quantity dictate that the liar include everything they consider true and relevant in their assertion, the fact that they did not include fish in the assertion implies that they either are not certain Baxton had fish or they consider that fact irrelevant. Since it is clearly relevant, we must conclude that "the liar is not certain that Baxton had fish for breakfast."

(13) Final Inference (Independent Quantity)

 $bel_s(\sim bf t_b(turkey)) \land \sim bel_s(bf t_b(fish))$

"The speaker believes Baxton did not have turkey and is not certain that Baxton had fish."

While at first these predictions seem rather similar in the sense that they both imply uncertainty about other food items, this changes if we generalize fish to "any food item not mentioned," as can easily be done since no other food items are mentioned, and then take the epistemic step. Suddenly, Independent Quantity would predict that "Baxton had nothing to eat" and Dependent would say "Baxton had everything except turkey to eat," two extremely divergent interpretations of the same sentence. Normally, our next step would be to look at what listeners actually predict in the scenario presented, as we will with our original game show scenario, but in this case, that may not be possible. Let's first take a look at some irregularities that may have been noticed while presenting the above information.

The first irregularity pertains to Independent Quantity. More specifically, how would it operate in this scenario at all? Our predictions above take the assertion in (10) at face value, and then apply the maxims to it. But how could (10) have been uttered at all under these conditions? Liar's Quantity says the liar can only say what they believe is untrue, yet, simultaneously, Independent Quantity says they must include everything they consider true in the assertion. This would mean, for example, that if the liar knows Baxton had kibble and nothing else for breakfast, that they would somehow have to express "Baxton had kibble" (as required by Independent Quantity) while also only saying things they consider untrue (as required by the liar's Quality). This simply isn't possible, and before even considering the actual inferences, drops Independent Quantity from the running.

Independent Quantity not working in this scenario isn't necessarily a win for Dependent Quantity, however. Our second irregularity is Dependent Quantity's prediction that the liar believes Baxton ate everything but turkey for breakfast. While there is certainly some wild scenario where this may be possible, in all reasonable cases this seems unreasonable. Additionally, as with Independent Quantity, there remains the question of how a sentence like (10) can be asserted at all, given the conditions. For instance, if we take the kibble example again, the liar's expected assertion would be "Baxton had everything except kibble for breakfast." Certainly more reasonable than the outright contradiction the conditions imposed by Dependent Quantity, but not accurate to what one might expect of a liar, either.

So what does this mean, then? Is the liar scenario proof that we actually need some sort of third Quantity that accounts for these sorts of situations? Not at all. Something this scenario critically assumes is that Quantity would be operative at all when speaking with a liar. Why should we make this assumption, though? A liar is an inherently uncooperative speaker. Maximizing informativeness is completely irrelevant when your assertions are inherently uninformative. The liar situation provides a simple alteration of Quality, sure, but fails to account for the fact that such a scenario would alter Quantity as well, leaving it unfit for our needs. Unlike the liar scenario, however, the conditions imposed by *To Tell the Truth* allow Quantity to remain operative even when considering the pretender's

26

perspective, as discussed above. Certainly, the game show chosen makes things a bit more complicated, but it allows us to control for the other maxims far better than other, more simplified scenarios.

3 Case Study: Attentional Pragmatics

This section examines Matthijs Westera's 2022 proposal *Attentional Pragmatics*, and, more specifically, how it addresses Danny Fox's Game Show Problem. We find that Westera's proposal, which introduces a set of 3 additional maxims parallel to the traditional neo-Gricean maxims, cannot account for the dilemma introduced by the Game Show Problem without assuming an independent Quantity.

3.1 Attentional Pragmatics

Matthijs Westera's *Attentional Pragmatics* (2022) bills itself as a solution to a number of problems concerning the pragmatic (neo-Gricean) approach to exhaustivity. One of the issues addressed is the Game Show Problem, the proposed solution to which this paper will focus on specifically. In order to tackle these problems, Westera argues, an additional set of Quality, Relevance, and Quantity maxims can be introduced, paralleling the originals but concerning themselves with what an utterance draws attention to, rather than what an utterance asserts. These attentional (or A-) maxims do not replace the original neo-Gricean maxims (which Westera refers to as informational, or I-, maxims), but instead provide an additional pathway for deriving exhaustivity when the I-maxims are insufficient.

A-maxims do not apply to an assertion, but rather to an "attentional intent" (A), a set of all things a speaker intends to draw attention to with an utterance. Since the A-maxims all deal with attention, they are all written in terms of the attentional intent:

(14) The A-Maxims (Westera, 2022)

a. A- $Quality(A) = \forall a(A(a) \rightarrow \diamond(`a \land \forall b((b \sub a \land A(b)) \rightarrow \sim`b)))$

'Intend to draw attention only to propositions you consider possible, in particular, possible [sic] independently of any thing stronger to which you intend to draw attention.'

b. *A-Relevance*(A, Q)³ = $\forall a(A(a) \rightarrow Q(a))$ 'Intend to draw attention only to relevant propositions.'

c. A-Quantity(A, Q) = $\forall a((A-\text{Quality}(\{a\}) \land A-\text{Relevance}(\{a\}, Q)) \rightarrow A(a))$ 'Intend to draw attention to all relevant propositions you consider independently possible.'

Though these maxims parallel the I-maxims, there are some key differences. Most notably, A-Quality requires that a speaker consider a proposition possible, whereas its counterpart I-Quality would require that a speaker consider a proposition true. This will later become relevant when these maxims are put into the game show scenario, which will be discussed in the next section. In order to solidify our understanding of these maxims, and how they would allow exhaustivity to be derived without using the I-maxims, let us reconsider the example of Baxton's breakfast, this time outside of the context of *To Tell the Truth*.

(15) Alan: "What did Baxton have for breakfast this morning?" Brenda: "Baxton had turkey and kibble for breakfast."

First, we should figure out what the attentional intent is for Brenda's response. Brenda draws attention to two things that Baxton had for breakfast: turkey and kibble. At first, one might assume that the attention set looks like this: {*bf* $t_b(turkey)$, *bf* $t_b(kibble)$ }. However, conjunctions like the one in Brenda's response are actually considered propositions of their own under Westera's model. This means that our intentional intent actually looks like this: {*bf* $t_b(turkey \land kibble)$ }.

³ The Q here refers to the Question Under Discussion, a set of all propositions relevant at a given point of the conversation. "Q(a)" can be considered roughly equivalent to the "bel_s(rel(a))" seen in the neo-Gricean maxims above.

⁴ Conjunction is the only connector that behaves this way. The attentional intent of the disjunctive version of (12B), "Baxton had turkey or kibble for breakfast," would not include the proposition *bf* t_b (*turkey V kibble*), for example.

With an attentional intent in hand, we can move on to the actual derivation of exhaustivity. To start, we consider a proposition not found in A, such as "Baxton had fish for breakfast." Under A-Quantity, Brenda would only have two reasons for not mentioning fish: either it is not relevant, or they do not consider it possible. One could easily imagine fish being a possible, and therefore relevant, response to (15A), so it must be the alternative. Thus, it must be that Brenda doesn't consider it possible that Baxton ate fish, our exhaustive reading. Additionally, as a consequence of conjunctions forming their own propositions, the absence of the individual propositions "Baxton had turkey" and "Baxton had kibble" implies that Brenda does not consider them individually possible that is to say, that Brenda doesn't think Baxton could have eaten turkey without also eating kibble, and vice-versa.

In contrast to the traditional neo-Gricean derivation, this process does not require an epistemic step. Instead, since A-Quality deals in possibility rather than necessity, the conclusion that a proposition does not obey A-Quality is enough to conclude a speaker does not believe something. Other than this lack of an epistemic step, however, the A-maxims derive exhaustivity in a very similar way to the standard neo-Gricean approach.

3.2 The Game Show Problem

Of course, Westera's solution cannot be examined without first having an understanding of what the Game Show Problem is. The problem was first introduced in Fox's *Cancelling the Maxim of Quantity: Another challenge for a Gricean theory of Scalar Implicatures* as a gap in a purely pragmatic approach to ex haustivity. Fox's paper, much like this one, presents a game show scenario where the Gricean maxims appear to be altered. Here, readers are asked to imagine a simplified game where a host provides clues as to which of a number of boxes contain a cash prize. The host knows the location(s) of the money, but does not reveal it directly. Additionally, the contestant is aware that the host knows the money's location. For example, a host may provide a clue like this:

(16) Host: "There is money in box 20 or 25." (Fox, 2014)

The maxim of Quantity requires that a speaker be as informative as possible. For the host to utter something like (16), Fox argues, they must not obey Quantity, as otherwise they would have chosen to reveal the exact location of the money, which is known to them.

An utterance like (16), outside of the context of this game show, should come with two inferences: that the speaker does not know exactly where the money is (the "ignorance inference"), and that there is not money in both boxes (the "exhaustivity inference"). Both of these inferences, critically, are derived from Quantity. In the case of the ignorance inference, the fact that the speaker does not

specify which of the two boxes the money is in, it must be the case that doing so would violate the conditions of either Quality or Relevance. Since the location of the money is clearly relevant, it must be that specifying would violate Quality, and thus, the speaker doesn't have reason to believe it's in one box or the other specifically. The exhaustivity inference follows much the same pattern, this time, however, considering the assertion that the money is in both boxes. As before, since the speaker doesn't assert this over the disjunction, and it is clearly relevant, it must be the case that the speaker doesn't have reason to believe it of assuming that the speaker would have reason to know it were in both boxes if it were true (the "opinionatedness assumption"). Thus, the fact that they don't have reason to believe it is in both boxes.

One would expect, of course, that in this game show, where Quantity is cancelled, neither of these inferences would be present. The ignorance inference is, indeed, absent; after all, the host, by the setup of the game, must know the location of the money exactly. However, as Fox observes, the exhaustivity inference persists. Take, for example, this retort by the contestant upon the reveal that the money was in both boxes:

(17) Contestant: "What you said was wrong. You said there was money in box 20 OR box 25. But, in fact, there was money in both boxes." (Fox, 2014)

If there was no implication that the money could not be in both boxes, the contestant could not take issue with the presence of money in both. The only way (17) is an acceptable retort is if (16) were truly exhaustive.

This presents a problem for the neo-Gricean method of deriving exhaustivity. If exhaustivity is truly a product of Quantity, why is it found here, where Quantity is not operative? Fox argues that this is evidence for an alternative, non-pragmatic pathway for inferring exhaustivity. Westera, however, proposes an additional set of maxims which should, in his estimation, provide a pragmatic method for finding exhaustivity that does not depend on Quantity.

3.3 Attentional Pragmatics and the Game Show Problem

Now that we understand the basics of Attentional Pragmatics, we can begin to unpack how Westera applies it to the Game Show Problem. Westera does not take issue with anything presented in 3.2, however, he does introduce a new assumption: that the host may pretend to know less about the locations of the prizes than they actually know, but may not pretend to know more (in other words, lie). With this in mind, Westera concludes that A-Quality must not be operative, as, by pretending to know less information, a host allows themselves to include propositions they do not consider possible. The other A-maxims remain untouched, however. Westera argues that since A-Quantity is still operative, the attentional pathway can still derive exhaustivity, using the method described in 3.1.

This conclusion is debatable, however. To see why, let us consider A-Quantity. At first, this seems to be a clear case of a dependent Quantity. After all, A-Quality and A-Relation are directly cited in the formalization. On the other hand, the base conditions (or their A-maxim equivalents), are not referenced at all in the informal version. Instead, the written form only refers to "relevant propositions you consider independently possible," without any mention of other maxims. While at first this would seem a simple case of ambiguity, the choice of an independent or dependent Quantity is crucial for being able to derive exhaustivity in this case.

Let's start with what is, as established in the above sections, the typical assumption for a Quantity: that it is, indeed, dependent on the base conditions. As Westera describes, A-Quality has been cancelled. At first, cancellation would appear to be outside the scope of this paper, which deals with Quality being altered, rather than outright absent. However, this cancellation can easily be considered a case of alteration, instead. More specifically, since the host can draw attention to propositions they don't consider possible, A-Quality is weakened to the point of being a tautology:

(18) Host's A-Quality

$A-Quality(A) = \forall a(A(a) \to \Diamond^{\vee} a \lor \sim \Diamond^{\vee} a)$

'Intend to draw attention only to propositions you consider possible or do not consider possible.'

Considering cancellation as a case of alteration rather than absence allows us to avoid the issue of what happens when A-Quantity refers to something that isn't there. Of course, (18) no longer places any restrictions on what fulfills A-Quantity, leaving relevance as the only condition with any force. This, however, would force the inference that a proposition like "there is money in boxes 20 and 25" is somehow irrelevant to the matter at hand, which it clearly is not. Worse than predicting an incorrect inference, a dependent A-Quality predicts nothing at all!

Since a dependent A-Quality is a non-starter, let's move on to the independent variation. This would allow for correct predictions, as the cancellation of A-Quality would have no impact on the conditions of A-Quantity. However, for this to be the case, Westera would need to make explicit the independence of A Quantity from the attentional base conditions, and, more importantly, account for why the A-maxims differ from the I-maxims in this regard.

Conclusion

At first glance, Dependent and Independent Quantity seem totally equivalent. It makes sense, then, that the formulation of Quantity as dependent has been taken for granted as long as it has. However, as we establish with the case of *To Tell the Truth* establishes, making Quantity dependent on the base conditions is crucial for its functioning in scenarios where the base conditions are altered.

References

Fox, Danny. 2007. Free Choice and the Theory of Scalar Implicatures. In Presupposition and

Implicature in Compositional Semantics, ed. Uli Sauer land and Penka Stateva, 71–120. London: Palgrave Macmillan UK. URL <u>http://link.springer.com/10.1057/97802302107524</u>.

Fox, Danny. 2014. Cancelling the Maxim of Quantity: Another challenge for a Gricean theory of Scalar Implicatures. *Semantics and Pragmatics* 7. URL http://semprag.org/article/view/sp.7.5.

Gamut, L. T. F. 1991. Logic, language, and meaning. Chicago: University of Chicago Press.

- Grice, H. P. 1975. Logic and Conversation. In Speech Acts, ed. Peter Cole and Jerry L. Morgan, 41 58. BRILL. URL https://brill.com/view/book/edcoll/9789004368811/BP000003.xml.
- Gunlogson, Christine. 2011. De-Maxim-izing Quality. In *Working papers in the language sciences at the university of rochester*, ed. Alex Fine. University of Rochester. URL https://www.sas.rochester.edu/cls/assets/pdf/working/Gunlogson.pdf.
- Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30:669–690. URL http://link.springer.com/10.1007/s10988-008-9029-y.
- Westera, Matthijs. 2022. Attentional Pragmatics: A pragmatic approach to exhaustivity. *Semantics and Pragmatics* 15. URL http://semprag.org/article/view/sp.15.10.

The Computation of Scalar Implicatures: Pragmatic, Lexical or Grammatical? © 2024 by Jordan Webster is

licensed under CC BY-ND 4.0

The Computation of Scalar Implicatures: Pragmatic, Lexical or Grammatical?

Jordan Webster

Department of Linguistics, McGill University LING 565: Pragmatics Professor Bernhard Schwarz

Abstract

The problem that I focus on in this paper is whether scalar implicatures are truly derived through pragmatic means. I provide an overview of the alternatives to the pragmatic approach that I have come across, as well as provide my own thoughts on the matter and offer some questions for further examination. I begin with an introduction for the non-linguist of scalar implicatures and the (neo-)Gricean theory as described by Geurts (2011), after which I identify two other competing theories laid out by Sauerland (2012), focusing on some-all implicatures. I then explain the advantages that the grammatical theory has over the pragmatic theory, with an emphasis on SIs embedded in disjunctions that Sauerland and Chierchia et al. (2013) point out. Considering these arguments, I offer a critical review of these papers and conclude that, despite the pragmatic theory facing challenges still, it seems that some degree or pragmatic reasoning is essential to understanding scalar implicatures, even if that means accepting a pragmatic-lexical hybrid approach.

1 Introduction: The Pragmatic Theory of Scalar Implicatures

Scalar implicatures have long been a phenomenon puzzling both linguists and philosophers of language. When taking an introductory logic course, students may be initially puzzled by the fact that in classical first-order logic, *some* is consistent with *all*. That is, given (1), it is often intuited that (1') is false:

- (1) Some students failed the exam.
- (1') All students failed the exam.

In reality, (1) is perfectly consistent with (1'), and a professor who announces to his class that "some students failed the exam" could do so truthfully in a world where—to the shock of his students—

they all failed. This confusion arises due to the phenomenon of *conversational implicature*: when one engages in conversation, not only is meaning taken from what is said, but assumptions are also made about the speaker's intention and knowledge to take meaning out of what is *not* said.

The type of conversational implicature observed in (1)-(1') is known as *quantity implicature* because its derivation involves reasoning with Grice's Maxim of Quantity: "Make your contribution as informative as is required" (Grice 1975). If the professor says (1) to his students, he could have said (1'), which is more informative. Nonetheless, he chose to say (1), violating the Maxim of Quantity. The students reason that he must have violated the maxim for a reason; in this case, it is likely *not true* that *all* students failed the exam. Thus, from the speaker's utterance of (1) alone, the listener infers that the speaker must believe (\neg 1'):

 $(\neg 1')$ It is not the case that all students failed the exam.

This path of reasoning is why the students would be so shocked in the case that they *did* in fact *all* fail the exam: it would contradict this quantity implicature. This variety in quantity implicature may be more specifically called *scalar implicature* (SI). SIs involve generating an alternative sentence by replacing one expression with another, more informative, expression in the same *Horn scale*. Given the Horn scale in (2), it is reasonable for (1') to be considered an alternative to (1), and, as (1') is more informative than (1), why it is employed in the computation of the SI (\neg 1').

(2) {*some*, *all*}

To be more precise, this conception of the computation of SIs involves the listener going through four steps that have been adapted from Geurts (2011) and laid out in (3).

(3) The speaker (S) has said (1).

i. S could have said the more informative (1'). Why didn't he?
ii. It's probably because S doesn't believe that (1') is true.
iii. Furthermore, S must have an opinion as to whether (1') is true: either S believes (1') is true or S believes (1') is false.
iv. By disjunctive syllogism, (ii) and (iii) entail that S believes (1') is false. (Geurts 2011, p. 32)
These steps are central to this approach towards SIs and will be referenced throughout this paper. Crucially, this conception is inherently *pragmatic*: it relies on reasoning about what the speaker believes. However, certain problems arise in the pragmatic theory of SIs (henceforth *the pragmatic theory*, or *Gricean theory*) that have led linguists to conceive of alternative approaches to provide an explanation for this phenomenon. Namely, the pragmatic theory seems unintuitive when considering cognitive efficiency, and makes the wrong predictions. In the following two sections, there is an elaboration on these challenges and an exploration of the alternative theories that have arisen in response.

2 The Lexical Theory

Sauerland (2012) lists three alternatives to the pragmatic theory: the lexical, pragmatic+lexical, and grammatical theories. The lexical and grammatical theories are discussed in this section, and the pragmatic+lexical theory returned to in the last section. When considering the pragmatic approach, what are perhaps the most obvious doubts that come about all pertain to the *efficiency* of computing SIs in this way. When a listener hears the utterance in (1), the inference that S believes (\neg 1') seems automatic and instantaneous; it seems that in an everyday situation, (1) will always include (\neg 1') in its meaning and there is no need to reason about speaker intentions in order to know this. Is it really necessary to go through steps (3, i) to (3, iv) to arrive at the same inference every time?

One may imagine that these inferences may actually be part of the *lexicon*. Lexical theories propose that SIs can be attributed to the memorized vocabulary of a language, rather than any logical derivation that relies on general, non-linguistic assumptions about the speaker's cooperativeness (e.g., Grice's Maxim of Quality). Perhaps *some* as it is used in English always contains a silent *not all*, so that (1), for example, is always taken to mean (4):

(4) Some *but not all* of the students failed the exam.

Surely, this is a much more cognitively efficient approach, as the derivation of the pragmatic inference is rendered obsolete. However, SIs are not quite as "automatic and instantaneous" as one might assume. There are cases where one may *not* wish to derive a SI at all, such as the use of *some* in (5) adapted from a class lecture:

(5) Speaker A: If some students failed, I will be disappointed.

Speaker B: #You will be disappointed even if all students failed.

In this case, Speaker A's utterance is not taken to mean "if some *but not all* students failed, I will be disappointed." This is because of the intuition that in this situation, Speaker A will be disappointed even if all students failed, hence the feeling of oddity towards Speaker B's response (he does not seem to be adding new information to the discourse). In sum, *but not all* cannot always be appended to the meaning of *some*, which suggests that SIs go beyond the lexical meaning of words.

Furthermore, Sauerland (2012) notes that the lexical theory makes the wrong predictions when "scalar expressions are embedded under an environment that reverses entailment relations" (Sauerland 2012, p. 12). One such example is adapted from Sauerland in (6):

(6) Elizabeth doesn't like all of Drake's songs.

(7) Elizabeth likes some of Drake's songs

(Sauerland 2012, p. 11)

(6) is predicted to give rise to the SI that (7), due to the negative predicate *doesn't like* reversing the entailment relation, *some* becomes more informative than *all*. By applying the pragmatic reasoning steps in (3) to sentence (6), one arrives at the SI that S believes that it is not the case that Elizabeth doesn't like some of Drake's songs, which is equivalent to S believing (7): the correct prediction. Under the lexical theory, however, a SI would not arise at all, as it is impossible to reach the correct prediction by adding a similar "silent" SI to the meaning of *all* as the lexical theory does for *some (but not all)*.

For this reason, and because of the problem observed in (5), the lexical theory is not widely accepted in current literature, and the remainder of this paper will pay more attention to the other theories mentioned by Sauerland.

That said, another weakness of the pragmatic theory is the inability to account for *local* SIs. The pragmatic theory is only able to correctly predict *global* SIs— that is, "it can only apply to entire speech acts" (Sauerland 2012, p. 9). Because the pragmatic theory involves S's beliefs about the utterance as a whole, the wrong predictions are made for "local" environments where the SI is *embedded*, such as within a disjunct. The lexical theory does account for these local SIs, but as seen, it falls victim to other

challenges. One theory that does not seem susceptible to these challenges, however, is the grammatical theory.

3 The Grammatical Theory

These next two sections focus on SIs that arise in sentences containing *or*. The Horn scale for *or* is $\{or, and\}$, meaning that the sentence adapted from Chierchia et al. (2013) in (8) would have the alternative, more informative sentence in (8'):

- (8) Jonas or Isabella will show up.
- (8') Jonas and Isabella will show up.

(Chierchia et al. 2013, p. 2299)

Intuitively, (8) is taken to mean that either Jonas or Isabella will show up, but not both; i.e., one arrives at the SI that (\neg 8'), which seems consistent with the pragmatic derivation of SIs detailed in (3). However, under disjunctions that make use of *Hurford's Constraint*, the pragmatic approach makes the wrong predictions. Provided below is a paraphrase of Chierchia et al. (2013)'s definition of Hurford's Constraint in (9) and just one example of the shortcomings of the pragmatic theory with regard to disjunction and Hurford's Constraint in (10).

(9) **Hurford's Constraint (HC):** A sentence with a disjunctive phrase is infelicitous if either one of the disjuncts entails the other.

(10) a. Aashiha finished some of her exams.

b. Aashiha finished some or all of her exams.

(Chierchia et al. 2013, p. 2309)

(10) makes use of HC in the capacity that (10b) can only be interpreted as felicitous if *some* is taken to mean *some but not all*, or else the second disjunct will entail the first. The problem for the pragmatic theory arises in the fact that even if *some* is taken to mean *some but not all* in (10b), (10a) and (10b) are equivalent.

The problem is only obviated if it is assumed that there is an *exhaustivity operator* that applies to the first disjunct. Such an operator would be *grammatical* in that it would apply selectively to this one constituent of the sentence. Only grammatical rules can predict this behaviour, as, once again, pragmatic theories apply globally to the entire utterance in order to reason about speaker intentions and thus cannot posit any grammatical operators on an individual constituent level. Because of this, the pragmatic theory could derive from (10b) the inference that "Aashiha finished some but not all or all of her exams," but this would not account for the slight difference in (implicated) meaning between (10b) and (10a).

On the other hand, Chierchia et al. discuss the possibility of the additional alternatives L and R under a grammatical theory. When an exhaustivity operator is applied to a disjunctive phrase, the Horn set for *or* is taken to be {*or*, *L*, *R*, *and*}, where *L* takes the left disjunct and excludes the right disjunct, and *R* takes the right disjunct and excludes the left (Chierchia et al. 2013, p. 2314). This is exemplified in (11):

(11) Aashiha finished some L all of her exams. \Leftrightarrow Aashiha finished some of her exams.

If it is assumed that the exhaustivity operator is applied only to the first disjunct of (10b) and observes the alternative in (11), it is acceptable for the additional alternative to (10b) of "Aashiha finished some *but not all* of her exams," period, right disjunct excluded. Because this alternative is not consistent with (10a), which leaves open the possibility that Aashiha *did* finish all of her exams, the problem with distinguishing the implicatures of (10a) and (10b) is obviated.

Yet another challenge involving HC is what Sauerland (2012) refers to as *intermediate implicatures*. Take this adapted example from Sauerland in (12):

(12) Either Josh didn't read every page or he read no pages.

(Sauerland 2012, p. 19)

The grammatical theory is readily able to account for an example like this. Because "no pages" entails "not every page," the only way that HC is not violated in (12) is if the first disjunct is taken to mean (13):

(13) Josh didn't read every page, but did read some pages.

Once again, this requires applying an exhaustivity inference at the constituent-, and not utterance-, level. What is unique about this example, however, is that not only is the SI not global, it is not *local* either. This is due to a similar problem to what is seen in (6)-(7); the negative environment of the first disjunct reverses the entailment relations so that it may give rise to the SI in (13), which cannot occur if the implicature is taken to be locally attached to the lexical meaning of *every*. This concept is briefly returned to in the following section. To summarize, the grammatical theory accounts for SIs that interact with Hurford's Constraint in a way that the pragmatic theory cannot, as a grammatical exhaustivity operator would allow for the computation of SIs embedded in a disjunct, which the pragmatic theory fundamentally cannot do.

4 Critique & Discussion

This section touches on an attempt that has been made to counter the local SI problem for the pragmatic theory, especially considering a combined pragmatic+lexical theory. Following this is an explanation of how this still seems to fall short, however, a problem will be brought up with the grammatical approach that has not been seen to be addressed enough in the literature. The section will then conclude with an argument that the challenge within the grammatical approach is significant enough to favour a pragmatic theory, but there is extensive work to be done in order to fully accept the pragmatic approach.

Firstly, note that Russell (2006) offers a solution to some forms of disjunctive sentences that have historically been seen as counterexamples to the pragmatic theory. In particular, he details a logical derivation that employs some of the same tactics as Sauerland and Chierchia et al. (namely, considering individual disjuncts as alternatives to a disjunctive phrase) to *sometimes* predict the correct outcomes, which the grammatical theory literature has not given credit for. For the purposes of this paper, this logical derivation will be excluded simply because the grammatical theory has yet to overcome the instances of SIs interacting with Hurford's Constraint, a great feat for the grammatical approach. However, Russell (2006) is mentioned because it does offer hope for the pragmatic theory which the papers advocating for the grammatical theory did not adequately address. It is plausible that what are seen as "local" SIs are really not so distinct from their global counterparts. With some additional assumptions about speaker intentions that may not have been historically included in the pragmatic theory and a more careful consideration of the logical structure underlying a pragmatic approach to SIs, the Gricean theory may be able to account for even HC examples with speaker assumptions and logic alone. The pragmatic theory does, after all, have a common-sense advantage over the grammatical approach and, as such, should be intuitively preferred over the grammatical approach if ever a method is offered to account for

instances such as HC. As Sauerland (2012) himself states, "[t]he pragmatic account is an outgrowth of common sense reasoning" (p. 16). To illustrate this point, consider a passage from Chierchia et al. (2013).

Referring back to (8) and (8'), according to Chierchia et al. and following the derivation outlined in (3), the pragmatic theory would lead to the inference that "*it is likely* that the speaker takes (8') to be false" (p. 2299). Chierchia et al. note that although this seems to resemble the target SI, "[w]hat we actually want to draw is that the speaker is positively trying to convey that [Jonas] and [Isabella] will not both come," rather than simply make a statement about the likelihood of what the speaker believes (p. 2300). To open a debate on this argument, consider the following example which has been modeled after another that was the center of an in-class discussion:

(14) Chloe took some of the pears from the bowl.

Grammatical theories are often quick to criticize the context-sensitivity and dependency on *speaker competence*—the assumption in (3, iii) that is crucial for the derivation of SIs under the pragmatic theory. However, it is exactly this assumption that makes the pragmatic theory so appealing. Considering (14), in the context that there was something blocking the speaker's view of the pear bowl (e.g., a cat, a vase), the strongest implicature that can arise is that the speaker *does not believe* that Chloe took all of the pears from the bowl. That is, the competence assumption disappears, and the derivation is halted at step (3, ii). On the other hand, in the context that the bowl is in plain sight for the speaker, not only is it inferred that it is *likely* the speaker believes that Chloe *did not take all* of the pears, as Chierchia et al. contend, one *knows* that the speaker believes this! The competence assumption allows for both the contextual cancellation of the SI as well as a "positive conveying" of the SI. If anything, the pragmatic theory is more accurate on this front, given that the grammatical theory *cannot* predict the implicature cancellation in specific contexts. As Russell (2006) puts it, "[t]he enrichment of weak implicatures with contextual inferences in a global, Gricean framework correctly mirrors the observed fine-grained context-sensitivity of scalar implicatures" (p. 364).

That said, the pragmatic theory is still far from overcoming the problems that proponents of the grammatical theory have posed. Most alarmingly, the failure of the pragmatic theory to account for many varieties of local implicatures (those varieties not accounted for in Russell (2006) and Geurts (2011), such as in (10)) must somehow be repaired. One solution mentioned by Sauerland (2012) is a combination of

the pragmatic and lexical theories into one (P+L theory) (p. 15). Where the lexical theory fails to account for SIs that *must* be global or those under reverse-entailment environments, pragmatic reasoning may step in, whereas local SIs may be as simple as the lexical theory suggests, all of which may somehow be married into a single account. However, Sauerland's intermediate implicatures such as in (12) would force such a theory to also account for those SIs that are neither local nor global, which raises another challenge yet.

5 Conclusion

This paper has investigated the varying approaches to SIs in search of an alternative to the pragmatic theory. Considering local implicatures, the lexical theory was evaluated, coming to the conclusion that the lexicon alone cannot explain implicatures that are necessarily global nor those in reverse-entailment environments. This was followed by a discussion of Sauerland and Chierchia et al.'s use of Hurford's Constraint to justify the grammatical over the pragmatic theory. In conclusion, it is more plausible to construct a theory that incorporates pragmatic reasoning than to abandon pragmatic influences entirely and thus somehow account for contextual effects on the derivation of SIs through grammar alone.

References

- Chierchia, Gennaro, Fox, Danny and Spector, Benjamin. "87. Scalar implicature as a grammatical phenomenon". Volume 3, edited by Claudia Maienborn, Klaus von Heusinger and Paul Portner, Berlin, Boston: De Gruyter Mouton, 2013, pp. 2297-2332. https://doi.org/10.1515/9783110253382.2297
- Geurts, Bart. 2011. Quantity implicatures. Cambridge: Cambridge University Press.
- Russell, Benjamin. "Against Grammatical Computation of Scalar Implicatures". Journal of Semantics 23:4, 2006, pp. 361-382. <u>https://doi.org/10.1093/jos/ffl008</u>
- Sauerland, Uli. "The Computation of Scalar Implicatures: Pragmatic, Lexical or Grammatical?" Lang. Linguistics Compass 6, 2012, pp. 36-49. https://doi.org/10.1002/lnc3.321

The Construction and Motivations for Immobile Verbs in German © 2024 by Simone Brown & Rachel

Mattie is licensed under <u>CC BY-ND 4.0</u>

The Construction and Motivations for Immobile Verbs in German

Simone Brown & Rachel Mattie

Department of Linguistics, McGill University LING 410: Structure of a Specific Language 1 Professor Bernhard Schwarz

Abstract

German is a complex language with unique syntactic characteristics. In this paper, one of these unique features will be examined, namely, verbs that resist typical patterns of movement. These verbs, called immobile complex verbs, combine a prefix with a verb to communicate additional meaning. Although rules governing the German language require these verbs to move, they remain in their underlying position. This paper will provide a summary of current linguistic discussions surrounding immobile verb structure in German, exploring several key theorists who have attempted to discern what differentiates immobile complex verbs from separable and inseparable verbs within German. Ultimately, this paper will argue that Haider's theories for complex and comparative immobile verbs are not successful in capturing the nature of this phenomenon. Vikner's theory for complex immobile verbs, which argues for the existence of the IP, as well as Meinunger and Freitag's theory for comparative immobile verbs, which capture the cross-linguistic applications of this verb immobility, provide more comprehensive theories of this (non) verb movement.

1 Introduction

German grammar presents linguists with many opportunities to expand linguistic theory regarding phrasal construction and movement rules. For German syntacticians, the language is both a blessing for the material it provides for study and a curse for the complex methodology necessary to explain phrasal grammaticality. Furthermore, these phenomena are not simply restricted to German; other languages contain similar constraints that allow linguists to form conclusions with wider cross-linguistic implications. Examining German immobile verbs contributes significantly to broader theories about the grammar constraints of phrasal structures and verb mobility. In this paper, immobile complex verbs will be discussed as one of Haider's arguments against the appearance of an Inflectional Phrase (IP) in German. Without the use of IPs in German, the language is effectively disregarding one of the elements believed to be a necessity for Universal Grammar, the overarching goal of many linguists to define a set of rules all languages adhere to. Furthermore, immobile verbs following comparative phrases imply an expansion of c-commanding and a reevaluation of S-structure realization. By studying German immobile verbs, linguists can then extrapolate these rules to other languages to mark patterns and possibly reconfigure previous assumptions about verb movement. Beyond their wider implications, immobile verbs (in both forms) pose an interesting dilemma as they operate outside the boundaries of normal verb movement. Furthermore, explanations for verb immobility typically entail the extension, or omission, of German grammar rules. For this reason, theories addressing immobile verbs must be careful to operate, as best they can, within the boundaries of German grammar rules.

This paper will summarize the various arguments posited by German linguists about the structures that prevent verbs from undergoing expected movement. The examination of immobile verbs will be divided into two categories; first, immobile complex verbs, which are constructed through affixation, and second, verbs restrained from movement in a comparative phrase. The focus will remain on immobile verbs as a parallel consequence of two different grammar processes, rather than as one single constraint applied across two different contexts.

Several concepts and terms will be referenced throughout this paper that should be defined. One such concept is the understanding of X' schema, where the building blocks of a sentence are assembled in ascending order from the most basic element to the most complex. The use of X denotes a placeholder where any word type, such as a noun or verb, can be substituted in to create a different type of phrase; this is known as the "head" of the phrase. There are three basic elements to each X' bar schema: The first denotes the most basic fundamental building block of a word or phrase, referred to as X0. Next is X', which denotes the level(s) intruding between X0 and XP, where additional information can be included by linking subordinate XPs. Finally, XP represents the entire phrasal unit and encompasses all the information that descends from that level. The same concept of building blocks explains this paper's references to N0, which marks the base from which more complex words are built by appending suffixes.

2 Immobile Complex Verbs

Complex verbs are constructed by the addition of prefixes onto a 'root' verb to communicate additional meaning. These complex verbs are differentiated by the types of prefixes they include; they are separable, inseparable, or immobile. Separable prefixes are a classification of prefixes that are separated from the 'root' verb when it undergoes movement, including prefixes such as '*auf*-' (meaning 'on'). These prefixes will appear where the verb was positioned in a sentence before it moved. Inseparable prefixes are the opposite of separable prefixes; they must remain attached to the verb as it undergoes movement and they will never appear in the base position after the verb has moved to the left. Prefixes such as '*ur*-' (meaning 'earlier, older or more primitive') that combine to the N0 are considered inseparable. The final type of complex verb is a combination of separable and inseparable prefixes. These

⁴⁵

verbs are typically referred to as "immobile" verbs due to their inability to undergo movement despite no visible obstructing factors. An example of an immobile verb would be '*ürauffuhren*,' meaning to 'put on (a play) for the very first time,' which combines both aforementioned prefixes to a verb base (Vikner, 2005).

German is similar to French in that it allows movement of all verbs between head positions, as opposed to English, which only allows the movement of auxiliary verbs to empty head positions. The theory of head movement states that when possible, head movement will occur within the constraints of a language (Murphy, 2021). Since German allows movement of all verbs, not just auxiliary verbs, the formation of complex sentences the verb would be expected to 'stop off' in the head of the IP (Inflectional Phrase) and leave a trace as it continues to the head of the Complementizer Phrase, henceforth referred to as the C head. This understanding of head movement is the motivation behind Haider's argument against the inclusion of an Inflectional Phrase in German X' theory, which will be discussed below.



Figure 1: Demonstration of basic verb movement based on Haider's argument against the inclusion of an Inflectional Phrase (IP). In this tree, movement leaves behind a 'trace' to occupy the I head in its movement upwards.

Immobile verbs are an interesting phenomenon that appear not only in German but in several other Germanic languages including Dutch and Swiss German (Vikner, 2005). These verbs, also called complex verbs, pose an interesting dilemma as they are an exception to the constraints established for verbs and their movement. Assuming that immobile verbs function differently, a way to categorize and understand them must be found as to understand why they are an exception to the otherwise universal rule in German that all verbs must undergo movement to a V2 position unless the C head is otherwise occupied (Murphy, 2021). Without this assumption, the entire structure established for German must be scrutinized, and grammatical constraints must be started from scratch. For some linguists, such as Haider,

the solution to this paradox is to introduce a new constraint that blocks the immobile verb's movement into C without affecting other types of verbs.

Haider (2010) argues that immobile verbs are the product of a verb conjoined with multiple prefixes that contradict each other. Returning to our previous example, '*urauffuhren*' contains both '*ur-*' which is inseparable and '*auf-*' which is separable. When a separable prefix occurs closer to the root than the inseparable prefix, it causes a paradox in which the verb must leave behind an embedded prefix while keeping the other (inseparable) prefix attached to the verb. Rather than violate the constraint on prefixes, the grammar instead chooses to violate the less important constraint that obligates verbs to move when they are presented with the opportunity (Haider, 2010). This understanding of German X' structure relies on a certain interpretation of immobile verbs as a combination of contrasting prefixes that figuratively 'weigh' the word down so as to prevent it from moving into C.

The ramifications of this theory are widespread. In Haider's construction of a Germanic X' structure, he controversially argues against the Inflectional Phrase (IP), based on the claim that the immobility of certain complex verbs negates its existence. This is due to the fact that if the IP were present in German, immobile verbs would be able to move, either in part or as a whole, to the I head in cases where it is previously empty. The immobile verb would then leave behind the separable prefix in the I position while carrying on to the C position with the inseparable prefix.



Figure 2: Demonstration of complex verb movement based on the inclusion of the Inflectional Phrase, where the separable prefix would be left behind in the I head. The star demonstrates this sentence is ungrammatical, arguing against this construction in German.

Haider provides several examples in his paper which categorically prove that this theory of movement is impossible. In sentences where the separable prefix has been left in the presumed location of I, even if there are no other violated constraints in German, the sentence will still be considered ungrammatical by a native German speaker (Haider, 2010). Under this interpretation of complex verbs, the IP cannot exist, as it would present an alternative solution to the paradox of verb prefixes that do not necessitate a restriction on verb mobility. This is Haider's argument as to the exclusion of an IP when constructing an X' structure for German, but considering alternative theories about immobile verbs, this entire argument is questionable. Haider only addresses certain examples of complex immobile verbs in his writing and leaves behind many other verbs which do not include the contrasting prefixes he deems necessary in explaining their immobility (Murphy, 2021). Without this theory of restrictions placed upon 'immobile' verbs, Haider lacks a clear piece of evidence used in his argument against IPs, meaning that the possibility of Inflectional Phrases must again be considered in German.

Alternative theories of the functionality and categorization of immobile verbs have been posited in an attempt to resolve the gaps in Haider's theory, or dismiss it entirely. Vikner (2002) criticizes Haider's theory as insufficient to address the breadth and variety of immobile verbs. Vikner lists many examples which lack the paradoxical prefixes that Haider associates with immobility and still are unable to undergo movement. This includes immobile verbs such as '*teilzahlen*', meaning 'to pay in installments' and '*ehebrechen*', meaning 'to break up a marriage or commit adultery' (Vikner, 2005). In both of these examples, the contradictory prefixes that Haider argues prevent these complex immobile verbs from movement are not present, yet they remain in their base position. Instead, Vikner proposes a new categorization of complex verbs which simultaneously expands the X' theory of German and other languages in which immobility occurs like Swiss German and Dutch. He differentiates between V0 and V*, which are similar to Haider's categorization of inseparable prefixes and separable prefixes, respectively. Just as X' theory designates different levels of constituent power, a V* is equivalent to a new constituent that is more powerful than V, but less powerful than V'. This permits a differentiation between inseparable verbs and separable verbs. Only V0 can undergo movement, meaning that within a V* only the V0 portion displays movement. This is demonstrated below:



Figure 3: Internal structure of a verb phrase based on Vikner's assertion of a V* between the V' and V, allowing the movement of the verb and inseparable suffixes while leaving behind separable prefixes.

According to Vikner (2005), immobile verbs differ from these two categorizations because they "are in the intersection" of complex V0 and complex V* verbs (Vikner, 2005, p. 101). Therefore, immobile verbs are unable to undergo movement to the C head, not directly due to the paradox of prefixes. He claims that since immobile verbs are typically composed of separable and inseparable prefixes, which usually indicate whether a complex verb is V0 or V*, it cannot undergo movement in any form. This alternative theory calls Haider's argument for the exclusion of inflectional phrases in German into question. Assuming this lack of movement is not due to an inability to separate the inseparable prefix from the root, verbs could reasonably move through the I to arrive in C without leaving behind a verbal trace.

Furthermore, Haider's theory does not acknowledge the group of complex verbs commonly referred to as pseudo-compounds. Pseudo-compounds are complex verbs created through the back-formation of a compound noun with prefixes. These verbs are similar in appearance to other complex verbs, but their resistance to movement makes their categorization controversial; if they were a single lexical unit they would be expected to move as one unit, and if they were multiple units conjoined together, the prefixes would be separable from the root. In comparison, Vikner's theory addresses these pseudo-compounds since their ambiguity is resolved by assuming these verbs have qualities of both V0 and V* (Vikner, 2005).

Vikner's theory is further affirmed in a study conducted by Chenchen Julio Song attempting to compare and contrast multiple languages. This study provides evidence in German for the expansion of the X' theory to accommodate these 'in-between' categories. Song (2016) remarks that inseparable prefixes function as X0 but that separable prefixes must exist above X0 but below X' in order to accommodate V2 movement, prefix topicalization, and inflection. He also describes inseparable prefixes as "more abstract than separable ones in a way similar to Type II [bound prefix + free base] and Type III

[free prefix + free base] prefixes in English" (Song, 2016, p. 4). An example of Type II in English would be 'disconnect' and an example of Type III would be 'overrun.' He further proposes that there are constraints on the number and forms of prefix attachment, which prevent mobility in instances where the complex verbs create a unit too large to move into the C0 position. Here, Song argues for the existence of a constituent higher than X, while acknowledging that verbs can be 'weighed down,' similar to Haider's contentions about the cause of immobility.

Yet another argument for understanding complex verbs is that particle verbs should be examined based on their morphological makeup rather than as a syntactically complex constituent. Previously, these complex verbs were thought to be restricted from movement, modification, or separation from the base verb, however, Müller (2002) argues that there are certain examples where these circumstances can occur and the sentence will remain grammatical. His argument is not that syntactic constraints are applied exceptionally to immobile verbs, but rather that there are other factors that are falsely attributed to a single grammatical rule. Two examples that Müller gives of these so-called 'other factors' are information structuring and contrast establishment. Information structuring dictates the ways in which a sentence must be structured so as to communicate the intended semantic meaning, occasionally at the expense of less important grammatical rules. Contrast establishment dictates a particle's ability to be modified based on semantic factors. If it were in fact a semantic concern, communicated through the morphological makeup of these complex verbs, then it would circumvent some syntactic constraints to ensure the sentence's grammaticality (Müller, 2002).

3 Comparative Phrase Immobile Verbs

Immobile verbs can also be created by German comparative phrases. In sentences where a comparative phrase modifies and exists before a verb in the sentence's base order, that verb will not complete the expected V to C movement that creates the V2 order. Meinunger (2006) gives a number of examples of these phrases, including "*mehr als*" (meaning 'more than'), "*so gut wie*" (meaning 'as good as') and others (55). In sentences with these phrases, the verb cannot move to C as usual, meaning sentences such as (1) "**Der Wert verdoppelte sich nun soviel wie*" are unacceptable (Meinunger, 2006, p. 55). At first glance, this appears nonsensical. In almost any other situation, the same verbs can move to the C position. This indicates that some property of the comparative phrase restricts movement, rather than some condition of the verb itself. Meinunger (2006), Haider (1997), and Freitag (2021) each present different explanations for this phenomenon.

For a syntactic explanation, Meinunger argues that immobility is due to the last word in the phrase. For example, '*als*' in "*mehr als*." According to his explanation, the blocking elements occur within C, and are therefore functional categories. He then explains how functional words are different from lexical words. Functional words, as they are not independently considered phonological words, need to become part of an actual phonological word, by having a phonological word come afterwards. Meinunger (2006) argues that the last word in the blocking phrase becomes a part of the phonological word that contains the verb, meaning that these two cannot be split up— putting the words in any other order creates a badly formed sentence. For example, in the sentence (2) "*Der Wert hat sich nun soviel wie verdoppelt*", '*soviel wie*' (meaning 'as much as') would be considered a lexical or semantic word, while '*wie verdoppelt*' (meaning 'as doubled') is considered a phonological word (Meinunger, 2006, p. 55). Because these three words are divided in this way, '*wie*' cannot be split from '*verdoppelt*' nor '*soviel*.' Furthermore, because these word boundaries are all assigned in base order, the verb cannot move and leave '*soviel wie*' behind. Therefore, Meinunger argues that this is more than just a problem of syntax.

Haider (1997), however, disagrees . He argues that the unacceptability of sentences like example 1 ("*Der Wert verdoppelte sich nun soviel wie") are due to a c-command constraint. In an acceptable sentence, such as "Der Wert hat sich nun soviel wie verdoppelt," the adverbial comparative phrase, 'soviel wie' is an adjunct to the VP containing 'verdoppelt', and c-commands it (Haider, 1997, p. 23). He argues that the comparative word, such as 'mehr,' in "mehr als", or 'soviel,' in "soviel wie" is the one that creates the constraint, which is the exact opposite of what Meinunger says. A sentence such as "*Der Wert verdoppelte sich nun soviel wie" breaks this constraint by having 'soviel' before the verb. Using Haider's rules, 'soviel wie' must c-command 'verdoppelte' at the surface structure. Surface structure, or Sstructure, is the sentence's final form, the way it is spoken, which can be different from the deep structure, or D-structure, where a sentence is originally generated. Words or phrases often move and change order between the D-structure and the S-structure. Moving the verb to the C position puts it before the comparative phrase, breaking the c-command constraint. Meinunger (2006), however, disagrees with this claim. He argues that in most other cases, there are no S-structure constraints on verb movement, a pillar of Haider's proposal. If Haider's theory were correct, previous conclusions drawn about constraints on movement derived from the S-structure in all other contexts would have to be reevaluated. Usually, the Dstructure is used to determine whether movement can be allowed, so accepting this theory would mean considering that the S-structure might have a greater impact on movement than previously thought. Also, Meinunger (2006) argues that there are other sentences with different constructions that this rule does not account for. For example, in other constructions, 'mehr' is permitted to go before the verb, which Haider's proposal would classify as ungrammatical. Also, there are other constructions that do not allow for verb

movement, which Haider's proposal does not acknowledge. While Haider's proposal is certainly more simple to explain and understand, it is a narrow explanation that does not include data from other sentence constructions. For this reason, Meinunger's theory seems advantageous, as it fits a wider variety of sentences.

Freitag (2021) provides a fresh perspective on this type of immobile verbs. He argues that this restriction on V2 movement is specific only to this type of comparison structure, which, just like Meinunger's theory, contradicts Haider (1997). Freitag demonstrates how the same type of sentences in different languages, including English and Italian, also restrict movement, showing that this is a problem that is not specific to German or to V2 movement (Freitag, 2021). In fact, he states that this pattern is not specific to verbs either, using the example "Dann war Christopher mehr als betrunken" (Freitag, 2021, p. 13). This sentence uses the comparative phrase 'mehr als' with an adjective, 'betrunken' (meaning 'drunk'). Freitag then shows that "*Betrunken war Christopher mehr als' ' is unacceptable, similar to previously discussed examples (Freitag, 2021, p.13). The widespread nature of this problem means that according to Freitag's analysis, explanations for these sentences' unacceptability cannot simply be due to a German-specific syntax constraint. Like Meinunger, Freitag views this as a problem beyond the scope of syntax. He states that movement blocking is due to the requirement of focus marking (Freitag, 2021). Freitag's contention is that the entire verb (or other constituents) must come after the comparative phrase in order for the focus and accent to be applied correctly, and for the semantic meaning to remain intact. Therefore, syntactic operations, such as V2 movement, would make sentences containing these comparative phrases unacceptable. Even verbs with particles, such as 'anschreien,' (particle 'an-' meaning 'at' and verb 'schreien' meaning 'to scream') which leave their particles behind after movement, cannot undergo V2 movement in a sentence with a comparative phrase, as the focus is then only on the particle, rather than the whole verb, which generates an incorrect semantic interpretation (Freitag, 2021, p. 15). This emphasizes the fact that any kind of movement that splits up or changes the order of the comparative phrase and the constituent it modifies (in this case, a verb) will create an unacceptable sentence. This does, however, permit the movement of the entire phrase that contains both the comparative phrase and the phrase it modifies.

Each of these three perspectives provides a different explanation for the unacceptability of V2 verb movement in sentences with comparative phrases. Haider's explanation is the most simple, concerned only with syntax. Having a constraint that requires the comparative word to c-command the VP that it modifies solves the problem in a direct and easily digestible manner. However, this theory was amended by Freitag, and directly refuted by Meinunger. While it may apply to isolated cases, it is hard to accept a constraint that cannot be generalized to the entire German language when more data is collected.

Therefore, both Freitag and Meinunger's theories seem to be more likely, despite their intricacies. Freitag's claim can be generalized not only in German, but cross-linguistically. His analysis that the focus marker must come after the comparative phrase and include the whole verb has remnants of Haider's theory, in that it requires the phrase being modified to come after the comparative phrase (and essentially be c-commanded by the comparative phrase). However, because he is concerned with comparative phrases, rather than comparative words like 'mehr,' Freitag's theory accounts for Meinunger's examples where the comparative word, when not a member of a comparative phrase, occurs after the phrase it modifies. As it does not only involve syntax, this contention also satisfies Meinunger's criticism of Haider's theory, that S-structure does not affect verb movement in other cases. Since this is a case where the focus assignment affects the acceptable syntax, rather than the other way around, the acceptability is therefore not determined by syntactic structures. Meinunger's analysis bears similarities to Freitag's analysis. Although a different mechanism, both are examples of syntax interacting with other aspects of the sentence, in this case, phonology. Meinunger successfully satisfies his own criticisms of Haider's theory, while creating a theory that accounts for specific cases of the immobile verbs following comparative phrases. He does so without creating a theory that has unwanted, incorrect implications for other sentence constructions.

4 Conclusion

Despite possessing similar surface level attributes, the two types of verbs discussed in this paper have very different causes for their inability to perform the V to C movement. Neither has a firm, agreedupon theory as to why they behave in this way, but each suggested theory can provide a little bit of insight into the German language.

In the first section, several theories were proposed for complex immobile verbs. Haider's theory, that verbs with multiple prefixes are immobile due to the lack of an IP in German would require many to rethink their understanding of German syntax. However, Haider's theory lacks thoroughness. In contrast, Vikner's theory, which would allow for the existence of the IP, emerges as the more likely theory.

In comparative phrase immobile verbs, Haider proposed a c-command restraint that would have a major implication on the way comparative words such as *'mehr'* could be used in a sentence. However, once again, he failed to consider other examples outside of those that worked for his theory, which allowed Meinunger and Freitag to propose their own theories, which directly refute Haider's contentions. Both of their theories allow for application not only in specific examples covered, but also throughout other German constructions, and in the case of Freitag's argument, even throughout other languages.

In order to know the cause of these immobile verbs, more research must be conducted with an expanded data set. However, in the meantime, the theories that have been proposed allow for the exploration of this topic and other related topics in the structure of German. Each theory helps to give a possible answer to a complicated problem and creates an area for future research to pursue, leading to a greater understanding of the German language.

References

- Freitag, C. (2021). The Syntax of V2. In *Verb-second as a reconstruction phenomenon: Evidence from grammar and processing* (pp. 7–63). Essay, De Gruyter.
- Haider, H. (1997). Typological implications of a directionality constraint on projections. *Linguistik Aktuell/Linguistics Today*, 17–33. <u>https://doi.org/10.1075/la.13.02hai</u>

Haider, H. (2010). The syntax of German. Cambridge: Cambridge University Press.

Meinunger, André. (2006). Interface restrictions on verb second. *The Linguistic Review* 23.127–160. Müller, S. (2002). Syntax or morphology: German particle verbs revisited. *Verb-Particle*

Explorations, 119–140. https://doi.org/10.1515/9783110902341.119.

- Murphy, A. (2021). Rightward verb movement: A reappraisal. *The Journal of Comparative Germanic Linguistics*, 24(3), 319–376. <u>https://doi.org/10.1007/s10828-021-09130-8</u>.
- Song, C. (2016). A Minimalist Study of Complex Verb Formation: Cross-linguistic Patterns and Variation. University of Cambridge. https://www.juliosong.com/doc/Song2016FYR.pdf.
- Vikner, S. (2005). Immobile complex verbs in germanic. *The Journal of Comparative Germanic Linguistics*, 8(1-2), 85–117. https://doi.org/10.1007/s10828-004-0726-9.

Effects of Community Involvement and Age on Nativization: A Sociolinguistic Analysis of Jewish

Communities © 2024 by Ella Brown is licensed under <u>CC BY-ND 4.0</u>

Effects of Community Involvement and Age on Nativization: A Sociolinguistic Analysis of Jewish Communities

Ella Brown

Department of Linguistics, McGill University LING 521: Sociolinguistics 2 Professor Charles Boberg

Abstract

This study examines age and community involvement of Jewish people in North America as independent variables affecting the nativization of two phonemes, (χ) and (i) when using Hebrew loanwords in English or when reading out an Aramaic passage. Participants in this study filled out demographic questionnaires, listened to recordings of differing pronunciation of the variables in question to identify which one is closer to their own pronunciation, and finally recorded themselves saying target words. The (i) variable was found to be an indicator of change in progress towards a more Hebraic pronunciation. The (χ) variable is salient and nativization to /h/ is stigmatized, so it is highly sensitive to community involvement.

1 Introduction

Growing up in a Jewish community, I heard many animated discussions of the correct English transliteration of Hebrew words that include the graphemes *chet* (π) and *chaf* (\Box), phonetically best represented by / χ /. While some believe that "h" is the most understandable and accessible spelling, others prefer "ch" or "kh" to indicate a phonemic contrast with the sound /h/, represented orthographically in Hebrew by the letter *hey* (π). Regardless of which orthographical variant each person selects, the reality is that these linguistic choices are intertwined with the social identity of North American English-speaking Jews. This orthographical debate represents something deeper – English speakers' variable pronunciations of the (χ) phoneme appear to indicate even more intraspeaker variation than their spellings.

Through anecdotal observation, I have noticed that more religiously observant and involved Jews tend to use pronunciations that are more closely aligned with Hebrew pronunciations, whereas Jews who are disaffiliated tend to further anglicize their pronunciation. This study examines age as well as the level of Jewish involvement and affiliation as social variables in North American Jewish communities and their impact on the nativization of both the linguistic variable (χ) in words that are spelled with the Hebrew letters "chet" and "chaf" and on the linguistic variable (i) as it appears in closed syllables preceding alveolar obstruents in an Aramaic text.

The Linguistic Variables

This variability of (χ) as it is articulated as /h/⁵ or / χ / in Hebrew and Aramaic borrowings is consciously salient for most monolingual English speakers. In a perceptual task as part of this survey, participants had the opportunity to hear two pronunciations of each of five different Hebrew loanwords, and then had to select the variant they believed was most similar to their own pronunciation or select that they could not detect a difference. Among all participants and all words, there were only two instances of participants saying that they could not hear a difference between the pronunciation with /h/ and that with / χ /. Importantly, this is a phonemic contrast in Hebrew. While most speakers can articulate / χ /, many find it more difficult and will instead articulate /h/, especially in rapid and casual speech.⁶ While the former variant is aligned with a more Hebraic pronunciation, the latter variant is nativized to be more pronounceable within the phonology of English. Given that this variable is salient and phonemic, the / χ / variant is perceived by many as a more educated and correct pronunciation. Accordingly, the /h/ variant is stigmatized.

In a preliminary sociolinguistic interview with one middle-aged woman – a leader in her synagogue community, but not especially religious, and not raised in a particularly observant community – I played two audio recordings of English speakers pronouncing "challah," the braided bread traditionally eaten on the sabbath. In the first recording, the word was pronounced with $/\chi$ / and in the second, /h/. I asked the interviewee what pronunciation was most similar to her own, and she quickly decided that the former was most aligned with her pronunciation, even as she proceeded to refer to the bread as /'ha_la/. Labov calls this phenomenon "linguistic insecurity," explaining that insecure speakers consciously strive for correctness, shifting toward prestige variants even in middle age (1972:117). This woman's assertion reflects her attitude towards the variable (χ) and her linguistic insecurity regarding her use of a stigmatized variant. In another interview, one speaker, a *Rebbetzin* (Rabbi's spouse) explained that she "would pronounce 'Chanukah' as / χ anuka/ but 'Hanukkah' as /hanuka/." Her style-shifting isn't unique – Sarah Benor's paper on Jewish linguistic distinctness suggested that Jews are more likely to use

⁵ Occasionally (χ) \rightarrow Ø word-finally, in the binary code used, this was valued the same as /h/ because of its status as "full anglicization"

⁶ In my experience teaching Hebrew to students aged 8-12

loanwords from Hebrew and Yiddish when talking with other Jews and about Jewish topics than with non-Jews (2009).

With regards to the (i) variable, this vowel is represented by the Hebrew *nikud* (diacritic) called *chirik*. It should remain a high tense front vowel when pronounced in accordance with Hebrew and Aramaic phonology (Blau 2010). In closed syllables, English speakers have a tendency to lax this vowel, articulating the variable instead as something most akin to /1/. In English, short lax vowels cannot occur in open syllables, and tense nuclei often glide in closed syllables. Given this, maintaining (i) as a tense vowel may feel unnatural for some speakers, prompting them to lax it in the closed syllables examined here. Among many tokens of (i), this study only examines the thirteen tokens of (i) elicited per speaker that occurred in closed syllables that were either /j_s/ or /j_t/. In contrast to the Hebraic pronunciation, Yiddish does not have the phoneme [i] and instead has [I] (Kleine 2003:263). While some speakers who produce /1/ may be influenced by Yiddish, for most English speaking Jews this influence is minimal and indirect since children are taught to read Hebrew rather than Yiddish. Additionally, the laxing of (i) and the weakening of (χ) to /h/ are both phonological processes that may be easier for English speakers, so (i)-laxing will be considered hereafter as nativization.

The laxing of (i) does not appear to be stigmatized, likely because this variable appears to be beneath the level of conscious awareness. In another preliminary sociolinguistic interview, I instructed a highly involved young woman who worked as a Hebrew instructor at a Jewish summer camp to recite an Aramaic passage the way her campers would, with a strong American accent. Her recitation included many features of English phonology – she epenthesized to re-syllabify words that did not fit English syllable rules, pronounced all instances of (χ) as /h/, even giggling at the ridiculousness of the task as she did so – but she maintained all instances of (i) as tense vowels. In this way, the variable (i) is best classified as a Labovian indicator, given that it appears not to be subject to style-shifting and so it can be considered non-salient to speakers (Labov 1972:178). Instead, it appears to be a function of group membership. When I demonstrated the difference between the tense and lax pronunciation in an explicit minimal pair format and asked her explicitly how she felt about the lax pronunciation, she announced with disgust that she "would never say *that*." While the variable may not be salient to speakers in their own pronunciation, this informant's reaction indicates that, when explicit, the variable does hold negative prestige for fluent Hebrew speakers.

The Social Variables

Jewish involvement and affiliation as social variables can be understood similarly to the loose simplex and dense multiplex networks discussed in Milroy and Milroy's 1977 examination of Belfast.

Jews who are more religiously observant tend to be more involved in Jewish social life, and vice versa (Mitchell 2020). Among Jews who are religiously observant, there are significant social implications. Their extensive adherence to Jewish law obligates them to live in a Jewish community where they can access the resources to practice their tradition, including Jewish schools and kosher grocery stores. The synagogue becomes a center of social and religious life. This facilitates the creation of a dense, multiplex social network where speakers may be more likely to prioritize linguistic variants that demonstrate community belonging and to be less influenced by the phonological processes of the English vernacular.

In contrast, speakers who do not practice stringent adherence to Jewish law do not need to live walking distance to a synagogue, shop only at Jewish kosher supermarkets, or attend Jewish schools. This creates a much looser social network with more significant interaction with speakers of other dialects. Notably, speakers who are less religious and less socially involved typically have less exposure to Hebrew and Aramaic, perhaps making them more likely to nativize (i) and (χ) to /I/ and /h/.

Also examined was age as a social variable, looking at the speech of young adults in their early to mid-twenties, as well as their parents' generation. Among the speakers in this study, age was pretty evenly distributed, with nearly equal groups of older and younger speakers. In terms of religious involvement and practice, none were Hasidic or strictly religious. All were my personal friends or friends-of-friends, meaning that although they all had varying levels of involvement and religious practice, they were all at least somewhat integrated into the general American or Canadian cultures of the cities in which they lived. All were native English speakers, and most were monolingual. In this way, social involvement and Jewish community practices can be interpreted as a way to measure a speaker's familiarity with Hebrew and Aramaic as Jewish contact languages and the extent to which they employ that familiarity in their nativization of foreign phonemes and application of English phonological processes. Some older Ashkenazi Jews (from the Baby Boomer generation or older, with Yiddish speaking parents) and Hasidim have more extensive influence from Yiddish, so among this population, producing lax /i/ would not necessarily be indicative of (a lack of) familiarity of Hebrew or Aramaic. However, these populations were not examined in this study.

The Jewish community is linguistically interesting given that nativization as well as other linguistic patterns may not follow the typical patterns found in previous research of other communities. While the dense multiplex networks described by Milroy and Milroy are most typical of working class communities, Jewish communities are comparatively quite well off, with 79% of non-Orthodox Jews reporting middle-class or higher family incomes (Mitchell 2021). Despite the fact that dense multiplex networks are not as typical in other well-to-do communities, the religious and cultural identity of Jews creates rather homogenous communities, which helps maintain linguistic distinctness in English. Previous research has found less raising of /æ/ before nasals in Jewish communities compared with the general

population, more released /t/ word-finally, and even syntactical differences (Benor and Cohen 2011). Given the distinct features of Jewish English, it is especially interesting to examine intragroup variation.

In previous research on Canadians' nativization of the foreign (a) phoneme, young Canadians appeared to be leading the trend towards the more American (a) patterning rather than /æ/, although it was lexically selective. With /æ/ being considered the more "native" English vowel, in the word *Iraq*, it could be that the /æ/ pronunciation "...is associated in Canada as much as in the United States with nativistic, anti-foreign attitudes and therefore rejected by those with a politically liberal outlook, despite neutral or negative feelings toward American culture" (Boberg 2020). In a similar way, the nativization of Aramaic loanwords can be associated with speakers' identification with Judaism, the Jewish community and their position within it.

Kroch explained that one of the main characteristics of "prestige dialects" includes the "blocking" of "phoneme assimilation" (1978:28-29). This aversion to full nativization by emphasizing one's knowledge of foreign language and culture allows speakers to indicate their social values and level of education, with variants closer to the original foreign pronunciation having greater prestige (Boberg 2020). In this population, the phenomenon of nativization is unique given that this is a language and culture that, at the same time, is both foreign and familiar. While these speakers grew up hearing these words, they do not use them in their vernacular speech as they are not fluent speakers of Hebrew or Aramaic. As such, a speaker's social identity will impact their categorization of the word as something more foreign or more familiar.

I hypothesized that Jews who are more involved in Jewish life are less likely to nativize their speech towards the more anglophone variants, with lower rates of /h/-pronunciation and (i)-laxing. Additionally, I hypothesized that younger speakers will also be more likely to articulate (χ) as / χ / compared with older speakers favouring /h/ given that there is evidence of a shift towards the more prestigious Hebraic pronunciation. Since I believe (i)-laxing to be below the level of conscious awareness, I suspect this will be more correlated with lower Jewish involvement rather than age.

2 Method

I solicited data from friends and friends-of-friends via a three-part survey, which I shared with Facebook groups and email lists for alumni of a Jewish summer camp, a Chicagoland synagogue and Jewish book club, and a Montreal Jewish community. All participants in this study were involved enough in the Jewish community to have joined virtual Jewish groups, seen the post, and self-selected to complete the survey. The majority of participants were based in the Midwest (see tables 1-3 for demographic information).

Table 1: Regional Demographics

Regional Background	Number of Speakers
Inland North Midwestern (Chicago Area, parts of NW Indiana, Milwaukee)	30
Other (Montreal, Vancouver, Toronto, New York, Bay Area, Pennsylvania)	10

Table 2: Age Demographics

Age	Number of Speakers
Over 30	18
Under 30	22

Table 3: Gender Demographics

Gender	Number of Speakers
Men	17
Women	23

In future research, it would be interesting to see if there are significant differences between Canadian and American Jews. However, Benor reports that, to a large extent, Jewish speech communities often have displaced dialectalism, tending to have variants typical of New York and not participating in some local sound changes (2009). This, combined with the small sample size solidified my decision not to examine regional influence at this time.

While the electronic survey methodology made it easier to collect data, it was not a random sample and not free of bias. This paper examines data from 40 participants, although there were nearly twice as many survey respondents. Yet, only 40 participants provided complete data for all three parts of the survey, likely due to technological difficulties and confusion, especially in the second and third parts of the survey – a perceptual task and audio recording.

The first part of the survey included demographic questions and questions about Jewish practice. Answers to these questions informed the study of the social variables discussed above. With the survey responses, I created an *Index of Jewish Involvement*, and was able to give respondents a numerical score to facilitate comparing speakers. I developed this index following the research tradition of Eckert and Benor and enlightened by the findings of the 2020 Pew study on Jewish American Identity, as well as through my own knowledge of these specific Jewish communities acquired through my own experience as a participant-observer.

The Pew study found that "people who are highly observant by traditional measures [...] also tend to report the highest participation rates in [...] cultural Jewish activities [...] Those who are low on the scale of traditional religious observance, meanwhile, tend to be much less active" (Mitchell 2020). Given this finding, it did not make sense to separate religious and cultural practices. The *Index of Jewish Involvement* (hereafter referred to as IJI) examines these factors simultaneously.

Also, I wanted to account for the fact that many people do not have the same Jewish practices now as they did as children. However, the childhood "critical age" is when speech forms are acquired and learned. Accordingly, I created a separate *Index of Childhood Practice* (referred to as ICP). For the ICP, I gave speakers a numerical score out of four, assigning one point for each of the following: having a b'nai mitzvah ceremony, attending a Jewish private school, going to a Jewish summer camp, and partaking in Jewish youth group programming. I selected these as variables to include in the ICP, separate from the IJI, because they are not ongoing activities in adulthood. Additionally, while Jewish practice in adulthood is one's personal choice and may be reflective of one's orientation and attitudes towards their religion and culture, children did not choose to take part in these social and educational experiences. Although their participation in these experiences may not have been voluntary, these activities include significant Jewish learning and time socializing within a Jewish network, and increased ease in decoding or speaking Hebrew.

The IJI was based on Sarah Bunin Benor and Steven Cohen's findings in their Survey of American Jewish Language. In their survey, they examined generational cohort, religiosity through synagogue attendance, the proportion and engagement of Jewish friends within social networks, visits to Israel and time spent there, and Orthodox identity as independent social variables (Benor and Cohen 2011).

The IJI is calculated with respect to six factors and scored out of ten as is illustrated in table 4, below. In keeping with Benor and Cohen's work, I examined religiosity with a slightly higher weight on Conservative and Orthodox denominational identity as well as synagogue attendance and Shabbat observance, as I believed that this would most differentiate speakers in my sample. In addition to asking about Shabbat observance in question one ("Do you observe Shabbat?") I added a second question ("Do you handle money on Shabbat?") to differentiate those who observe Shabbat casually with those who keep Shabbat in a strict sense, given that handling money is prohibited by *halacha* (Jewish law). I used the fifth and sixth questions ("Do you work for a Jewish organization?" and "How many of your friends are Jewish?") to measure social networks. While Benor and Cohen did not look at occupation, those who work for Jewish organizations have more contact with other Jews even if they are not close friends with

their coworkers, so I thought that this was valuable to include as part of information about social networks. Additionally, those who are employed as members of clergy have had extensive schooling and training in Jewish texts and liturgy as well as immersing themselves in Jewish social spheres, which is why working as a member of clergy was weighted more heavily than working in another Jewish job.

Index of Jewish Involvement		
Do you observe Shabbat?	+0 no, +1 sometimes, +2 yes	
Do you handle money on Shabbat?	+0 yes, +1 no	
How often do you attend synagogue or participate in	+0 rarely, +1 roughly monthly, +2 weekly or more	
other Jewish spaces?		
Which of the following best describes your	+0 Reform, Reconstructionist, Humanistic; +1	
denominational identity?	Conservative, Orthodox, Traditional	
Do you work for a Jewish organization?	+0 no, +1 yes but not as clergy, +2 clergy	
How many of your friends are Jewish?	+0 some, +1 roughly half, +2 all/mostly Jewish, +2	
	all/mostly Jewish	

Table 4: Summary of the Index of Jewish Involvement

While Israel visits proved to be significant in Benor and Cohen's 2011 study as well as in the 2020 Pew study, I did not ask about this variable given my belief that attitudes towards Israel are changing, especially among liberal Jews such as those in the Facebook groups where I posted the survey. Many of these Jews are trying to define their Jewish identity and practice without centering Israel and trying to speak Hebrew as a Jewish language rather than as a way to connect with the state of Israel⁷. Within this sample, Israel visits would not be uniformly indicative of the strength of the speaker's involvement. Additionally, I worried that asking about Israel would politicize the survey, leading to fewer respondents.

While the IJI allows for a score between zero and ten, the highest score in this sample was an eight and the lowest score was one. This is largely due to the limitations of the sampling method used. It may also be related to the audio file that was solicited as part of the second portion of the survey, which asked speakers to record their recitation of the Mourner's Kaddish (also called *Kaddish* or *Kaddish Yatom*), a prayer traditionally said by a child mourning the loss of a parent, but also used widely as a general prayer for mourning. Although many Orthodox Jews would be uncomfortable reciting and recording for non-religious purposes and therefore unfortunately excluded from the survey, this passage

⁷ This decision was largely informed by my personal experiences as a participant-observer in these communities with many friends, acquaintances, and clergy asserting a non-Zionist position with respect to the Hebrew language and Jewish involvement.

was optimal for several reasons. Firstly, although it must be acknowledged that sociolinguistic interviews are wonderful for their ability to elicit more natural speech in several different styles, the nativization of these linguistic variables would be largely inaccessible due to their absence in casual speech. Given this, asking speakers to record a passage was the most efficient way to gather uniform elicitations. Additionally, this passage was ideal given that it is spoken and not sung, and has several tokens of the relevant linguistic variables.

With regards to style, the passage cannot easily be categorized on the Labovian speech style continuum of casual speech and reading style. Given that the passage is a prayer for the dead, speakers may feel emotional when reciting this passage, causing them to be less cognizant of their speech. On the other hand, the fact that speakers are prompted to read and record the passage may heighten their awareness of their pronunciation. That being said, many speakers do have this passage memorized and have familiarity reciting it from childhood, which suggests that perhaps it could be better classified as a "childhood rhyme," and a substyle of casual speech (Labov 1972:70-109). Perhaps though, this illuminates the need for more or different classifications of speech style, with "prayer style" being considered as distinct.

Another important factor in the selection of this passage was the language. While most Jewish prayers are written in Hebrew, this prayer is in Aramaic. Given that no respondents are speakers of Aramaic and all speakers are native English speakers, all speakers were inclined to nativize the phonology of their speech. That being said, while Aramaic is distinct from Hebrew, this study is built on the assumption detailed by Shelomo Morag in his book, *The Vocalization Systems of Arabic, Hebrew, and Aramaic*: Aramaic and Hebrew share the same phonology and allophones, and that all Aramaic vowel phonemes can be denoted by the vocalization systems of Hebrew. Morag elaborates, observing that "it appears that the vocalizers did not encounter any difficulty in applying the same vocalization systems to both Hebrew and Aramaic, in spite of the phonetic and phonemic differences between the two languages. This is to be explained by the fact that the vowel phonemes of Hebrew and Aramaic differ from each other mainly in their distribution..." (Morag 2019:45-47). Respondents who pronounced the passage as if it were Hebrew were not nativizing the Aramaic passage to L1 Hebrew, but rather attempting to minimize their nativization to English phonology compared with those who used the variants /h/ and /t/.

The final part of the survey was a perceptual task. Using the web based platform Edpuzzle, speakers heard two pronunciations recorded by myself, the researcher, of each of five different loanwords and phrases: chag sameach, challah, chai, Hanukkah, and Pesach⁵. They then were able to respond by

selecting the variant that most aligned with their own pronunciation. The goal of this perceptual activity was to better understand the salience of the $/\chi/$ variable.

After collecting all survey data, I calculated scores on the IJI and ICP for each speaker. I then used a binary coding system to analyze all variables. All instances of laxed-(i) were coded as 0 while tense /i/ was coded as 1; all instances of /h/ were coded as 0 and / χ / was coded as 1. These numerical values were assigned impressionistically while the audio was played at reduced speed. After adding all the tokens together, I calculated a percentage of total tokens that were articulated in accordance with the Hebraic forms /i/ and / χ /, respectively. I also calculated a variable I called *Non-Anglicized Tokens* or NAT, which was the sum of the total realizations of /i/ with the total realizations of / χ /. Then, I was able to compare the average percentages of tokens articulated in the target form with the scores on the indexes and age to understand the way the social and linguistic variables interact.

To determine the significance of the findings, I used a two-tailed t-test with a p-value of 0.05. All data presented below was found to be statistically significant with p<0.05. I also ran a linear regression model looking at the NAT with the explanatory variables of age, IJI, and ICP.

3 Results & Discussion

As predicted, differences between men and women were found to be null as determined by a ttest. This is likely because the differences are superseded by the effects of involvement. In more liberal Jewish communities, such as the ones surveyed, differences in Jewish education and practice are negligible for men and women, who pray together. In Jewish communities that adhere to traditional gender roles, these roles are enforced by strict interpretations of Jewish law, which occur in every aspect of community life, meaning that these communities are exceptionally isolated from the larger anglophone community. Having received a Jewish education from a young age, members of these traditional Jewish communities have been exposed to lots of Hebrew and Aramaic, making them likely to be categorical users of $/\chi$ and avoid (i)-laxing. While this variable showed differences to be negligible, it is likely that significant gender differences could be found when examining a different variable.

As was expected through anecdotal observation and with consideration of previous research, Jewish involvement proved to be a highly significant predictor of both $/\chi$ / and /i/ realizations. Speakers who were more involved in Jewish life, as indicated by their higher index scores, were less likely to "fully" nativize their speech, avoiding both placeless /h/ articulations and (i)-laxing.

Figure 1	8: Jewish	Involvement	: and /χ/	Realizations
----------	-----------	-------------	-----------	--------------

Number of Speakers	Jewish Involvement by Indexical	Average Percent of Tokens Realized
	Score IJI	as /χ/
22	More involved, 4-8	85.0%
18	Less involved, 0-3	59.1%

As demonstrated in Figure 1, above, speakers who had relatively high scores realized significantly more tokens as $\chi/$ compared with their less engaged peers. While this pattern held for /i/ realizations compared with the /I/ variant (see Figure 2 below), overall there were higher rates of $\chi/$ realizations, perhaps due to the fact that this variable was more salient to speakers.

Figure 2⁹: Jewish Involvement and/ i/ Realizations

Number of Speakers	Jewish Involvement by Indexical	Average Percent of Tokens Realized
	Score IJI	as /i/
22	More involved, 4-8	82.2%
18	Less involved, 0-3	57.7%

The effects of the ICP, as illustrated below in figure 3, were similar to the effects of the IJI. While higher involvement levels in childhood practice were associated with a higher percentage of NAT, these ICP and IJI scores do not appear redundant.

Number of Speakers	Childhood Involvement by	NAT (Average Percent of
	Indexical Score ICP	Tokens realized as /i/ & / χ /
		rather than /I/ or /h/)
26	More involved, 3-4	85.0%
14	Less involved, 1-2	47.8%

Figure 3^{10} Childhood Involvement and /i/ and / χ / rather than / $_I$ / or /h/ realizations

⁸ The difference between speakers with indexical scores of 0-3 vs. speakers with scores 4-8 was found to be significant at p = 0.0004

 $^{^{9}}$ The difference between speakers with indexical scores of 0-3 vs. speakers with scores 4-8 was found to be significant at p=0.0489

 $^{^{10}}$ The difference between speakers with childhood indexical scores of 1-2 vs. speakers with scores 3-4 was found to be significant at p = 0.0005

With regards to the average NAT percentage when compared with both indexical social variables, the average percentage in the aggregate appeared highly significant, as shown in the graphs below.









Despite the large significance in average values when aggregated, there was a large standard deviation in the sample, meaning that there was minimal significance when examining outside the aggregate, suggesting the need for further research with a larger sample.

With regard to the salience of $/\chi/$, speakers' intuitions about their usage appear relatively accurate. Although no speaker reported exclusive usage of /h/ in the perceptual task, one speaker did have categorical /h/ used in their recording of the passage. Among speakers who self-reported variable usage of / χ / in the perceptual task, they recorded significantly fewer tokens of (χ) as / χ / than those who selfreported exclusive usage of / χ /. However, even the self-declared categorical users of / χ / articulated about 12% of tokens as /h/ (see figure 4, below).

Number of Speakers	Self-reported use of $/\chi/$	Average Percent of Tokens
		Realized as /χ/
24	Speakers who report variable usage	62.5%
	of χ and h	
16	Speakers who report exclusive	87.9%
	usage of /x/	

Figure 4¹¹

This could be because $/\chi/$ was highly salient perceptually and carries a strong positive prestige, so speakers were more likely to report using $/\chi/$ even if they do not actually use it. Alternatively, this could be due to a difference in contextual styles. Nearly all speakers reported pronouncing *chai*, a Hebrew word meaning "life," with $/\chi/$, even if they reported pronouncing *Hanukkah* with /h/. This could be because *chai* would be most typically used when speaking with Jewish interlocutors, whereas *Hanukkah* has made its way into mainstream English vernacular. In addition to further research on the variability of these phonemes between different types of borrowings, further research is needed to determine if $/\chi/$ is more common in Hebrew loanwords than in an Aramaic prayer passage.

With respect to age, there were no significant results for the (χ) variable. This appeared to be significant only for (i). This is likely due to the status of each variable, as speakers are more consciously aware of (χ) and hold pragmatic values for speakers, whereas (i) is more variable. In the sample surveyed here, the older group was skewed as it included a high proportion of Rabbis, who are highly involved in the Jewish community, having completed many years of education. To really examine the effect of age without the confounding effect of involvement and religiosity, I excluded all Rabbis from the analysis of age and (i). In the graph below, it is clear that among this group, there is no real correlation between age and involvement, illustrating that the distinction is needed.

¹¹ The difference between speakers who self-report variable usage of /h/ vs. / χ / compared with speakers who report exclusive use of / χ / was found statistically significant at p=0.002

Graph 3: Age and IJI



Figure 5 below illustrates that speakers over thirty realized significantly more tokens of I/I than the younger speakers.

Number of Speakers	Age	Average Percent of Tokens
		Realized as /i/
13	Older, Age > 30	49.1%
22	Younger, Age < 30	78.7%

Figure 5¹² Age and /i/ realizations, excluding Rabbis

With the (i) variable, older speakers were much more likely to lax their vowel. This suggests that it is not merely involvement and knowledge that impacts the laxing of (i), and that this could be indicative of a change in progress towards a more Hebraic variant rather than full nativization.

Overall, in this sample, the IJI proved to be the most significant in a linear regression model. The effects of age and ICP were also significant, but the large standard deviation and small sample size means that further research is required to better understand the relationships between these social variables and the linguistic variables.

¹² The difference between speakers over the age of 30 vs. speakers under the age of 30, excluding Rabbis, was found to be significant at p = 0.0429

4 Conclusion

While this study illuminated some intra speaker variation, further research is needed to better understand the larger picture. It would be interesting to examine the nativization of borrowings in other contextual styles. It would be especially interesting to further examine the scope of the (i)-variable among a highly involved and religious sample who had more extensive contact with Yiddish. Likewise, it would be helpful to examine age in a larger sample with a more representative number of Rabbis. Alternatively, it would be interesting to compare the speech of clergy members with the general population.

Despite its limitations, the use of "prayer style" was quite helpful to compare the same passage with different speakers and it would be interesting to employ this cross-linguistically in other research beyond just examining the nativization of Aramaic. Recordings of prayers are also available on youtube and other internet sources, and we can speculate demographic information with some certainty by examining the channels that publish such recordings.

Similar to the nativization of foreign (a), with speakers' social attitudes influencing their linguistic choices, speakers' relationship with their Jewish faith and culture was predictive of their pronunciation of both linguistic variables examined. Contrary to my initial hypothesis that (χ) would be significantly impacted by age, it was not. Age was, however, significant for the (i) variable, suggesting a potential change in progress. Younger people lead the trend towards the more Hebraic pronunciation, similarly to the way younger people lead the Canadian shift towards the more American production of (a).

References

Benor, Sarah Bunin. (2009). "Do American Jews Speak a "Jewish Language"?: A Model of Jewish

Linguistic Distinctiveness." *Jewish Quarterly Review* 99, no. 2 (2009): 230-269. doi:10.1353/jqr.0.0046.

- Benor, Sarah Bunin and Cohen, Steven M. (2011). "Talking Jewish: The 'Ethnic English of American Jews" *Ethnicity and beyond: Theories and Dilemmas of Jewish Group Demarcation*. Hebrew Union College-Jewish Institute of Religion
- Blau, Joshua. (2010). Phonology and Morphology of Biblical Hebrew : An Introduction. Linguistic Studies in Ancient West Semitic, 2. Winona Lake, Ind.: Eisenbrauns.

Boberg, C. (2020). Foreign (a) in North American English: Variation and Change in Loan Phonology.

Journal of English Linguistics, 48(1), 31-71. https://doi.org/10.1177/0075424219896397

- Kleine, Ane. (2003). "Standard Yiddish." *Journal of the International Phonetic Association*, 33(2), 261–265. doi:10.1017/S0025100303001385
- Kroch, Anthony S. (1978) "Toward a Theory of Social Dialect Variation." *Language in Society* 7, no. 1 (1978): 17–36. <u>https://doi.org/10.1017/s0047404500005315</u>.
- Labov, William. *Sociolinguistic Patterns*. Philadelphia, Pennsylvania: University of Pennsylvania, 1972.
- Milroy, J. and L. Milroy 1978. Belfast: Change and Variation in an Urban Vernacular. In P. Trudgill

(ed.), Sociolinguistic Patterns in British English. London: Edward Arnold, 19-36.

Mitchell, Travis. (2021). "Economics and Well-Being among U.S. Jews." Pew Research Center's

Religion & Public Life Project. Pew Research Center, April 5, 2023. https://www.pewresearch.org/religion/2021/05/11/economics-and-well-being-among u-s-jew/

My Jewish Learning. "Text of the Mourner's Kaddish." My Jewish Learning, November 2, 2022.

https://www.myjewishlearning.com/article/text-of-the-mourners-kaddish/.

Mitchell, Travis. (2020). "Jewish Americans in 2020." Pew Research Center's Religion & Public Life

Project. Pew Research Center, October 6, 2022. https://www.pewresearch.org/religion/2021/05/11/jewish-americans-in-2020/.

Morag, Shelomo. (2019). The Vocalization Systems of Arabic, Hebrew, and Aramaic : Their Phonetic

and Phonemic Principles (version 2nd printing. Reprint 2019.) 2Nd printing. Reprint 2019 ed. Janua Linguarum. Series Minor, 13. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110812398
Measuring the Voice Onset Time of Navajo and Tlingit Stops © 2024 by Simone Brown is licensed under CC

BY-ND 4.0

Measuring the Voice Onset Time of Navajo and Tlingit Stops

Simone Brown

Department of Linguistics, McGill University LING 411: Structure of an Indigenous Language Professor James A. Crippen

Abstract

VOT (Voice Onset Time) is a measurement used to understand a language's phonetic profile, i.e. the unique production of phonemes within a language. This paper will examine two languages in the Na-Dene language family of North America— Navajo and Tlingit— and measure the VOT of stop production in the continuous speech of native speakers. Both languages are unique for their three-way stop contrast between plain voiceless, aspirated voiceless, and ejective stops. Included is a discussion of the various considerations accompanying Indigenous language research, including orthographic variation, transcription bias, and dialectal and idiolectal variation. The VOT measurements show a distinct contrast between stop categories (plain voiceless, aspirated voiceless and ejectives), but evidence suggests there is a range rather than a specific length that speakers target in their production. Both languages demonstrate a preference for plain stops and include several unmeasurable 'unreleased stops.' Tlingit demonstrates a wider range and more accurate VOT targeting than Navajo, while Navajo has significantly longer VOTs, suggesting separate repairing mechanisms in the two languages to manage perception difficulties. The phonological and perceptive processes of ejectives are also discussed, particularly the ejective and non-ejective contrast. Suggestions for further research are provided.

1 Introduction

The Na-Dene language family is well known for its incredibly rich phonetic inventories. Previous research on these sound systems is unfortunately inadequate for building a complete phonetic profile of individual languages or the language family. Many linguists have recognized this deficit in phonemic analysis and attempted to repair it via purposeful research studying the phonetic and phonological processes of Na-Dene languages. The study of consonants, particularly obstruents, reveals intricate phonological processes in Na-Dene languages that affect their production. Particular emphasis has been placed on variable sound length (gemination) in the Athabaskan language family, a subset of the Na-Dene language family encompassing Navajo, Lheidli, Apache, and more. In Lheidli, intervocalic consonant length is substantially longer than consonants in other positions and even the intervening vowels (Bird 2002). Similar proof has been found for intervocalic consonant lengthening in Navajo (Sapir & Hoijer

1967, Young & Morgan 1987, McDonough & Ladefoged 1993). In Apache, evidence suggests a combination of stress placement and morphological position impacts stop length. Still, Apache stop voice onset time is not impacted when placed at a morphological boundary, instead consistently surfacing as longest in phrase-initial positions (Tuttle 2005). Following in the footsteps of these studies, this paper will study sound length in the Na-Dene language family by analyzing the stops in two languages, Tlingit and Navajo, hoping to provide a basis for future research studying its variation and conventions.

Contextualizing this research are two theories surrounding the selection of phoneme contrasts cross-linguistically: the Aperture theory by Keating (1990) and the 'continuum' theory by Cho & Ladefoged (1999). While each addresses how languages define the contrastive nature of voicing, they do so by proposing entirely different mechanisms. Keating's Aperture theory states that all languages universally select their phonemes from three distinct phonetic categories: voiced, voiceless aspirated, and voiceless unaspirated. Each of these distinct phonetic categories manifests as a singular length or 'target' the speaker is trying to produce. Under Keating (1990), Navajo and Tlingit select two of these phonetic categories and ejectives (generally omitted from this theory on voicing universals due to the dramatic differences in their production mechanisms). Each category would have a 'target', in which each singular length represents the type of stop being produced, and any variability in production is due to natural variation in human production, not encoded into the theory itself. Still, this theory fails to address the consistent variability within stop classifications and the language-specific constraints on sound length.

Cho & Ladefoged (1999) propose that the voicing of phonemes is selected from a continuum. This continuum explains the inherent variability in the pronunciation of stops despite their clear categorizations and occurrences of partial voicing, unreleased stops, and similar phenomena. This scale is applied language-specifically or in groupings to explain the relationship between different languages and language families that occupy the same continuum. Under this theory, Navajo and Tlingit would exist on the same continuum, but each language selects different positions on the scale— close enough to be recognized for their voicing category but not uniform.

In contrast to Keating (1990), which neglects the question of ejectives, Cho & Ladefoged (1999) explicitly dismiss the contrast between ejectives and other stops relying on post-release length. Both studies fail to address the adequate similarities between ejectives and other stops for measuring Voice Onset Time (henceforth called VOT) and the possible implications of these measurements in languages where ejective and non-ejective equivalents contrast. Ejectives must therefore be treated as meaningful data to be gathered and analyzed alongside other stops in Navajo and Tlingit, with the caveat of understanding their differences.

Rather than presupposing that Navajo and Tlingit simply distinguish between uniform groups (voiceless, voiceless aspirated and ejective stops), using a spectrum-based theory allows researchers to

determine with more specificity what the manifestations of these stops surface as. By comparing VOT measurements, this paper seeks to determine the phonetic profile of stops in Navajo and Tlingit. The three questions driving this research are: (1) What are the average VOTs for the stops in Navajo and Tlingit? (2) How do VOTs compare across each category (voiceless, aspirated, and ejective)? (3) What might the discerned patterns tell us about the relationship between these two languages and beyond?

2 Language(s) Background

Tlingit (*KLING-IT*) and Navajo (*NAH-VUH-HO*) are two languages of the Na-Dene language family, an Indigenous language family of North America. Tlingit is spoken along the Western coast of Alaska, British Columbia, and the Yukon. The language is considered critically endangered and has only a few hundred living native speakers, spurring intense language documentation and revitalization efforts by Tlingit communities. Historically, Tlingit split from the Na-Dene language family long before Navajo's formation, making them distant relatives descending from a singular Proto-Na-Dene. Navajo is spoken within Southwestern North America, including parts of Arizona, New Mexico and Utah. Comparatively, Navajo is a much younger language and maintains cross-generational transmission, but is still considered highly vulnerable to erasure and requires revitalization interventions.

Tlingit and Navajo share many phonetic characteristics associated with the Na-Dene language family, particularly the use of contrastive tone (high and low), a large consonant inventory and a complete lack of voicing in stops and affricates. Tlingit and Navajo then have a three-way contrast for stops that distinguishes plain voiceless, aspirated voiceless, and ejective voiceless as separate phonemes (Hargus 2010). These languages also contrast regular back obstruents against labialized back obstruents marked with a 'w' diacritic. The distribution of consonants in Na-Dene languages is influenced by the morphological and syntactic constraints of the language; for more information on morphophonology, see McDonough (2013).

Despite these similarities, these two languages are discernable from one another in several key ways. Tlingit has several ejective fricatives (/t', k', k^w'q', q^w'/) that Navajo lacks, while Navajo has a bilabial plosive (/p/) that Tlingit lacks. Tlingit also has very few sonorants— one nasal (/n/) and two approximants (/w, j/)— compared to Navajo's two nasals (/m, n/) and three approximants (/w, j, u_l/). Consonants in Tlingit prefer dorsal places of articulation, while Navajo is primarily composed of coronal sounds (Hargus 2010).

The voicing contrasts of the stops and affricates are indicated by a unique type of orthography shared by both languages. Traditionally, Na-Dene languages use the IPA symbols for their voiced

counterparts to mark 'plain' voiceless stops (/d/ = [t]). The aspirated voiceless stops use the IPA symbol for the equivalent voiceless segment $(/t/ = [t^h])$. The ejective voiceless stops use the IPA symbol for the equivalent voiceless segment followed by an apostrophe (/t'/ = [t']). Despite minor differences in symbol usage between linguistic adaptations of the orthographies of Navajo and Tlingit, this division is currently the most popular writing system. This writing system will be maintained throughout the paper except when the distinction between sound segments necessitates clarification.

	Bilabial	Laterals	Alveolar	Post-	Palatal	Velar	Glottal
				alveolar		(+Labialized)	
Unaspirated	(p)		t			k k ^w	3
Stop							
Aspirated	(p ^h)		t ^h			k ^h k ^{wh}	
Stop							
Ejective Stop			t'			k'	
Unaspirated		€¶	fs	t∫			
Affricate							
Aspirated		ff h	$\widehat{\mathrm{ts}}^{\mathrm{h}}$	f∫ ^h			
Affricate							
Ejective		ff '	ts'	tĵ"			
Affricate							
Voiceless			S	ſ		x	h
Fricative							
Voiced			Z	3		Y	
Fricative							
Laterals			1				
			1				
Nasals	(m)		n				
Approximant	W				j		

Figure 1: Navajo consonant inventory (IPA) adapted from Young & Morgan (1980) and McDonough (2013).

	Labial	Alveolar	Post-	Lateral	Palatal	V V	/elar	U	vular	Glottal
			alveolar			(+Lab	vialized)	(+Lat	vialized)	
Unaspirated Stop	(p)	t				k	kw	q	\mathbf{q}^{w}	3
Aspirated Stop	(p ^h)	t ^h				k ^h	\mathbf{k}^{wh}	q^{h}	$q^{\rm wh}$	
Ejective Stop		ť'				k'	k ^w '	q'	q ^w '	
Unaspirated Affricate		fs	f							
Aspirated Affricate		tsh	fʃħ							
Ejective Affricate		ts'	tĵ"							
Plain Fricative		S	ſ			x	\mathbf{X}^{W}	χ	$\chi^{\rm w}$	h
Ejective Fricative		s'				x'	X ^w '	χ'	χ^{w}	
Nasal Stop	(m)	n					(ŋ)			
Approximant				(1)	i	w	щ			

Figure 2: Tlingit central consonant inventory (IPA) adapted from Maddieson, Smith & Bessel (2001).

3 Methods

The anatomy of a stop can be divided into three main sections: (1) the closure, where the airflow is obstructed even as pressure mounts in the vocal tract, (2) the release, where the obstruction is removed, creating a sharp relief of the pressure in a 'popping' sound, and, depending on the stop, (3) a period of exhalation before the next sound segment begins (Lisker & Abramson 1964). Typically, stops are produced with pulmonic eggressive airflow, with the obstruction originating somewhere in the mouth. Plain voiceless stops lack a lengthy period between the release and the beginning of the following sound. In the brief period following the release, the stops lack energy or 'noise' on the spectrogram and the oscillogram, as no exhalation accompanies the release. In contrast to plain stops, aspirated stops have a longer pause between the release and the beginning of the built-up pressure, called aspiration (Lisker & Abramson 1964). Ejective stops differ from plain and aspirated stops as they utilize a glottalic egressive airstream mechanism, with the sound originating in the contraction of the glottis rather than air travelling from the lungs (Wright, Hargus & Davis 2002).



Figure 3: Example of a plain voiceless alveolar stop in Navajo, /t/. Its anatomy is marked with red lines, i.e. the stop's three components: the closure, the release, and the end of the sound.

Since its conception in 1968, Voice Onset Time (VOT) has become a major phonetic tool used to identify contrasts between individual stop tokens and discern larger patterns in the lengths of different stop classifications within a language (Lisker & Abramson 1964). Phoneticians commonly use voicing to reliably distinguish stops from one another within voice recordings or in cases where the larynx activity is not easily measurable. Beyond phonetics, VOT is also an important tool for phonological work as it helps to measure the variable pronunciation of stops depending on their context (Abraham & Whalen 2017). This paper will focus solely on VOT averages for the three contrastive structures of stops within Navajo and Tlingit. Still, it will include a discussion of phonological variability and the next steps for research within the discussion portion. Cho & Ladefoged (1999) describe a continuum for universal voicing that problematizes using VOT to measure the contrast between stops as the variable production can lead to frequent miscategorization of stops. This indicates that VOT is insufficient for addressing the breadth of contrastive stops cross-linguistically, since languages like Hindi or Korean exhibit unusual voicing and aspiration contrasts (Abraham & Whalen 2017).

VOT is a measurement of the time that elapses between the 'release' of the stop and the 'onset' of glottal pulses typically associated with the subsequent sound segment. The release is marked as the 'zero,' or starting point, and the onset is placed somewhere along a negative to positive scale centered around that starting point. Although the negative to positive scale (i.e. whether they occur before or after the starting point) is consistent, classifying stops based on their VOTs depends on traditional patterns for stops within a language. Following standards in phonetic analysis, this paper will assume that: (1) aspiration will lengthen the VOT, (2) the 'release' is crucial to the measurement as a starting point, and (3) in plain stops, the voiced segment will traditionally be longer than its voiceless counterpart (Abraham & Whalen 2017).

In their original paper, Lisker & Abramson (1964) identified the optimal context for measuring VOT as the onset of a CV syllable, and subsequent research identified intervocalic, clustered, and codaposition stops as more likely to experience "bleed"— accidental inclusion of the phonetic properties of surrounding sounds— and provide less accurate data. Previous research in Navajo has also indicated that speakers are likely to target the optimal CV syllable through affixation and epenthesis when necessary to avoid 'marked' syllable structures (Wright 1984). To accurately identify the VOTs of different stops within this study, only stops within CV syllables will be measured. Focusing on stops within CV syllables will ensure that research is not influenced by secondary phonological constraints that limit the inclusion and pronunciation of stops in casual speech.

Also, any stops that do not contain a prominent 'release' are discounted, as there is no accurate way to differentiate the starting point of the sound segment from the closure that precedes it. These sound segments are traditionally called 'unreleased stops' and are characterized by minimal activity in an area where the stop would be typically produced. Despite these unreleased stops lacking the necessary qualities for measuring VOT, they still contain a closure and may hold phonological significance. Across the 83 syllable-initial CV clusters analyzed for each language, each had four unreleased stops. See the 'Discussion' portion for further information on 'unreleased stops' and glottal stops.



Figure 4: Example of an unreleased stop that is targeting $/k^{h}/$. Its anatomy is marked with red lines, i.e. the closure and the end of the sound, with a noticeable lack of a release point.

Researchers identified 166 sound segments (83 sound segments in each language) and, excluding the eight unreleased stops, analyzed 158 individual stops encompassing the complete phonetic inventory of each language. These stops were categorized based on language, then further divided by stop type. The unequal distribution of stops is due to the measurement technique and the asymmetrical distribution of the stop types. The first 83 tokens in a 3-minute recording of casual speech were selected to minimize the introduction of research bias through the hand-selection of tokens.

The source for the Tlingit portion is *Four Tlingit Stories*, a collection of oral stories narrated for the October 2014 Tlingit Literacy session at the Yukon Native Language Centre. The storyteller is Sam Johnston, a fluent speaker and leader of the Ishkitàn clan, and Dr. Jeffry Leer transcribed and translated the audio. The source for the Navajo portion is the *2009 Valentine's Day message* by the 2008-2009 Miss Navajo, Yolanda J. Charley, broadcast on KTNN radio. The audio from the broadcast was later transcribed and translated by Daybreak Warrior on Youtube as a part of his ongoing efforts to increase the accessibility of Navajo.

The sources were selected based on speaker criteria, legibility, accuracy and clarity. Speakers were required to be fluent native speakers of their respective language and communicate with relative confidence in their production. Researchers ensured that the audio had minimal background noise, the speaker was enunciating, the transcription was explicit, and the translation was accurate. The 2009 *Valentine's Day message* and *Four Tlingit Stories* were selected based on their ability to meet all necessary criteria. The decision to use one source for each language is partly meant to simplify the datagathering process and partly due to insufficient sources that met the above criteria. For further discussion about the impacts of single sources on research, see the subsections under 'Consideration' about idiolectal variation and orthography.

The process for the identification and measure of VOT was performed on Praat, a software used for phonemic analysis. All 'stop' tokens in the recordings were identified, and the three parts of the stop were marked and labelled. The beginning of the 'burst' was defined as the first meaningful spike on the oscillogram following a quiet period, i.e. the closure. The end of the sound segment was determined to be the transition in the oscillogram/spectrogram that coincided with the beginning of a vowel. The distance between the spike, i.e. the 'burst' and the transition, i.e. the end of the sound segment, was measured in milliseconds. All measurements were collected at the nearest zero crossing (the boundary between negative and positive measurements) on the oscillogram, in line with phonetic measurement standards. Once the 83 tokens had been measured in their order of appearance, they were recorded in a table and averaged to determine the mean VOT of each stop type for Navajo and Tlingit. Other data calculated from the measurements included the range and distribution of the tokens based on various parameters.

4 Considerations

i) Transcription/Orthographic Bias

In any transcription or translation, the choices made by the transcriber can seriously affect the accuracy of the new written communication with respect to its original intended message by the speaker. This is unfortunately one of the unavoidable side effects of transcription and translation work.

Transcribers are tasked with first interpreting the message through the lens of their own experiences and then relaying that interpretation of the message in a different format. The transcriber's choices can lead to gross misinterpretations of source material, which problematizes the source's usage when seeking to study its qualities. Often in Indigenous storytelling, the author must represent grammatical elements in roundabout or insufficient ways because English has no equivalent. Dauenhauer & Dauenhauer (2011) list several challenges that translators must address in Tlingit, where different representations can entirely change the meaning and composition of the utterance. These include the representation of imperfectives that lack English equivalents and the whole of the aspectual system in grammar for which English has no comparable system. Since Tlingit and Navajo share many grammatical elements that lack an English equivalent, it can be assumed that both translators have repaired these 'gaps' based on their own experiences as bilingual speakers of English and Navajo or Tlingit (Dauenhauer & Dauenhauer 2011). The transcriptive choices are significantly more likely to affect this research, but vigilance of translator bias remains crucial.

Similar to problems created by transcription bias, the choices made in how to represent the sounds of the language best can introduce orthographic bias. An 'orthography' refers to a language's spelling convention associating sounds/segments with certain symbols. Generally, the Indigenous languages of Canada contain much of the same orthography as English, especially for those sounds that English itself has. However, orthographic representation becomes more tricky when representing sounds in Indigenous languages not traditionally considered phonemic in English, such as ejectives or aspirations. The orthographies of Navajo and Tlingit are highly contested among different groups, each following different linguistic standards on how best to represent these unique sound segments.

Some of these concerns can be addressed through a clear methodology that minimizes the opportunity for bias in data collection and analysis. The Navajo speaker was providing a speech on the radio, and the Tlingit speaker was telling a story to a room of people, creating a setting that encouraged projective speech. Similarly, both speakers understood that their voices would be recorded and transmitted through speakers, so their mediums would encourage careful and emphatic speech productions. Such circumstances likely created the conditions for the speakers to project their voices to increase their prominence and enunciate their speech sounds more to increase the recording's intelligibility to listeners. The clarity of speech recordings minimized the transcription bias due to misinterpretation or difficulty discerning the sound segments originating in the recorded conversation's poor quality or casual nature.

The familiarity and experience of the transcribers would also aid in minimizing the bias found within their transcriptions. Daybreak Warrior is a bilingual native speaker of Navajo and English and had several years of experience by this time in transcribing Navajo speech and translating it to English in the process (Daybreak Warrior 2006). Dr. Jeffry Leer is a professor emeritus at the University of Alaska Fairbanks who specializes in studying Na-Dene languages, making him familiar with transcription conventions and research for the language family. He is also a bilingual speaker of both English and Tlingit, giving him ample experience in Tlingit (Alaska Native Language Center 2023).

Despite these interventions, the inevitability of transcription and orthographic bias necessitates the constant consideration of these factors in tandem with any conclusions drawn from the data gathered. The limited amount of sources to choose from when studying Indigenous languages and the limited control by researchers on how these sources are recorded makes strategies such as multimodal transcription difficult to implement in practice (Mondada 2018). Several common orthographic and transcription biases are listed below to prepare readers on what biases are relevant to the results of this study.

Non-verbal information:

Voice recordings, and therefore the researchers who study them, are unable to account for the considerable amount of information that is conveyed using non-verbal communication. Non-verbal communication refers to the information communicated by the speaker separate from the utterance and its properties. Non-verbal communication includes hand gestures, facial expressions, pointing, eye contact, posture, and other movements or body language (Wiener, Devoe, Rubinow & Geller 1972). Without any way to track anything beyond the spoken word, all nonverbal communication is lost with audio recordings. Speakers often use non-verbal communication in tandem with verbal communication, such as pointing to a location being discussed. It can also provide additional context that changes or expands the meaning of what is being said. Since this information is missing from audio recordings, transcribers cannot record or incorporate it into their transcriptions or translations. Two common solutions are providing video recordings of the speaker for the transcriber to watch or having the transcriber has more opportunity to consider nonverbal cues and how they may affect the utterance. These solutions still require further revision, considering non-verbal communication is only sometimes easily translatable into words and often up for interpretation, leaving it vulnerable to bias-introduction (Mondada 2018).

Metacommunication:

Much like entirely nonverbal communication, information can also be communicated by secondary qualities of speech that are separate from the actual conventions of utterance building. These qualities include paralinguistic and prosodic information, such as pitch, duration, speaking style (e.g. passive vs active), and stress. Most importantly, they are not a contrastive quality of language. Pitch is a prosodic quality of language that can change the intended meaning of a phrase, while tone is a phonemic quality of language that can change the direct meaning of a sound segment/word. For example, pitch-raising at the end of an utterance may indicate it is a question, while a high tone would change the phoneme being used by the speaker. Transcribers often neglect meta-linguistics since it does not involve observing what is being said but rather the utterance's implications.

Orthographic representation:

The history of orthography invention for Indigenous languages is often intertwined with the increasing politicization and colonization of Indigenous identity, and Navajo and Tlingit are no exception. Much of the early orthographic work was created by colonial settlers and academics using either Cyrillic or Latin conventions, rarely receiving input from native speakers (Sebba 2007). This makes much of the foundation of modern orthographies of Indigenous languages fragile at best and unrepresentative at worst. In recent decades, linguists and native speakers have worked to establish orthographic standards that would better represent the sounds of the language and avoid the colonial bias introduced in earlier manifestations of the writing system. Still, these orthographic standards are not without problems. Often orthographic standards are developed based on a single dialect of a language, then extrapolated onto all other dialects, whether or not it accurately represents their sound system. Orthographic standards are frequently only standardized within certain spheres of usage or among certain individuals, such as academic settings, limiting its accessibility to the language's speakers (Littell et al. 2018). Authors and transcribers are rarely explicit about their orthographic choices, choosing whichever style they have the most experience in without acknowledging its insufficiencies in representing the speaker's pronunciations. The writing conventions of transcribers are also commonly focused on representing what the speaker means to say rather than what they actually say, leading to a misrepresentation of the data. For example, neither Daybreak Warrior or Dr. Leer marked the unreleased stops identified by researchers, representing them instead based on their target sounds. Since researchers are basing their categorizations on the transcribers', these classifications might lack the nuance within the data and therefore be unfairly

represented within this research. While there is little evidence that orthographic bias has significantly impacted this paper's conclusions, it is still an important consideration in examining the data.

ii) Dialectal and Idiolectal Variation

Much of Tlingit's dialectal variation focuses on non-phonological elements or the production of tones or vocoids. There is also ample evidence for dialectal variation in Navajo. Yet dialectal variation is unlikely to affect VOT measurements as it typically concerns lexical change or non-stop phonemes like vowels and fricatives (Saville-Troike 1974). Both languages demonstrate a parallel form of dialectal variation that involves a change in the place of articulation for obstruents. In Tlingit, this arises with some speakers' delabialization of stops word-finally. In Navajo, this materializes as the regular substitution of $/t/ \rightarrow /k/$ among certain dialects, including Kiowa Apache and Jiricallia. While these differences affect the pronunciation of stops, there is no evidence that VOT in Navajo is impacted by place of articulation (Saville-Troike 1974). Thus, despite the difference in where each speaker produces stops, there is little possibility that these phenomena would impact VOT stop measurements.

Similarly, the largest idiolectal variation found within Tlingit is regressive palatalization and progressive labialization, and while they target the place of articulation, they affect vowels rather than obstruents (Wolfe 1977). No research into the idiolectal variation of Navajo speakers has been conducted thus far, limiting the opportunity to examine the phenomenon's impact on research. All the same, research shows that on an individual level or within a dialect of these languages, changes in the production of stops concern the place of articulation rather than the stop's manner of articulation or any secondary phonemic qualities. There is no evidence in either language that dialectal or idiolectal variation influences the categorization of stops or their VOT measurement. It is unlikely that these subtle differences between stops will impact data accuracy. New research on the dialectal and idiolectal variation in Na-Dene languages is necessary, but it can be assumed that these changes are not substantial enough to warrant data intervention for the purposes of this study.

5 Results

As mentioned above, 162 sounds were selected, and 156 stops were analyzed following exclusions. These 156 stops were divided equally among Tlingit and Navajo, with each language having 78 analyzed stops. In Navajo, 54 were plain stops, 13 were aspirated, and 11 were ejectives. In Tlingit, 43 were plain stops, 26 were aspirated, and 9 were ejectives. The ratio of stop types (plain, aspirated and

ejective) were proportionate between the two languages, with the number of aspirated stops reduced in Navajo in favour of plain stops.



Figure 5: Graphs demonstrating the distribution of stop types (plain, aspirated, ejective) in Navajo and Tlingit.

The average length of VOT for plain stops was 21.1 ms in Navajo and 11.5 ms in Tlingit. The average length of VOT for aspirated stops is 90.5 ms in Navajo and 36.9 ms in Tlingit. The average length of VOT for ejective stops is 76.5 ms in Navajo and 31.8 ms in Tlingit. Ignoring their categorization, the average VOT for all analyzed stops was 48.4 ms in Navajo and 22.3 ms in Tlingit.

		Navajo		Tlingit			
	Plain	Aspirated	Ejective	Plain	Aspirated	Ejective	
Avg. VOT (ms)	21.1	90.5	76.5	11.5	36.9	31.8	
Avg. VOT (ms)	48.4			22.3			

Figure 6: Summary of the results divided by stop types and languages.



Figure 7: On the left, a graph demonstrates average VOTs compared between the stop categories of Navajo and Tlingit. Figure 8: On the right, a graph demonstrates the average VOT of each stop category, combining Navajo and Tlingit data.

6 Discussion

i) Stop Type Comparison

There is an evident correlation between stop types and VOT measurements: plain stops are associated with shorter VOTs, aspirated stops are associated with longer VOTs, and ejective stops are considered to exist somewhere between these two boundaries. Ejective stops are less uniform in their productions than aspirated and plain stops, leading to higher variability in their VOTs. Ejective tokens vary by 63 ms in Tlingit and 89 ms in Navajo, while plain stops vary by 44 ms and 25 ms, respectively. Aspirated and ejective voiceless stops are closer in their VOTs than plain voiceless stops are to either type of stops.

In both languages, the most common stop is plain, followed by aspirated, the least common being ejectives. The majority of the tokens analyzed were plain stops (62% and 55%), demonstrating a clear preference for plain stops over aspirated (18% and 33%) and ejective stops (14% and 11%). This preference may indicate that plain stops are the least 'marked' of the phonological categories in Na-Dene languages, likely due to their simplicity in production compared to ejective and aspirated stops.

ii) Navajo and Tlingit Comparison

While Navajo and Tlingit may select different points along the VOT 'continuum' as their production target, they still pattern similarly. Speakers consider the distinction between these stop structures crucial to their production, which explains the insignificance of minute differences in the data. Navajo has significantly longer VOTs than Tlingit, averaging 36 ms longer in stop VOTs than Tlingit. The most significant difference in VOT length occurs in aspirated stops; Navajo aspirated stops have the longest VOT at 90.5 ms compared to Tlingit aspirated stops averaging 36.9 ms. This data suggests that, on average, Navajo speakers produce longer stops than Tlingit speakers produce their equivalents.

In general, the range of VOTs is much more narrow in Tlingit compared to Navajo. The data clearly shows a difference in how Tlingit and Navajo arrange their phonemic inventories. Navajo is less accurate in targeting phonemes but repairs this inaccuracy through wider divides between individual sound categories. Tlingit is more accurate at targeting phonemes, so wide distinctions between sound categories are unnecessary. These differences highlight how languages can repair indiscernibility as it arises, either by promoting the differences between phonemes or increasing phoneme production's accuracy.

An unexpected value found in both languages was unreleased stops, which could be identified but not measured, given the lack of the necessary 'release' for VOT measurements. Three unreleased stops were present in the first 81 sounds selected for analysis in both languages. They did not appear in similar contexts among the 81 sounds or contain any significant similarities. Further research is necessary to determine if the equivalent number of unreleased stops indicates a pattern or is coincidental.

iii) Language Family Comparison

Unfortunately, research is currently confined to Tlingit and Navajo, neglecting the many other languages in the Na-Dene family that may contribute to our understanding of VOT. This lack of research poses limitations primarily around the accuracy and form of conclusions about the Na-Dene language family based solely on the data analyzed in this study. Speculations will be made to address this on the assumption that Tlingit and Navajo are at least somewhat representative of the whole language family, with the caveat that this may turn out false, pending new research on the subject.

Since Navajo and Tlingit have similar patterns, we can assume these standards are upheld throughout the Na-Dene language family. The high contrast between plain and aspirated stops is likely to parallel across all languages simply due to the nature of each stop type, with one requiring significant soundwaves following the release and the other requiring an absence of that activity. The distribution of stops for the two languages is not necessarily demonstrative of a wider pattern for the distribution of stops in the Na-Dene language family since it may be incidental that plain stops are the preferred stop type. The distribution, especially that of Tlingit, is ambiguous enough that it could be debunked through a series of examples of other languages that favour aspirated or ejective stops over plain ones.

This research can also be used to explain the phonological processes surrounding ejectives in the Na-Dene language family. Ejective voiceless stops have the highest degree of variability in their VOTs, with the tokens ranging from 11-74 ms in Tlingit and 20-109 ms in Navajo. Aspirated and ejective voiceless stops are closer in their VOTs than plain voiceless stops are to either type of stops. The ejective/non-ejective contrast for stops is consistently difficult for listeners to discern in other Na-Dene languages (Wright, Hargus & Davis 2002). The lack of distinction may be due to the variability with which ejectives are produced and the sharp difference in their production compared to aspirated stops. If ejectives can be anywhere along the VOT range of the language, then VOT length is no longer a reliable feature in categorizing stops. Similarly, if ejective stops consistently mirror the VOT of aspirated stops, listeners may struggle to accurately identify sounds from one another. While more research is necessary to study the extent to which other languages of the Na-Dene family are affected by the minimal ejective/non-ejective contrast, these VOT measurements may suggest this indiscernibility is due to similar VOT realizations.

iv) Further Research

Following the measurements and conclusions provided in this study, several avenues for future research present themselves: (1) A study into the appearance and mechanisms of unreleased stops, whether at a language or individual-specific level. (2) A study into discernibility between the ejective/non-ejective contrast in Navajo or Tlingit. (3) A study examining the change in VOT length or the distribution of stop types across morphological domains. (4) A study measuring the VOT of stop categories in other languages beyond Navajo and Tlingit in the Na-Dene language family.

7 Conclusion

As discussed in the introduction, three questions guided this research: (1) What are the average VOTs for the stops in Navajo and Tlingit? (2) How do they compare across each category (voiceless,

aspirated, and ejective)? (3) What do the discerned patterns tell us about the relationship between these two languages and beyond? To answer these questions, the stops identified in the continual speech of two native speakers, Sam Johnston, via 'Four Tlingit Stories' and the 2008-2009 Miss Navajo Yolanda J. Charlye, via her Valentine's Day message, were measured.

Stop categorization (voiceless plain, aspirated, and ejective) dictated the length of VOT for sound segments, where longer sound segments were aspirated, and shorter sound segments were plain. The length of ejectives was much more variable in both languages than the production of plain and aspirated stops, likely contributing to the perception difficulties previously identified along the ejective/non-ejective division. The distribution of stops in Navajo and Tlingit shows a preference in both languages for plain stops, with ejectives being the least frequently produced. This preference may be due to the 'unmarkedness' of plain stops compared to the other phonemic stop categories. Navajo and Tlingit show different mechanisms for increasing the discernibility of stop contrasts; Navajo increases the distinction between stop classifications and Tlingit increases accuracy in targeting productions. The similarities between Navajo and Tlingit were predicted to represent wider patterns of the Na-Dene language. Such theories necessitate more data collection and analysis of other languages in the language family.

Returning to the theories in the Introduction, this data supports a perspective on voicing that is more similar to Cho & Ladefoged (1999) than Keating (1990). Despite emulating the unconscious distinction between stop types, the speaker's production of stops is highly variable and resembles a target 'range' rather than a singular point. This suggests that phonemes are selected from a continuum of voicing, with each language selecting ranges that, although similar, show minute differences in their realization of each stop classification. The differences in how Navajo and Tlingit repair indiscernibility further support the soundness of a range-based rather than a point-based theory. Tlingit appears to align more with Keating (1990)'s theory due to its narrower realization of stops, while Navajo aligns more with Cho & Ladefoged (1999) due to its higher production variability. However, Cho & Ladefoged (1999)'s theory is more easily adapted to encompass evidence from Tlingit than Keating (1990)'s is to readily accept Navajo. This research may suggest that the selected 'range' on a continuum is not done uniformly and varies depending on language-specific factors, much like the target ranges each language selects. For example, applying Tlingit to Cho & Ladefoged (1999)'s theory may entail acknowledging that the selections are of much shorter ranges, leading to what appear to be more accurate 'target' productions. The question of ejectives' role in these theories remains unaddressed. Still, increasing the available data may lead to a revision of voicing theories to encompass ejectives as well.

Furthermore, this paper discussed several considerations that may impact phonetic research, including transcriber influence, orthographic bias, and dialectal and idiolectal variation. The weight of these concerns was debated, and several solutions were provided to ensure the accuracy of any conclusions drawn from the data. A brief discussion was provided about the exclusion of glottal stops and the commonality of unreleased stops in the two languages. Finally, further research suggestions were provided to confirm the hypotheses proposed in this paper or expand the degree of analyses available to linguists about the Na-Dene language family.

References

Abraham, Arthur & D. H. Whalen. 2017. Voice Onset Time (VOT) at 50: Theoretical and

practical issues in measuring voicing distinctions. Journal of Phonetics 63: 75-86.

Bird, Sonya. 2002. *The phonetics and phonology of Lheidli intervocalic consonants*. University of Arizona: Ph.D. dissertation.

Chodroff, Eleanor, Alessandra Golden, & Colin Wilson. 2019. Covariation of stop voice onset

time across languages: Evidence for a universal constraint on phonetic realization. *The Journal of the Acoustical Society of America* 145: 109-115.

Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27.2: 207-229.

Dauenhauer, Richard & Nora Marks Dauenhauer. 2011. Tense, Aspect, and the Organization and

Translation of Tlingit Oral Narrative. In *Working Papers in Athabaskan Languages*. Fairbanks, Alaska: Alaska Native Language Project.

Daybreak Warrior. 2010. Miss Navajo's Valentine's Day Message.

https://www.youtube.com/watch?v=waELx-wZm0c (Accessed 30 January, 2023).

- Edwards, Keri. 2009. Dictionary of Tlingit. Juneau, Alaska: Sealaska Heritage Institute.
- Hargus, Sharon. 2010. Athabaskan phonetics and phonology. *Language & Linguistics Compass*4.10: 1019–1040.
- Johnston, Sam. 2016. Four Tlingit Stories. *Tlingit Literacy Session*. Jeffry Leer (trans.).

Whitehorse: Yukon Native Language Centre. <u>https://ynlc.ca/tlingit/#four-tlingit-stories-by-</u>sam-johnston.

- Langford, Peter & Ian Maddieson. 1993. *The Sounds of the World's Languages*. Hoboken, New Jersey: Wiley-Blackwell Press.
- Lisker, Leigh & Arthur S. Abramson. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD* 20.3: 384-422.

Littell, Patrick et al. 2018. Indigenous language technologies in Canada: Assessment, challenges,

and successes. In *Proceedings of the 27th International Conference on Computational Linguistics:* 2620-2632. Santa Fe, New Mexico: Association for Computational Linguistics.

Maddieson, Ian, Caroline L. Smith & Nicola Bessell. 2001. Aspects of the phonetics of Tlingit.

Anthropological Linguistics 43.2: 135–176.

- McDonough, Joyce. 2013. The Dene verb: how phonetics supports morphology. *18th Workshop* on Structure and Constituency in the Languages of the Americas. University of California, Berkeley.
- Mondada, Lorenza. 2018. Multiple Temporalities of Language and Body in Interaction:

Challenges for Transcribing Multimodality. *Research on Language and Social Interaction* 51.1: 85–106.

- Saville-Troike, Muriel. 1974. Diversity in Southwestern Athabaskan: A historical perspective. *Navajo Language Review* 1.2: 18-35.
- Sebba, Mark. 2007. *Spelling and Society: The Culture and Politics of Orthography around the World*. United Kingdom: Cambridge University Press.
- Tuttle, Siri G. 2005. Duration, intonation and prominence in Apache. In Sharon Hargus and Keren Rice (ed.), *Athabaskan prosody*, 319–44. Amsterdam: John Benjamins.
- University of Alaska Fairbanks. J. Leer: Comparative Athabaskan Lexicon (CAL), Alutiiq lexicon. National Science Foundation. <u>https://www.uaf.edu/danl/project-updates/jeff-leer/</u> (Accessed 20 April, 2023)
- Wiener, Morton, Shannon Devoe, Stuart Rubinow & Jesse Geller. 1972. Nonverbal behaviour and non-verbal communication. *Psychological Review* 79.3: 185-214.
- Wolfe, Gaelan de. 1977. *Tlingit Phonology in a Generative Framework: An examination of phonological processes and abstract representation*. University of British Columbia: MA thesis.
- Wright, Martha. 1984. The CV Skeleton and Mapping in Navajo Verb Phonology. *North East Linguistics Society* 14.28: 461-477.
- Wright, Richard, Sharon Hargus, & Katherine Davis. 2002. On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association* 32.1: 44–77.
- Young, Robert & William Morgan. 1980. *The Navajo language: A grammar and colloquial dictionary*. Albuquerque: University of New Mexico Press.

The Gay Way to Say: A Review of Sociolinguistic Research on Gay Speech © 2024 by Kaikoura Gutteridge is

licensed under CC BY-ND 4.0

The Gay Way to Say: A Review of Sociolinguistic Research on Gay Speech

Kaikoura Gutteridge

Department of Linguistics, McGill University LING 320: Sociolinguistics 1 Professor Charles Boberg

Abstract

The elusive concept of "gay voice" has intrigued researchers for decades, prompting investigations into the intricate relationship between language, sexuality, and identity. This paper synthesizes findings from five seminal studies and a meta-analysis to shed light on the production and perception of gay speech. Beginning with Rudolf Gaudio's seminal study on pitch properties in gay and straight speech, we then turn to Erez Levon's exploration of prosodic variables, Robert Podesva's examination of phonetic features, and Drew Rendall's biosocial hypothesis on gay speech, and concluding with Deborah Cameron and Don Kulick's elucidation of gay language's sociolinguistic evolution. This synthesis underscores the multifaceted nature of gay speech, influenced by sociocultural, biological, and contextual factors. It prompts a re-evaluation of assumptions regarding the gay community and calls for a nuanced understanding of identity construction through language. As society's understanding of gender and sexuality evolves, future research may explore the interplay between gay speech and popular culture, particularly in the digital realm, offering fresh insights into this rich sociolinguistic phenomenon.

1 Introduction

The gay community has doubtlessly long been subjected to the ever-changing, and often oppressive attitudes and notions imposed by the outside population, but one particular question has remained unchanging across the decades: why do gay men speak the way they do? Researchers from backgrounds in sociology, physiology, linguistics, and psychology have taken a stab at understanding which factors might account for the phenomenon often referred to as "gay voice" and have always come up short. Sociolinguists have attempted time and again to unite a knowledge of social structures and processes with an understanding of language production and perception in order to explain this sociolect of sorts, but to little avail. The following discussion of five previous papers (four studies and one metaanalysis), aims to identify the specific features which account for the production and perception of sexual orientation through language use. Furthermore, we will explore gay men's speech in association to women's speech and I will offer my own ideas about this rich and nuanced topic.

2 Review of Previous Research

A brief look at American cultural anthropologist Rudolf Gaudio's 1994 study, "Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men," from *American Speech* will serve as a starting point for the journey to understand "gay voice". This paper, which Gaudio describes as an "initial attempt toward the study of gay male speech" (43), involves a study of the speech of eight American graduate students—half of whom are gay men, half of whom are straight men—as judged by thirteen other student participants. This study seeks to understand "the normative process whereby certain types of speech become identified as more masculine or feminine" (32) in relation to sexual orientation. Gaudio was particularly interested in pitch dynamism (pitch variation) as a possible contributor to the perception of a male speaker as gay. Gaudio criticises the lack of perception-based studies dealing with effeminateness, and illuminates the effects of listener judgements on somebody's own view of themselves as effeminate or masculine in his own research. This study is centred around four polar adjective scales typically related to the contrast in speech patterns between hetero- and homosexuals, particularly among men. The adjective scales, as assessed by the thirteen listeners, rated each speaker on how straight/gay, effeminate/masculine, reserved/emotional, and affected/ordinary their speech sounded.

The main finding of Gaudio's study was that listeners were able to correctly identify the sexual orientation of a speaker seven out of eight times. This serves as strong evidence that listeners have very accurate judgements on the sexual orientation of gay men from their speech alone. It does not, however, answer the all-important question of "why?" Although we can conclude from this study that the concept of a "gaydar" might be real, pitch variability, which is noted in female speech, cannot in isolation convince a listener of a speaker's sexual orientation. And so the search for speech qualities strong enough to affect this judgement in an experimental context continues.

Staying within the realm of perception-centred research, there is Erez Levon's 2006 study from the same academic journal, titled "Hearing 'gay': Prosody, interpretation, and the affective judgments of men's speech." Here, Levon provides an overview of four previous perception studies and their drawbacks, then sets out to conduct a new study using a completely different methodology wherein one male speaker's speech is manipulated in various ways (using Praat), and then each recording (four in total) is played for a large group of listeners. This study examines two prosodic variables—pitch range and sibilant duration—often associated with gay speech. Despite the deliberate design of his study and the extensive care he took to obtain clear and effective results, Levon failed to provide any notable findings in the perception of sexual orientation from speech. The manipulation of the speaker's pitch range and sibilant duration had no significant effect on listeners' perception of the speaker's sexuality. This study did, however, unveil significant correlations between speaker homosexuality and listeners' associations with the more positive qualities on the polar adjective scales. For example, perceived gayness was found to be strongly correlated to neatness and friendliness. Although the results of Levon's study showed no novel discoveries about the relationship between sexuality and language, his methodology is worth noting. Levon's approach to examining the perception of homosexuality from manipulated speech recordings established a promising starting point for the studies that followed.

Next, we have studies concerned with the production, rather than perception, of gay-sounding speech, with "Sharing resources and indexing meanings in the production of gay styles," by Robert J. Podesva, Sarah J. Roberts, and Kathryn Campbell-Kibler, published in 2002. This paper examines the speech of a gay, American lawyer's segment on a radio show in which he converses with a straight man about whether private organizations should be allowed to exclude gay people from membership. The study examined five phonetic variables and found that the most notable was the releasing of word-final stops, which occurred more frequently in the speech of the gay "Speaker A" than in that of his straight counterpart, "Speaker B." The authors of the paper urge readers not to jump to the assumption that releasing word-final stops is thus a characteristic of gay speech, but instead to see this variable in a wider context, explaining how this feature has also been linked to "geek girls," as well as Orthodox Jews. This phonetic quality possesses a "culture-wide relation to education or precision," (186) which may well have been a subconscious motive for Speaker A's use of it in the context of a radio debate. In this way, this paper reveals the need to study gay speech not only as compared to straight speech, but also—and perhaps more importantly—within the gay community, among its various intersectional members.

This study is valuable not only for its investigation into gay voice, but also in a wider sociolinguistic context, as it offers a unique approach to the notion of style. This study heavily emphasises the idea of identity as constructed by language, and suggests that the study of gay speech should consider the various subcommunities of gay men and with which goals its members use language. The speaker studied in this paper is a man who belongs to various social groups, and so his particular phonetic qualities cannot simply be attributed to his belonging to only one of these groups. Although Speaker A is a gay man, he is also a lawyer and an activist. In the particular context in which he is studied, his goals are most likely to come across as those of an intellectual with vast knowledge of the topic at hand, rather than simply as a gay man. The phonetic qualities identified in Speaker A's speech should therefore be further studied in different contexts, and among a more varied range of gay speakers to reflect the actual diversity of the community. The second production-based study of interest is that of Drew Rendall, Paul Vasey, and Jared McKenzie from 2008, titled "The Queen's English: An alternative, biosocial hypothesis for the distinctive features of 'gay speech'." This study examines the speech of roughly sixty gay speakers and sixty straight speakers (both men and women) from universities in Alberta. It looks at vowel formants from twenty-three vowel tokens, and also takes body size into account, a detail which sets this study apart. The introduction reviews three hypotheses for why homosexuals develop a way of speaking that is distinct from heterosexuals. This study is unique not only for providing findings on gay men and women's speech as opposed to their respective straight counterparts, but also for drawing a comparison between men and women overall, which may be key in understanding the production and perception of "gay voice". Gay men were found to have vowels that are slightly higher and shifted right in the acoustic space (i.e. higher F1 and F2 values) compared to those of straight men. This is a more modest version of the same difference observed between men and women's speech overall, and may therefore help us gain a better understanding of the stereotype that gay men speak like women–a notion clearly rooted in misogyny and implicit bias–but perhaps also a way of signalling belonging, solidarity, and allyship within and across the communities of women and gay men.

A complementary pattern was observed in gay women, who tended to have vowels that were farther down and left than straight women, which could contribute to a more masculine perception of their voices. Though the differences related to sexual orientation this study revealed were relatively small, they did nonetheless provide some clear answers to questions about the production of gay and lesbian speech. The "Queen's English" study thus stands apart from many of its perception-based counterparts, which fail to identify any notable differences between the speech of gay and straight people. This study also reaffirms what a few previous ones had discovered: namely, that a higher pitch, despite often being thought of as a "gay voice" quality, is not significantly correlated with homosexuality. In fact, the speakers with the deepest voices happened to be gay men.

This paper also offers the hypothesis that the same hormones that have been attributed to the development of sexual orientation might be responsible for the ways in which gay speech is distinct from straight speech. The degree of lip protrusion and retraction while speaking as well as the facial gestures and expressions that might co-occur with speech are potential causes of the difference in vowel-space position between gay and straight men and women. These attributes shorten or lengthen the vocal tract and thus create a difference in formant height. The authors touch on the correlation between lower frequency voices and "assertive behaviour," observed among humans and other species, in both the physical and social sense. They go on to cite a number of previous studies on body size, formant frequency, and assertive demeanors to form the basis of their biosocial hypothesis on gay speech: Gay

men may speak with less lip protrusion and higher frequencies which in turn reflects their less assertive demeanors and causes the formant value variation discussed earlier.

A final dive into research on this rich subject brings us to Deborah Cameron and Don Kulick's "Sexuality as identity: gay and lesbian language," a chapter from their book *Language and Sexuality*, published in 2003. This chapter provides a brief history of sociolinguistic research on gay language spanning from the 1940s to the end of the 1990s. With a focus on lexicon and surface structures, Cameron and Kulick examined the ways in which conscious word-choice, as opposed to phonetic features, can serve to construct gay identities.

Beginning in 1941 with Gershon Legman's lexicon from the back of a medical text on homosexuality, people have sought to understand what is often regarded as a sort of "secret language of identification" (Read 1980: 11) among gay men. By the 1980s, a large shift had taken place and a newer generation of gay people—and thus notion of gayness—emerged. This newer gay generation had no use for the older vocabulary Legman had documented in the 1940s, and often rejected users of this sort of gay slang because of the setback they threatened to the gay liberation movement. The language used by gay men in the 1980s was centred on "assertion, not concealment" (Cameron & Kulick 2003), and language was thus used openly to distinguish members of various emerging gay subcultures. The more modern idea of gay language was courtesy of Keith Harvey's establishment of a grammar for what he calls "camp talk," wherein speakers use "paradox, inversion, ludicrism, and parody" (99) to present a gay identity. The subversion and juxtaposition of language among many gay speakers show how far the gay vernacular has managed to come in the short period between Legman's time and that of the 1980s gay liberationists, indicating just how dynamic this community is. This chapter also provides insight into the absence of research on lesbian speech and the unfortunate fact that only the easily identified "Gayspeak" has been able to capture the interest of most sexuality-based sociolinguistic research. Pairing sociolinguistic research with Foucauldian insight into sexuality and identity, Cameron and Kulick contextualise the fascination with gay speech while providing various insightful perspectives on gender and sexuality in relation to language.

3 Discussion

In reading these five pieces of research, as well as having a look at Bruce Rodgers' *The Queen's Vernacular: A Gay Lexicon*, I find myself increasingly interested in today's gay slang, which I hypothesise may be salient in non-gay social contexts now more than ever. Thanks to popular reality TV shows centred around gay men, such as *RuPaul's Drag Race* and various fashion competition shows, and

language-diffusing platforms such as TikTok, the queer community, especially when dominated by men, seems to me to be a presiding source of slang words among young people (primarily women, who play a large role in language diffusion). This is unsurprising due to the long-standing cultural influence of queer communities such as the African-American and Latino underground ballroom scene which no-doubt shaped much of our modern popular culture. On social media, young women often employ slang terms like "slay," "purr," "boots," "queen," and a range of other gay slang terms comfortably and frequently. It seems these terms have become solidarity markers for a broader online community, sometimes referred to as "the girls and the gays" (young women and the younger queer community). This emerging community, whether it be real or merely aspirational, serves to unite its members, perhaps over a shared sense of oppression that sets both groups apart from the dreaded cisgender, straight, white male. Although it is harder to notice a convergence of non-lexical speech qualities among "the girls and the gays", I do feel that the Kardashian-esque, "valley girl" way of speaking comprises qualities which have come up frequently in my research on gay speech. Whether it be lengthened sibilants (associated with lisping), higher fundamental frequency, or greater pitch variation ("swoopy voice"), many phonetic features are characterised as effeminate, and are shared by women and gay men. Rejecting the assumption that gay men and straight women simply imitate each other, I subscribe to Gaudio's understanding that gay men "have particular ways of speaking which challenge conventional notions of what constitutes proper male and female behaviour" (Gaudio 1994: 32) and further extend it to the speech of young women, who also have much to gain by their intentional use of language.

I believe all five aforementioned pieces of gay speech research to be fundamental in the study of sexual orientation, gender, and identity, which are all constructed— as well as influenced—by language. I adopt a bi-directional view of the influence linking language and identity, since anything less open than this might exclude some vital component in the uncovering of how language and identity work both discretely and in tandem. Though Cameron and Kulick (2003) reject the identity-based approach to this field of study, stating that it "compels us to circumscribe inquiry," I feel that sexual identity does not exclude but rather encompasses the "fantasy, repression, pleasure, fear and the unconscious" (105) that are of great concern to Cameron and Kulick. I am also in agreement with the perspective of Podesva, Roberts, and Campbell-Kibler (2002) that identity is intrinsically linked to style and that "the individual negotiates identity across situations" (187), leading to a rich and diverse inventory of possible gay speech styles, such as that of the gay lawyer and activist whose speech they studied. Their new definition of style as "the linguistic means through which identity is produced in discourse" (179) is, in many ways, more valuable to the field of sociolinguistics than the historical understanding of style as speakers' formality level. Furthermore, the idea that "identity and style are co-constructed" (179) deserves attention in

sociolinguistic research of all sorts, especially in relation to the "processes through which power accomplishes [the] production [of identities]" (Cameron & Kulick 2003: 78).

The studies that stood out the most were those that recognised the importance of abandoning any notion of a homogenous community of gay speakers. Since gay sociolinguistics is a relatively young field of study, it might be undergoing a process similar to that which shaped sociolinguistics in the 1960s. Much in the same way as Chomsky assumed a general homogenous speech community in the early days of sociolinguistic research, gay sociolinguistic research seems to have often relied on an archaic idea of a static and homogenous gay community. These assumptions are lazy and accomplish nothing, and so must be abandoned, as many of these papers suggest. Podesva, Roberts, and Campbell-Kibbler (2002) provide a solid starting point for re-evaluating sociolinguistics as a whole in order to attain more valuable results in the studies of speech communities as large and diverse as that of gay men, since this step is paramount to understanding "gay voice." Levon (2006) shares Cameron and Kulick's (2003) view that typical methodologies for gay sociolinguistic research "ignore the reality that linguistic practice is highly variable, both within and between speakers" (Levon 2006: 61). Rendall, Vasey and McKenzie (2008) also stands out to me as a bold study of gay speech which asserts the idea that sexual orientation has some biological implications— namely, certain hormones present during puberty— and that this should not be ignored in the search for an explanation for "gay voice". The existence of "attempts by [gay] speakers to project through a body size that seems larger or smaller than they truly are" (190) is a particularly interesting idea that exemplifies a speaker's own intuitions about their physical qualities in relation to language-use and definitely warrants further research. Among the various suggestions for future research on gay speech, I agree with Rendall, Vasey and McKenzie (2008) that "body size considerations should be routinely integrated" (200) into research on homosexuality and speech production.

4 Conclusion

In this foray into research on the sociolinguistic phenomenon of gay speech, various potential qualities for perceiving sexual orientation from speech alone have been proposed. Since no single feature can apparently influence listeners' judgement of a speaker as either gay or straight, it is clear that this is a nuanced process of both production and perception that warrants investigation from a range of social science perspectives. Since the focus within this field of study shifts from archaic lexicons to speaker judgements using polar adjectives, then on to speech manipulation and various more advanced forms of research, it is safe to say this is a rich and worthwhile topic of study. The widespread use of social media and the consumption of queer media (or at least media with queer origins) have vaulted the question of

how gay identities are constructed using language into a place of more pertinence than ever. The five studies on gay speech examined above have highlighted how "perceptions of sexuality are ideologically linked to other perceptions of personality and personhood" (Levon 2006: 73), which solidifies the importance of language's ability to "construct particular identities" (Cameron and Kulick 2003: 102). Among the various potential candidates for markers of gay speech, formant height variation (Rendall, Vasey and McKenzie 2008), pitch dynamism (Gaudio 1994; Levon 2006), and sibilant duration (Gaudio 1994; Levon 2006; Podesva, Roberts & Campbell-Kibbler 2002), have stood out as possible components that could, in sum, constitute the ever-elusive phenomenon known as "gay voice." In future studies on this topic, much could be gained by examining gay speech in the context of social media and popular culture among teenagers, as well as the relationship and commonalities between gay speech and that of "valley girls", as well as of young women more generally.

References

Cameron, Deborah, and Don Kulick. Language and sexuality. Cambridge University Press, 2003.

Gaudio, Rudolf P. "Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men."

American Speech, vol. 69, no. 1, 1994, pp. 30–57.

- Levon, Erez. "Hearing "gay": Prosody, interpretation, and the affective judgments of men's speech." *American speech* 81.1 (2006): 56-78.
- Podesva, Robert J., Sarah J. Roberts, and Kathryn Campbell-Kibler. "Sharing resources and indexing meanings in the production of gay styles." *Language and sexuality: Contesting meaning in theory and practice* (2002): 175-189.
- Read, Kenneth E. Other voices: The style of a male homosexual tavern. Chandler & Sharp Publishers, Incorporated, 1980.
- Rendall, Drew, Paul L. Vasey, and Jared McKenzie. "The Queen's English: An alternative, biosocial hypothesis for the distinctive features of "gay speech"." Archives of sexual behavior 37 (2008): 188-204.
- Rodgers, Bruce. The queens' vernacular: A gay lexicon. [San Francisco, Calif.]: Straight Arrow Books, 1972.

licensed under <u>CC BY-ND 4.0</u>

Sounding Dyke-ish: Demystifying the Conventions of Lesbian Speech

Natalia Ventura Carrero

Department of Linguistics, McGill University LING 320: Sociolinguistics 1 Professor Charles Boberg

Abstract

The domain of queer studies as it pertains to sociolinguistic research has been predominantly concerned with analyzing the speech of gay men, while very little has been said about the nature of lesbian speech patterns. This failure to consider the speech of gay women as theoretically significant is a direct product of male supremacy as well as the unique combination of oppressive forces that ostracize lesbian existence, including misogyny and homophobia. This paper navigates a linguistic-anthropological approach towards dissecting the particulars of lesbian speech, drawing from a number of quantitative studies in tandem with queer theory. Data collected on phonetic and stylistic variables such as pitch, intonation, articulation, and politeness elucidate the complexities of lesbian speech habits, while shedding light on the phenomenon of "enforced invisibility" that hinders further research on the matter. Qualitative analyses on the social ramifications of perceived lesbian speech contest the reductive notion that lesbians merely seek to mimic the speech of heterosexual males, suggesting that the subversiveness of lesbian speech is inherent to the perceived deviance of the identity.

1 Introduction

Much of the literature concerned with the production and perception of markedly queer speech have dealt exclusively with the speech of gay men, while very little choose to focus on lesbian speech. This disproportionate attention, rooted in patriarchy, is indicative of the discomfort surrounding the subversion of gender roles as it manifests in lesbianism. To the extent that traditional womanhood is necessarily anchored to the preservation of male dominion, the romantic and sexual exclusion of men foundational to the lesbian identity necessarily threatens this social order. Given this deviance, the few sociolinguistic studies that have attempted to identify markers of lesbian speech often default to the assumption that lesbian speech is simply an elaborate imitation of male speech patterns. This reductionist assumption, yet to be proven, fails to capture the nuances of lesbian identity which aim at subverting traditional notions of femininity by eradicating their ties to heteronormative expectations, rather than simply replicating a male dialogue. To the extent that lesbian existence is characterized by its own subversion, the linguistic apparatus must reflect this atypicality. Utilizing various quantitative studies on the matter, in tandem with scholarship in queer theory, this paper navigates a linguistic anthropology approach towards identifying the parameters and manifestations of lesbian speech, its origins, and its perception within the heteronormative matrix.

2 Discussion

The Birch Moonwomon-Baird's paper, published in the 1997 anthology Queerly Phrased, reflects on the complexities surrounding a quantitative evaluation of a lesbian speech style, as well as the politics of lesbian identity. Moonwomon-Baird uses the term "enforced invisibility" to describe the silencing of queer experience, which works two-fold on lesbians who, as women, are already positioned on the periphery (Moonwomon-Baird 1997: 204). They argue that this expectation of submissiveness and quietness expected of both women and queer individuals compounds into the inaudibility of lesbianism, in that lesbians are not only depicted as deviants, but are ultimately not supposed to exist. This intersectional perspective of systematic oppression serves as an explanation as to why the topic of lesbian speech is underrepresented and under analyzed, and why, therefore, lesbians are "hard to hear" (Moonwomon-Baird 1997: 205).

Moonwomon-Baird references their earlier 1983 study in which they tested for a correlation between listener's perceptions of a speaker's sexual orientation and the willingness of listeners to, consequently, identify particular speakers as lesbians in tandem with other marginalized identifiers that encompass linguistic variables associated with "unladylike" speech (Moonwomon-Baird 1997: 206). The participants of this study included six heterosexual women and six lesbian women, of which all twelve were white, native speakers of North American English. The listener pool was composed of 21 students at U.C. Berkeley who were all also native speakers of North American English. The latter group listened to 30-second stretches of the pre-recorded speech of the heterosexual and lesbian speakers and were asked to answer a questionnaire in which they identified various social and vocal characteristics based on their own judgment. The questions on sexual orientation were enmeshed with other social characteristics such as class, age, educational background, region of upbringing, and ethnicity, while the selected voice traits involved rate of speech, pitch, amplitude, and forcefulness. Analysis on the correlation between selection of "lesbian" and selection of other traits revealed that listeners were reluctant to identify certain traits, such as "lesbian". Additionally, when such traits were identified, they were much more likely to be selected alongside other "marked" traits such as "Jewish" (Moonwomon-Baird 1997: 208). MoonwomonBaird summarized these findings to conclude that, although accuracy in listener's perception was not particularly meaningful, this aforementioned reluctance was reflective of a stigmatization of identifiers such as "lesbian" and "Jewish". They remark further that, albeit the study dealt with a small sample size, listeners did in fact have some vague sense of what lesbian speech ought to sound like. While this preconception did not translate into selection for this particular trait in the questionnaire most of the time, it proved reflective of the discomfort lesbian presence is prone to invoke.

In the second 1984 study expanded upon in their article, Moonwomon-Baird focuses on intonation patterns and how certain speech strategies come to be exhibited by heterosexual women in contrast to lesbians. In this study, conversations between two pairs of women, one lesbian and one heterosexual, were recorded and an analysis on acoustic contours such as direction of pitch, steepness of rises and falls, and glides were to be indicative of the ways in which women executed gendered behaviour in relation to one another. Moonwomon-Baird found that, although a great deal of individual variation was discovered, the lesbian pair tended more towards one trend than another, and vice-versa for the heterosexual pair. In particular, the lesbians utilized a smaller intonation range than the heterosexuals, as well as lower and narrower pitch ranges. The heterosexual pair revealed increased usage of steeper glides, which coincided with speech strategies deployed primarily by women. Overall, this study elucidated the effect of conversational roles on gendered performance in speech, and how certain strategies that manifest in particular vocal traits are reflective of situational and communicative needs (Moonwomon Baird 1997: 210-211).

Auburn Barron-Lutzross' 2015 paper on the production and perception of a lesbian speech style combined two experiments; the first recorded and analyzed the speech of lesbian, bisexual, and straight women, while the second tested for a lesbian speech stereotype. Barron-Lutzross made use of a larger pool of speakers, as well as a larger sampling of speech recording, aiming to provide more depth to this area of study that has previously been compromised by small numbers. This study also included supplementary variables in addition to sexual orientation, such as how many friends a speaker has who identify as homosexual, and how familiar a speaker is with queer culture. The phonetic variables studied relative to these social variables included those typical to previous works, such as pitch, vowel formants, and fricative variability, as well as speech rate, word-final /t/ release, and creaky voice .

All 54 speaking participants were native speakers of North American English; 12 identified as homosexual, 18 as bisexual, and 24 as heterosexual. Ages ranged from 18-54, with most falling between the range of 18-21. From a predetermined list, speakers were instructed to read aloud monosyllabic and multisyllabic words at random, followed by increasingly longer utterances, sentences, and questions, with

this methodology serving to contrast hypothesized variation in broken versus continuous speech, and its effect on structure. The results revealed no direct correlation between the sexual orientation of any given speaker with any of the aforementioned linguistic variables and did not indicate any explicit characteristics of a lesbian speech identity (Barron-Lutzross 2015: 20-26).

A correlation was identified between the deployment of these linguistic variables and a given speaker's familiarity with queer culture; in particular, a higher familiarity yielded generally lower median pitch and faster rate of speech (Barron-Lutzross 2015: 30-31). This correlation was only meaningful with the speech of straight speakers, leading Barron-Lutzross to theorize that straight people might attempt to reinforce their association and alignment with queer culture through certain linguistic tendencies (Barron-Lutzross 2015: 31). In other words, straight speakers may be unknowingly deploying marked characteristics of a lesbian speech style in order to signify a queer affinity.

In the second experiment of the study, the focus shifted to an attempt to pinpoint a phonetic stereotype of lesbian speech, using the above variables as markers of judgment. The same speakers read from the same list of words as in the previous experiment, and listeners were instructed to rate subjects on three scales in two separate divisions: "very educated" to "very uneducated", "very formal" to "very casual", "very masculine" to "very feminine", "very shy" to "very outgoing, "not at all likely to be a lesbian" to "very likely to be a lesbian", and "very compassionate" to "very uncaring" (Barron-Lutzross 2015: 34). The results showed that listeners were generally accurate in rating a speaker's sexual orientation, and straight speakers were significantly less likely to be judged as sounding like a lesbian than lesbian speakers (Barron-Lutzross 2015: 35). Sexual orientation also correlated with the other ratings; straight speakers were significantly more likely to be rated as feminine than their lesbian counterparts, and lesbians in contrast were more likely to be rated as uncaring.

With regards to uncovering a blueprint for a lesbian speech stereotype, among those speakers who were rated most likely to be a lesbian, correlates included a lower median pitch, a lower F2 (or vowel backness), and a higher proportion of creaky voice. Those rated higher on the scale of femininity exhibited longer vowel durations and less creaky voice. These findings point to an unconscious internalization and conception of what ought to characterize lesbian sounding speech and affirm the hypothesis that there is a metric upon which listeners are able to discern and make generally accurate judgments about speakers' sexual orientations (Barron-Lutzross 2015: 47). Additionally, the ratings on sexual orientation were not only consistent with a specific speech style, but were accurate more than half the time, suggesting that lesbian speakers do in fact adopt certain traits of a lesbian speech style.

In his 2007 study on the acoustic correlates of perceived masculinity, femininity, and sexual orientation, Benjamin Munson provides insight on the ways in which gendered norms are exercised through language, and how listeners perceive adherence or deviance from this binary norm. In particular, the study comprised an exploration of the independence of perceived masculinity and femininity to that of perceived sexual orientation, and how the acquisition of a queer speech style ought to be more complex than simply resulting from sex-atypical ways of speaking (Munson 2007: 129). The participants consisted of 44 speakers, male and female, originally involved in Munson's earlier 2005 study, while the listeners were 10 adults who did not participate in this earlier study. All were native speakers of English. Listeners provided ratings based on acoustic measures on a scale of 1-5, with 1 denoting "definitely sounds heterosexual", 3 indicating "sounds neither GLB nor heterosexual", and 5 indicating "definitely sounds GLB" (Munson 2007: 130-131). Male and female speakers were rated separately along the axes of masculinity and femininity, respectively. Overall, self-identified lesbian and bisexual women were rated as sounding less feminine than their heterosexual counterparts, however, evidence points towards a distinction between perceived femininity versus perceived sexual orientation in women's speech, i.e., these two axes acted independent of one another. For instance, women were rated as less-feminine sounding when ratings were produced in response to words with sibilant fricatives (such as gas, said, and some) as opposed to words lacking these sounds, yet this same factor did not affect judgments of perceived sexual orientation. From these findings, Munson concludes that constructions of perceived sexual orientation and perceived masculinity/ femininity are not identical, but rather, may be related to unique acoustic parameters respectively (Munson 2007: 138). Analyses revealed that average F0, or fundamental frequency, was more strongly correlated to perceptions of masculinity and femininity, while average F1, or vowel height, was more strongly linked to judgments of perceived sexual orientation. The overlap revealed between these dimensions indicates that queer people are more likely to be identified as gay or lesbian when they exhibit speech characteristics typical of that of the opposite sex, and that it is a general tendency of queer speak to envelop patterns atypical of gendered expectations. Munson also adds nuance to these gendered subversions and argues alongside Moonwomon-Baird that where there is a lack of a lesbian speech stereotype, general listeners may default to assuming that the speech of lesbians must be similar to that of men.

In their groundbreaking 2006 book "Gender Trouble", Judith Butler deconstructs the prescriptive notion of gender as inherent to one's biological sex and argues that gender is not what someone necessarily is, but rather is something that one consciously performs. Butler asserts that "gender is the repeated stylization of the body, a set of repeated acts within a highly rigid regulatory frame that congeal over time to produce the appearance of substance, of a natural sort of being." (Butler 2006: 45) In

alignment with the studies previously mentioned, if gender constitutes a performance, then the linguistic variables associated with an ideal masculinity or an ideal femininity are variable upon the individual exercising them. If gender is to be understood as nothing more than a socialized repetition of discrete acts, then the suggestion that "one is not born a woman, but, rather, becomes one" by Simone de Beauvoir necessitates a cultural and social expectation (Butler 2006: 11). De Beauvoir further dissects the latter's analyses, arguing that women occupy a deficit within a masculine economy, and therefore inherently elude the gendered dyad in absence of a definition (Butler 2007: 13). In the case of lesbian lived experience, an identity founded upon the exclusion of men, the lesbian ought to constitute a "third gender" that operates beyond the boundaries of sex and gender (Butler 2007: 26). If womanhood is constructed in such a way that is complementary to the centrality of proper masculinity, then lesbianism cannot fit neatly into this dichotomy. Whether lesbianism constitutes a third gender becomes a grey area of discourse; however, when analyzed through the lens that gender is a conscious enactment of prescribed norms, speech as a critical component of social performance becomes the vehicle through which the lesbian modality asserts itself. Operating within a binary that draws strict divisions between masculinity and femininity renders it increasingly difficult to define the parameters of lesbian speech if it is to be acknowledged that the lesbian necessarily operates and sustains herself beyond these gendered boundaries.

Robin M. Queen's article on locating lesbian language, published in the anthology Queerly Phrased, contends that lesbian speech is a broad amalgamation of several stylistic conventions, and is not necessarily derived from an approximation of traditionally feminine or masculine ways of speaking. Queen draws upon comic-book portrayals of lesbians to extrapolate the assumptions underlying these media representations into the broader social arena. Characterization of a particular trope, Queen argues, necessitates a real world anchoring upon which specific traits must serve to differentiate a lesbian character.

Queen identifies four stylistic tropes, based on personal observation as well as analyses on the characterization of lesbians in comic strips, that constitute a unique lesbian language: stereotyped women's language, stereotyped nonstandard varieties, stereotyped gay male language, and stereotyped lesbian language (Queen 1997: 240). These tropes serve as a linguistic umbrella, from which lesbians deploy specific features in various combinations that serve to recontextualize their associations in a distinctly lesbian light. In the first of these four tropes, Queen gives particular attention to the *butch* and *femme* dichotomy, a categorization of lesbian identities which encapsulates respective negative and positive appropriations of stereotyped women's language, in particular: "Lesbians may use structural elements that do not conform to stereotyped women's speech in order to distinguish themselves from the

stereotyped woman, or they may use particular aspects of stereotyped women's language in order to index their identity as women." (Queen 1997: 241). Queen makes the case that the masculine and feminine constructions of *butch* and *femme* identities are not simply due to an imitation of traditionally male or female speech, but rather an appropriation of these stereotypes that reinforces a queer relationality. This notion, in tandem with Munson's study on the perception of femininity and masculinity in speech, disproves the reductive assumption that lesbian speech is simply a homogeneous approximation of male speech. While lesbian speech typically rejects markers of women's speech such as elaborate adjectives, politeness, and increased intonation, it may also make use of these markers in different social scenarios to index some approximation to womanhood.

The second of these four tropes draws primarily upon Labov's 1972 analyses on the speech of working-class males and the use of non-standard varieties such as [in'] versus [ing], postvocalic /r/ deletion, contracted forms such as "gonna", and profanities (Queen 1997: 240). Queen notes that, unlike stereotypically male speakers, the lesbian characters in the comic strips they analyze use these features variably and arguably in more subtle applications. For instance, where male characters in the comic strips use [in'] consistently, lesbians may use the standard form in some cases, and where male characters exhibit vowel shifts such as "fooken" (fucking) and "yiz" (plural form of 'you'), lesbians do not (Queen 1997: 244). This contrast dispels the idea that lesbians aim to replicate the speech of men by showcasing a perpetual linguistic fluctuation that cannot be predicted nor categorized. Third, deployment of stereotypically gay male speech such as specific lexical items and an H*L intonational contour (high tone to low tone pitch accents), disrupts the idea of a continuum of female and male speech patterns upon which lesbians are assumed to operate. The interjection of increased exclamation and lexical items such as "fabulous" demonstrate an indirect approximation to markedly feminine speech traits that have been co-opted by gay males, and in turn a uniquely queer association (Queen 1997: 248). Finally, the fourth trope is identified as operating beyond the stereotypes of male and female speech as well, involving "hyperarticulation, with little pitch modulation or apparent language play" as well as "little or no nonstandard phonology" (Queen 1997: 252). These features are in alignment with Moonwomon-Baird and Barron-Lutzross' analyses which revealed the prevalence of a lower and narrower pitch range in lesbian speakers, supplementing the validity of the claim that some markers of lesbian speech are indeed uniquely lesbian.

Additionally, Queen suggests that the difficulty surrounding an identification of lesbian speech patterns in the dominant literature is due to the sheer complexity and ambiguity surrounding the notion of a "lesbian community". She draws parallels with several sociolinguistic scholars, including Labov 1972, who emphasize the centrality of a "speech community" in order to make claims about the social use of
language (Queen 1997: 235). Moonwomon-Baird's argument of enforced invisibility coupled with Butler's deconstruction of gender reinforces Queen's argument that the very attempts at categorizing a lesbian community proves not only meticulous but futile. By extension, efforts at identifying particular traits that may be exclusive to such a lesbian speech community proves insurmountable. Through this, then, the argument is made that scholarship surrounding the topic of lesbian speech should not be grounded in the centering of an ideal community, but rather in a "linguistics of contact", as coined by Mary Louise Pratt (1987) (Queen 1997: 238). This framework, rather than emphasizing membership as the genesis of lesbian marked traits, denotes the unique exploitation of language in specific relational contexts that indexes a lesbian identity, similar to Moonwomon-Baird's second study concerned with the deployment of gendered speech in distinct conversational settings. It is the very heterogeneity of an assumed lesbian community that takes precedence in this model, and the usage of marked linguistic tools based on this diversity that foregrounds lesbian social cohesion, whether conscious or unconscious.

3 Conclusion

In "Subversive Bodily Acts", the third section of their 2006 book, Butler contends with Monique Wittig's views on lesbian existence in which the latter claims that a lesbian transcends the binary opposition of woman and man. In particular, it is through such non-conformance that the lesbian exposes "the contingent cultural constitution of those categories and the tacit yet abiding presumption of the heterosexual matrix." (Butler 2006: 153) The various studies explored here illustrate the hindrance that an adherence to the gender binary poses to an understanding of lesbian identity and expression. The very discomfort with lesbian existence is what plagues research on the matter. Simultaneously, it is through the very social ostracization that lesbians experience that they construct their visibility. Therefore, if an authentic panorama of lesbian speech is to be developed, it must break free of the framework that foregrounds the speech of men and women as a theoretical yardstick. The lesbian's inherent deviance and refusal to conform is what characterizes her speech, and while certain linguistic features typically associated with men and women are certainly borrowed, it is her tongue that transforms them into utterances that are uniquely lesbian.

References

- Barron-Lutzross, A. (2015). The Production and Perception of a Lesbian Speech Style. UC Berkeley Phonology Lab Annual Reports, 11, 1-68. https://doi.org/10.5070/p77f6332bh
- Butler, J. (2006). *Gender trouble: Feminism and the Subversion of Identity*. Routledge. https://doi.org/10.4324/9780203824979
- Moonwomon-Baird, B. (1997). Toward a Study of Lesbian Speech. In A. Livia & K. Hall (Eds.), *Queerly Phrased: Language, Gender, and Sexuality* (pp. 202-213). Oxford University Press.
- Munson, B. (2007). The Acoustic Correlates of Perceived Masculinity, Perceived Femininity, and Perceived Sexual Orientation. *Language and Speech*, 50(1), 125-142. <u>https://doi.org/10.1177/00238309070500010601</u>
- Queen, R. M. (1997). "I Don't Speak Spritch" Locating Lesbian Language. In A. Livia & K. Hall (Eds.), *Queerly Phrased: Language, Gender, and Sexuality* (pp. 233-256). Oxford University Press.