LSURC

Language Sciences Undergraduate
Research Conference

7th Annual Conference

MEDIUMS OF COMMUNICATION
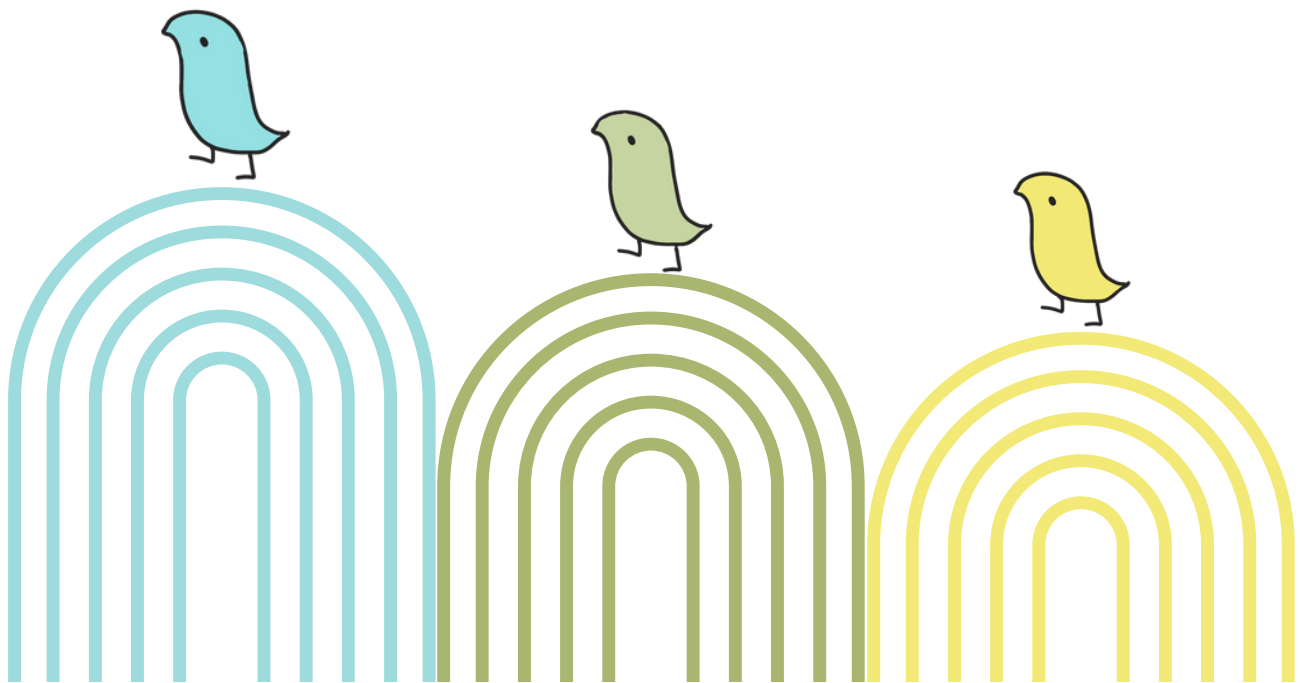
FEBRUARY 17-18, 2023

The University of British Columbia
Hybrid (In-Person and Online)

LSURC
Language Sciences Undergraduate
Research Conference

TABLE OF CONTENTS

ABOUT LSURC

We are proudly partnered with the UBC Language Sciences Initiative and UBC's Speech and Linguistics Student Association (SALSA), who tirelessly support our committee's efforts to organise this event. This year we are proud to add the University of Calgary's Verbatim to our family, and we encourage everyone to submit and attend their hybrid colloquium, Saturday April 1st 9am-3pm MST.

Established in 2017 by a committee of five Linguistics and Speech Science undergraduate students, the Language Sciences Undergraduate Research Conference (LSURC) was formed with the goal of creating an opportunity for undergraduate students to explore language and communication research with an interdisciplinary perspective.

While inaugurated at UBC Vancouver, we attempt to expand our reach to undergraduate students studying in other disciplines and at other institutions in order to facilitate undergraduate engagement with language research.

Our multidisciplinary conference involves input from various post-secondary institutions across the globe. In particular, we have worked together with Linguistics student committees from UVic's Underlings & UCalgary's Verbatim and are extremely grateful for the effort that has been put into this two-day event. The LSURC planning committee began working on this conference since LSURC 2022 wrapped up and is excited to have you all attend.

2023 CONFERENCE THEME:
MEDIUMS OF COMMUNICATION

Dynamic and ever-changing, communication is a ubiquitous force that shapes our understanding of ourselves and transcends the boundaries of our limited experiences to share our collective perceptions of the world. From the expansion of face-to-face meetings and phone calls to texting and Zoom meetings, one thing remains constant: these methods of communication connect humanity across space, time, and cultures. From individuals' use of voice, body language, and expression to large-scale media networks, let us explore communication in its forms across all time.
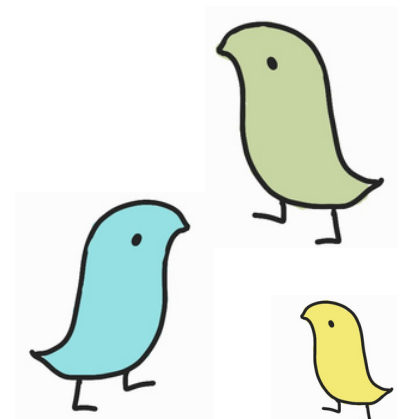
REVIEWERS

Dr. Amanda Cardoso

Arian Shamei

Dr.Christopher Hammerly

Ezra Yu

Dr. Hee Yeon Im

Katharine Davis

Dr. Kathleen Currie Hall

Dr.Lisa Matthewson

Dr.Márton Sóskuthy

Dr. Miikka Silfverberg

Rachel Soo

Dr.Stefan Dollinger

Suyuan Liu

ADJUDICATORS

Dr. Christopher Hammerly

Ezra Yu

Dr. Hee Yeon Im

Dr. Janet F. Werker

Professor Nigel Howard

Dr. Stefan Dollinger
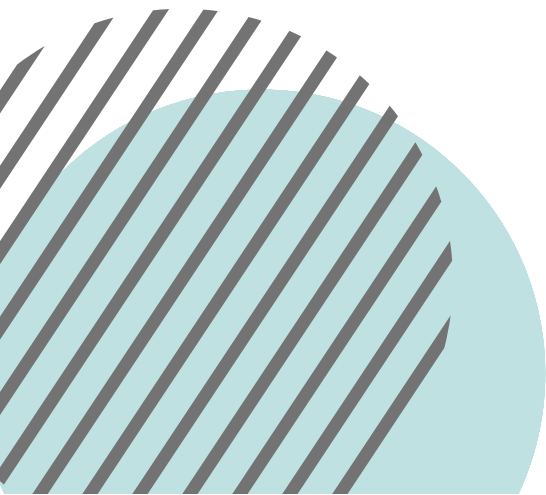
CONFERENCE SCHEDULE

DAY 1: FRIDAY, FEBRUARY 17TH	
3:00 - 4:00	UBC Emerging Media Lab Tour (Drop-Ins) IBLC 183
5:00 - 5:30	Arrival (IBLC 182) and join the Zoom and ask any questions (Zoom: QR code on page 33)
5:30 - 5:45	Opening Address
5:45 - 6:45	**Speaker Session 1**
5:45 - 6:00	Reversing a Cantonese tone-merger through conversation tasks. *Presenter(s): Samuel To & Ivan Fong (SFU)*
6:00 - 6:15	How to Cut Long and Rich-Format Text, Correctly, Once and for All *Presenter(s): Ziya Wang, Eunkyul Leah Jo, Grace Tianjiao Zhang & Angela Yoonseo Park (UBC)*
6:15 - 6:30	What lexical variable (vacation) may tell us about the Canada-US border and the role of multilinguals in Metro Vancouver and beyond *Presenter(s): Sophia Abundo & Sarah Lefebvre (UBC)*
6:30 - 6:45	Stories and Screens: Mother and child behavior during shared reading of electronic books with and without narration (virtual presentation) *Presenter(s): Sofia Cordeiro (UBC)* - Virtual
6:45 - 7:00	15 min break
7:00 - 8:00	**Speaker Session 2**
7:00 - 7:15	Social perception of [n] and [l] initials in Cantonese *Presenter(s): Sandy Chow (UBC)* - Virtual

Time	Event
7:15 - 7:30	Does Delivery Matter? *Presenter(s): Charlize Ma, Effie Kao, Stephanie Wang & Raechel Kitamura (UBC)*
7:30 - 7:45	Benchmarking the corpus annotation scheme for French NER *Presenter(s): Grace Yang (UBC)*
7:45 - 8:00	Variability on vowel production for three-year-old Spanish-English bilingual *Presenter(s): Elizabeth Alvarez-Alfaro (SFU)*
8:00 - 8:15	Closing Remarks
8:15 - 9:00	UBC Slam Poetry The Boulevard Coffee Roasting Co.	5970 Univeristy Boulevard

DAY 2: SATURDAY, FEBRUARY 18TH

Time	Event
9:00 - 9:30	Arrival (IBLC 182) and join the Zoom and ask any questions (Zoom: QR code on page 33)
9:30 - 9:35	Introduction to UBC Speech and Linguistics Student Association (SALSA)
9:35 - 9:45	Introduction to UBC Language Sciences Coordinator: Ella Fund-Reznicek
9:45 - 11:05	**Plenary Speakers Session 1 (Zoom link: see website)**
9:45 - 10:25	Word order or world knowledge? The role of early language on resolving conflicting cues by deaf individuals *Dr. Qi Cheng (Neuroplasticity & Language Lab)	Department of Linguistics, UW)* - Virtual

10:25 - 11:05	Aikamotsiipohtoo'p: Blackfoot language revitalization with families and in the home *Dr. Heather Bliss (Department of Linguistics, SFU) & Annette Fox-BruisedHead* - Hybrid
11:05 - 11:30	Break and Independent Poster Browsing and Boothing (https://blogs.ubc.ca/lsurc/) (Some presenters available for Q&A via a Zoom link posted below their poster)
11:30 - 12:10	**Plenary Speakers Session 2 (Zoom link: see website)**
11:30 - 12:10	Social Perception and Evolution in the use of Signed Languages *Professor Nigel Howard (Department of Linguistics, UBC)*
12:10 - 1:00	Lunch Break
1:00 - 2:00	**Speaker Session 3 (Zoom link: see website)**
1:00 -1:15	Metalinguistic Awareness in Young Children and Toddlers *Presenter(s): Julianne Bittante (SFU)*
1:15 - 1:30	Phonetic variation and perceptual processing of code-switching *Presenter(s): Khushi Nilesh Patil (UBC)*
1:30 - 1:45	Language Learning in a Global Babywearing Community of Practice *Presenter(s): Katherine Brand (UBC-O)* - Virtual
1:45 - 2:00	Part of Speech Tagging for Levantine Arabic *Presenter(s): Saughmon Boujkian (UBC)*
2:00 - 2:30	Break and Independent Poster Browsing (https://blogs.ubc.ca/lsurc/) (Some presenters available for Q&A via a Zoom link posted below their poster)

2:30 - 3:30	**Speaker Session 4 (Zoom link: see website)**
2:30 - 2:45	My Pronunciation is Éxcusable: A Diachronic Look at Stress Placement with English Derivational Suffixes *Presenter(s): Charys Russell (UCalgary) -* Virtual
2:45 -3:00	Indigenous Language Identification *Presenter(s): Yee Siong Pang (UBC)*
3:00 - 3:15	Measuring brain response when exposed to infant-directed speech: A meta-analysis *Presenter(s): Jaimie Muller (UBC)*
3:15 - 3:30	Environmental Entropy as a Means to Further Explore Child Bilingualism. *Presenter(s): Farzana Ali (SFU)*
3:30 - 3:45	Introduction to University of Calgary Underlings *Presenter: Mina Guan* 15 min break
3:45 - 4:15	Awards Presentation & Closing Address

Day 1: Friday, February 17th

Reversing a Cantonese Tone-merger Through Conversation Tasks

### Presenter(s): Samuel To & Ivan Fong (SFU)

Keywords: phonetic adaptation; conversation task; puzzle-solving task; heritage and native Cantonese speakers; Cantonese tone-merger

Cantonese has six contrastive lexical tones. Some of these tones have been observed to be undergoing a tone merger (Mok et al., 2013; Fung & Lee, 2019). This refers to a phenomenon whereby the distinctive contrasts between two tones are blurred. Tone merging can occur in one or both of speech perception and speech production (Fung & Lee). Phonetic adaptation, also referred to as conversational adaptation, phonetic accommodation, or conversational accommodation, is a phenomenon where interlocutors alter their speech pronunciation, either convergently or divergently, in order to facilitate communication and intelligibility (Lewandowski & Jilka, 2019). Past research has investigated phonetic adaptation through the use of shadowing (imitation) tasks (Pardo et al., 2017; Lin et al, 2021), map (wayfinding) tasks (Pardo, 2006), matching tasks (Dias & Rosenblum, 2011), and the Diapix (difference-finding) task (Baker & Hazan, 2011; Lee et al., 2011; Lewandowski & Jilka, 2019). Of these, only the latter two tasks create interactions that attempt a balanced amount of utterances between interlocutors, while also attempting to create natural conversational interactions that do not rely on imitation tasks. Accordingly, the task used in the present study draws from both of these tasks.

The present study asks whether non-imitation based tasks can cause adaptations of suprasegmentals, such as tones, through natural conversations. Lin et al. (2021) investigated the reversal of a tone merger through a shadowing task. The present study aims to investigate this adaptation through a conversational task, rather than an imitation task.

Participants will complete pre- and post-test tasks involving an AX-discrimination task, a lexical decision task, and a production task (Mok et al, 2013; Fung & Lee, 2019; Soo & Babel, 2020). The pre- and post-test tasks will be separated by a main task where pairs of participants will be asked to play an escape room video game (Pine Studio, 2021) together. In the game, they will be asked to work together to solve matching and password deciphering puzzles. The participants will be placed in separate rooms within the game, which will have slight variations. Participants are to discuss what they see and use similarities and differences in the rooms to work out puzzle solutions. Participants will complete these tasks either with or without the ability to see their interlocutor, as well as in a quiet or simulated noisy environment. The task itself is meant to elicit natural conversation-like interactions, while between-room differences are meant to elicit tone-merger-related utterances.

While the study is still ongoing, it is expected that phonetic adaptation will occur. It is expected that the degree of adaptation will be affected by factors such as baseline tonal distinction, age, tone pairs, conversation length, presence of visual cues, and amount of environment noise.
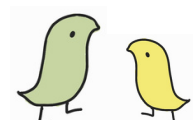
How to Cut Long and Rich-Format Text, Correctly, Once and for All

Presenter(s): Ziya Wang, Eunkyul Leah Jo, Grace Tianjiao Zhang & Angela Yoonseo Park (UBC)

(No Recording/Picture)

Both tokenization and sentence boundary detection (SBD) have been considered long-solved Natural Language Processing (NLP) challenges, especially in view of their seemingly consistent high performance and the prevailing pre-tokenized corpora. However, although often disregarded, the low error rates of current approaches are virtually task-specific, and the rule-based tokenization process is typically language-specific and often high maintenance in interoperability. Although subtle, these restrictions in comparability and reproducibility are nontrivial in the NLP pipeline. When such state-of-the-art tokenization and SBD implementations are re-applied to new use cases, most of the inaccuracies and errors will likely be ignored. Since most prevailing implementations are widely accepted as well-established and accessible for direct re-implementations, most of these ignored inaccuracies and errors will remain latent and opaque in the intermediate preprocessing results. Nevertheless, these text segmentation tasks are in such an initial and foundational stage of the NLP pipeline. Any latent inaccuracies or errors in such early steps could be unexpectedly amplified in downstream tasks, and significantly but unnoticeably impact the entire NLP pipeline.

In the literature, there are attempts to address and rectify these restrictions of tokenization and SBD, respectively, with comprehensive comparison and detailed reassessment (Dridan and Oepen, 2012; Read, Dridan, Oepen, and Solberg, 2012). However, although the results of their re evaluations are thorough and compelling, the claimed flexible and reproducible configurations are not provided for reimplementation. Furthermore, perusing these attempts side-by-side, it becomes conspicuous that these two tasks of segmenting text into words and sentences are practically more intervening and complementary of each other rather than two separate steps as conventionally accepted. In consideration of merging word and sentence segmentation tasks, some innovative sequence labelling method has been proposed and achieved compelling results in reproductivity crossing domains (Evang, Basile, Chrupała, and Bos, 2013). Unfortunately, some noticeable issues of alignment mismatches with sentence boundaries remained unsolved.

Therefore, in this presentation, we integrate tokenization and SBD tasks more seamlessly with a different alignment method and compile extensive evaluations of both preprocessing tasks into one systematic pipeline. For tokenization results, the assessment includes F1 measures and the Levenshtein Edit Distance; for SBD results, F1 measures are calculated on adjusted results (with realignments). To highlight this realignment method and demonstrate the core tasks of our approach, here is one adjusted example from our results.

Figure 1 shows the proposed evaluation by alignment. For tokenizations, the system result and the gold file are identical, so the F1 measure for tokenization is 1.0 (100%) for this short example text. For SBD, there is a discrepancy. With previous methods, Click here and To view it. would be considered as two separate and detected sentences in SBD, with two S labels for sentence boundaries under Click and To respectively (Figure 2, Evang et al. (2013)).

In such a case, the true positive would be 2, and the F1 measure is 0.8. However, as it can be easily perceived, this is one sentence messed by some rich format. The first sentence boundary marked by Click is, in effect, merely a partial match, while its second half is separated into another sentence by the second sentence boundary marked by To. With our proposed pipeline, we aim to avoid such partially matched sentences being true positives. By re-aligning sentences in the system results and the gold file, we catch this false positive case and reduce the true positive from 2 to 1, which reduces the F1 measure from 0.8 to 0.4 instead. That is to demonstrate, with the proposed pipeline, such latent but commonly-found inaccuracies and errors in the foundational prepossessing steps would be easier to capture and avoid.
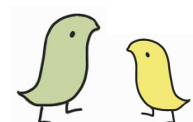
With this type of adjustment, our ultimate goal is to provide a reusable implementation for the evaluations of both tokenization and SBD tasks along with a corresponding one-stop demonstration of the results for these vital preprocessing steps to ease downstream tasks of the NLP pipeline.

Figure 1: Evaluation by alignment

Figure 2: Evaluation in previous work

What lexical variable (vacation) may tell us about the Canada-US border and the role of multilinguals in Metro Vancouver and beyond

Presenter(s): Sophia Abundo & Sarah Lefebvre (UBC)

This research paper analyses how the linguistic variation of the terms vacation, trip, holiday, and other more minor variants, are structured by social characteristics and patterns of regionality in a variationist framework. Most notably "homogenization on a continental level" (Dollinger 2015: 214) is a major prediction factor for Canadian and American variants of our lexical variable.

The question of interest is a closed answer question of the format: "Greece is my dream ___." The question form was optimised with Questionnaire Item Testing, with the intention of offering a maximally transparent question. In accordance with the WQ methodology in The Written Questionnaire in Social Dialectology (Dollinger 2015), we analyse demographic variables of age, BC Region and Regionality Index (developed from Chambers & Heisler 1999), English Use Index (EUI), gender, education, among others, from 1602 responses collected in-person and on-line in October and November 2022 with our ENGL 323 class. Blank entries for our variable and infelicitous responses were purged and demographic groupings were recategorized for effective analysis.

By cross-examining BC Region by age, ages 20-29 in the Metro Vancouver (MV) and Lower Mainland (LM) regions expressed 65.17% and 52.12% respectively in favour of the variable vacation. The MV and LM 20-29 age group's increase in vacation may demonstrate the phenomenon of "Americanization" and "homogenization" of linguistic variants due to the geographic proximity of MV to the American-Canadian border as British terms once "systematically distinguished" Canadian English (Dollinger 2015: 214). Further, apparent-time hypothesis suggests that speakers from different generations mirror the language they learned when they were young adults (Dollinger 2015: 176) and allows backdating. By 1973 the preference for vacation was noticeable, suggesting that the British variant, holiday, which then was also dominant in Canada, had begun to lose relevance in the younger generations. This was found to align with Dollinger's findings that British variants had "considerable currency until the third quarter of the 20th century" (Dollinger 2015: 216). Additional analysis revealed that gender did not significantly influence the selection of linguistic variables, with all gender categories displaying a similar distribution of responses. However, a cross-examination of EUI –how frequently the English language was used – and education revealed that the use of vacation among respondents of a low EUI category (i.e. less multilingual speakers, EUI = 0 means a monolingual English speaker) significantly increases with greater education suggesting no overt prestige for either vacation or holiday.
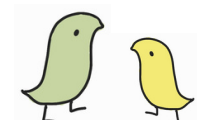
Stories and Screens: Mother and child behaviour during shared reading of electronic books with and without narration
Presenter(s): Sofia Cordeiro (UBC)

Keywords: digital; literacy; scaffolding; vocabulary; development

Educational outcomes are associated with children's early literacy environment (e.g. Gottfried et al., 2015). Research suggests that specific styles of interaction facilitated by caregivers during shared reading, such as vocabulary/narrative scaffolding and joint attention, are particularly important features of this environment (Wasik et al., 2016). In recent years, there has been an increase of digital books among families and schools, where children often experience books narrated through applications on tablets or other online platforms. While it has been shown that children learn effectively from multimedia e-books when read with an adult who provides high-level scaffolding (Savva et al. 2022), it is unclear what happens to parent-child behavior in the context of lower-resource digital platforms.

In this study, we ask whether a minimalist e-book with recorded narration elicits different parent-child interactions from a book that the parent must read aloud to the child. Parent-toddler dyads read two counterbalanced digital stories, one with recorded narration and one without. Participants are video-recorded through the entire session. Each dyad's interactions are then examined for scaffolding and joint attention behaviours. These are operationalized as follows: scaffolding includes parent responses that (1) connect, (2) explain, (3) prompt, and (4) respond to the child's question/comment. Joint attention includes an initiating and response interaction (e.g., the child points and parent responds to that initiating action). We are looking for differences in number, type, and quality of scaffolding/joint attention behaviours between the two digital book experiences. Data collection is ongoing.

We anticipate two possible outcomes. First, we may see more joint attention and scaffolding between parent and child in the unnarrated book. This would suggest that (some) digital books are less likely to invoke educationally rich caregiver behaviours during shared reading. On the other hand, we may find increased amounts of scaffolding and/or joint attention episodes during the narrated as opposed to unnarrated book. This would suggest that even minimalist e-books provide educationally rich literacy support for children. This research will thus contribute to both theoretical and practical goals: we hope to delineate the optimal features of shared reading, and how to implement/promote those features, in the context of the digital literacy landscape.
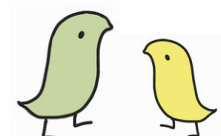
Social perception of [n] and [l] initials in Cantonese
Presenter(s): Sandy Chow (UBC)

Sociolinguistic research has shown that particular pronunciation variants in sound changes bear social weight (King et al., 2022). Listeners may have perceptions about a talker's age, socioeconomic status, or educational level based on their pronunciation (Lawrence, 2017; Pan, 1981). For example, the historical syllable-initial /n/ is unconditionally merging with /l/ in Cantonese, meaning Cantonese speakers tend to pronounce /n/ initial words with a [l] sound (Cheng et al., 2022; Zee & 徐, 1999). Speakers have metalinguistic awareness of this sound change, associating historical [n] pronunciations with prestige and "properness" while innovative [l] pronunciations are socially stigmatized. Pan (1981) found that speakers who used more historical [n] pronunciations were judged more positively on social dimensions (e.g. place of residence, educational attainment) than those who used [l] pronunciations. However, we do not know if the effect continues today. Therefore, through a conceptual replication of Pan (1981), this study examines whether social perceptions about [n] and [l] pronunciations are still present today.

A Cantonese story with 307 Chinese characters was created, of which 37 were critical historical /n/ words. Two versions of the base story were then generated. In the L-version of the story, 20% of the words were produced with [n] and 80% of the words were produced with [l] (7 [n] and 30 [l] pronunciations). In the N-version of the story, these percentages were flipped; 80% [n], 20 % [l] (30 [n] and 7 [l] pronunciations). Participants were randomly assigned to listen to one of the two story versions and then completed a questionnaire to assess whether the pronunciation of historically /n/ words affects social perception. The questionnaire included 5-point Likert scale questions (1 lowest --- 5 highest score), asking participants to rate the speaker's characteristics (e.g. correct mistakes, follow rules) and some multiple choice demographic questions about socioeconomic status and educational level. If the social biases observed by Pan (1981) still exist today, the [n] and [l] pronunciation variation should affect the scores of the rating scale. Specifically, participants who heard the N-version of the story would provide higher ratings in each criterion than those who heard the L-version of the story.

 The preliminary rating data collected from 26 Cantonese listeners (10 L-version and 16 N-version listeners) is provided in Figure 1. Listeners who heard the L-version of the story rated the speaker as more likely to double check one's work and value traditional cultural values, while ratings were well matched for all other speaker characteristics across story versions. These results do not match the original predictions. However, 7 out of 10 L-version story listeners think the speaker is from Hong Kong, while only 5 out of 16 N-version story listeners think so. Participants consider the [l] pronunciation to be more familiar and symbolic of Hong Kong Cantonese. The results of this study will shed light on how participation in a sound change affects social perception in Hong Kong Cantonese.
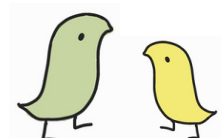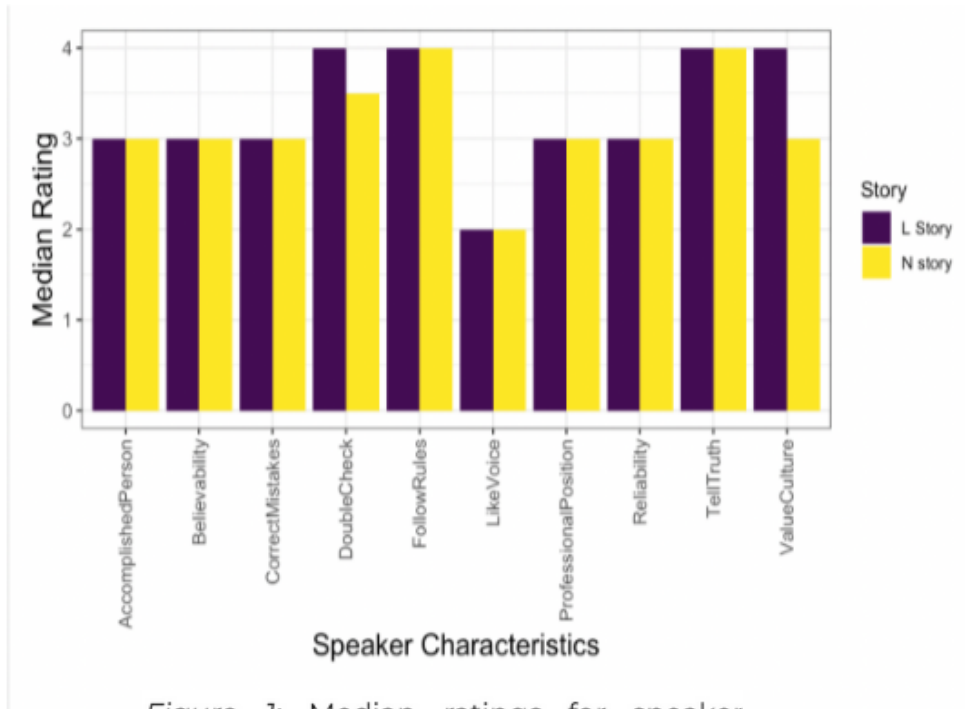
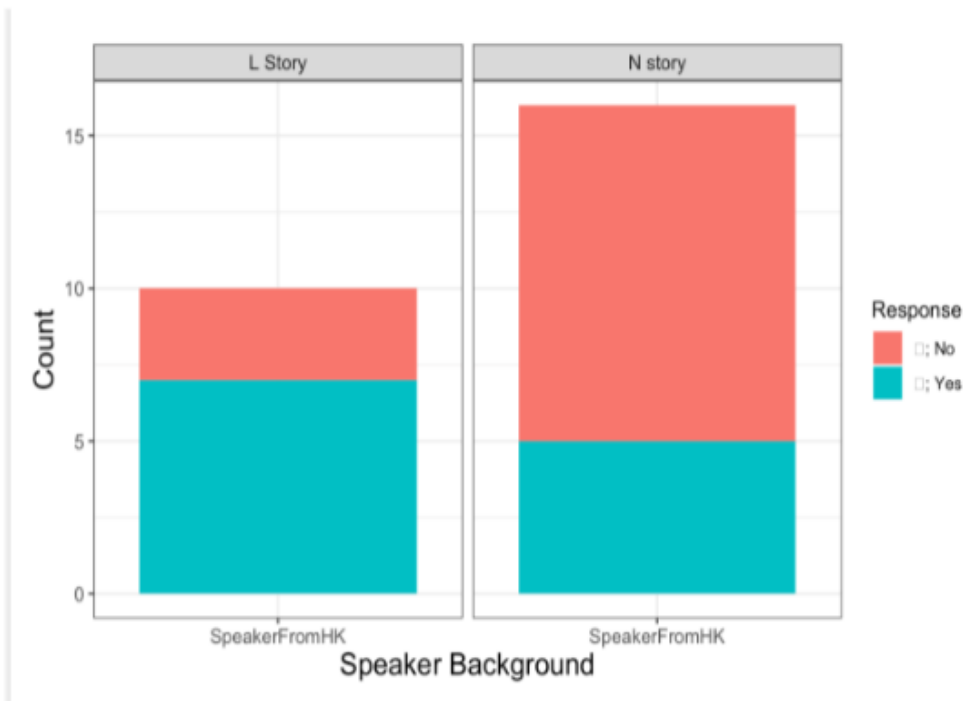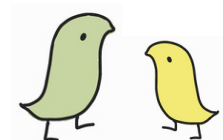Figure 1: Median ratings for speaker characteristics by 26 Cantonese listeners.

Figure 2: Count of listeners who think the speaker is from Hong Kong.

Does Delivery Matter?
## Presenter(s): Charlize Ma, Effie Kao, Stephanie Wang & Raechel Kitamura (UBC)

When conversing with another individual, we can observe change, or accommodation, in many aspects of our speech. Convergence in speech occurs when two speakers come to an agreement after the discussion, whereas divergence in speech indicates the two speakers do not share a general consensus after the conversation. This accommodation has been observed in physical lip and brow movements(Ménard, L., Leclerc, A., & Tiede, M., 2014). There has also been acoustic evidence of convergence and divergence when measuring factors such as pitch (F0). We may converge (or diverge) due to a multitude of factors. Most notably factors like age, social status, stereotypes about outgroup members, and current historical context (Giles, H & Ogay; T., 2007), all of which may have an impact on our speech and acoustic patterns (Babel, M. & Bulatov, D., 2011).

Our experiment investigates whether physical and acoustic changes occur during converging or diverging in opinion with another speaker. We predict that the acoustic and physical movements of two individuals will also converge when the conversation converges.

In order to investigate our hypothesis, 20 pairs of students between the ages 18-24 from the University of British Columbia were planned to recruit via UBC Linguistics SONA to participate for course credit. The participants are asked to have a 3-5 minute conversation on ZOOM. Video and audio recordings were taken of their speech throughout the conversation. Additionally, a 30-second baseline speech sample was taken prior to their conversation to sample their speech when unaffected by opinion. Facial movements such as brow and lip movements were measured with the facial landmark detection toolkit, OpenFace. Acoustic aspects were also accounted for to determine whether there were signs of acoustic convergence (or divergence), such as F0, using Praat. Pre-test and post-test questionnaires were distributed to collect participants' demographics and their opinion changes after the discussion with another participant.

In analysis, a cross-comparison between the first and last minute of speech is conducted to identify any evidence of accommodation, specifically looking at the facial movements and F0 of the participants.

The two pilot studies that we have conducted show preliminary results. The acoustics data displays that when the conversation between two participants is converging in speech, their F0 will have a smaller difference across participants at the last minute of the conversation (on average, 6.83Hz for two same gender participants) compared to the first minute (on average, 9.64Hz for two same gender participants) and the baseline speech (on average, 14.48Hz for same gender participants). The intensity of action units that ranges from 0.0 to 1.0 from OpenFace also showed similar results. When the conversation between two participants converges, the difference in lip and brow movements between the two participants decreases at the last minute of the conversation (on average a 0.72 intensity difference of action units) , compared to the first minute (on average a 0.56 intensity difference of action units). This suggests that their F0 and intensity of facial movements converge when the conversation is converging.
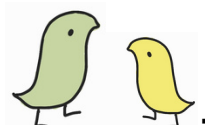
Benchmarking the corpus annotation scheme for French NER
Presenter(s): Grace Yang (UBC)

Keywords: computational linguistics, named entity recognition (NER), annotation representation, conditional random fields, French

A named entity (NE) is an information unit that can be denoted with a proper name, such as a person, organisation, location, etc. In information retrieval, named entity recognition (NER) is used to identify references to the entities in text (Tjong Kim Sang and De Meulder, 2003). The three most common annotation representations are IO, BIO, and BIOES. The simplest representation, IO, labels named entities with I and words outside the entity with O. BIO, adopted by the Conference on Computational Natural Language Learning, assigns a tag to the beginning (B), inside (I), and outside (O) of a named-entity. The most complex representation, BIOES, marks the end (E) of a named entity and single-token entities (S), in addition to the tags used by BIO. Figure 1 shows an example of different annotation schemes by IO, BIO and BIOES.1

TOKEN	IO	BIO	BIOES
The	O	O	O
NBA	I-ORG	B-ORG	S-ORG
player	O	O	O
Michael	I-PER	B-PER	B-PER
Jordan	I-PER	I-PER	E-PER
is	O	O	O
from	O	O	O
the	O	O	O
United	I-LOC	B-LOC	B-LOC
States	I-LOC	I-LOC	I-LOC
of	I-LOC	I-LOC	I-LOC
America	I-LOC	I-LOC	E-LOC
.	O	O	O

Figure 1: Example of IO, BIO and BIOES

The objective of this study is to explore the effect of different annotation representation schemes on NER tasks in French. The choice of annotation representation for NER tasks is not always straightforward. It has been suggested that the optimal representation scheme depends on the language given, as grammatical structure may affect the way named entities appear.

The models for this study were experimented using conditional random fields CRFs (Lafferty, McCallum, and Pereira, 2001). We use Wapiti's CRF implementation Wapiti (Lavergne, Cappé, and Yvon, 2010)2 and results are evaluated with the standard F1 metric using conlleval3. Models were trained on French NER data provided by Europeana Libraries4. This data is based on OCRed and manually annotated historical newspapers from the National Library of France and are originally in IO format. For training and predicting, we converted IO annotations to BIO / BIOES and added sentence boundaries. We then split the corpus in a ratio of 90/10 as training/test data sets. For a fair comparison, we converted all predicted data back to IO before evaluation. Our results are shown below. The values in brackets represent the total number of entities.

RESULT	IO	BIO	BIOES
PER (200)	57.35	58.52	58.40
LOC (634)	69.04	67.80	69.09
ORG (107)	28.57	22.82	26.67
Overall	62.41	61.33	62.57

Overall, the BIOES annotation scheme performed as expected, producing the best results. This paper shows the effects of different annotation representations for NER and can improve NER performance for French by evaluating the different annotation representation methods. Future work could include error analysis of whether the poor performance of ORG type is due to lower frequency in the dataset, in addition to why IO outperforms BIO when the IO scheme usually performs worse. Moreover, using better classification algorithms, such as Transformer, along with a wider range of datasets may allow for further insight into the impact of annotation representations on French NER tasks.

Variability on vowel production for three-year-old Spanish-English bilingual
Presenter(s): Elizabeth Alvarez-Alfaro (SFU)

Behavioral science studies demonstrate that infants can discriminate general sound categories before they reach six months of age. After this period, they become language specific learners, tuning their receptive abilities to the language they are exposed to (Munro, 2021). In the case of bilinguals, infants are able to tune into the two distinct codes they are immersed into. Since the specific reception skills have a direct effect over language production. What happens with the vowel production of Spanish-English bilingual learners by age three? Earlier studies have noted that there are differences in the realization of vowel categories for bilingual children.

The aim of this study was to assess the production of Spanish and English vowel phones produced by a three-year-old growing up in a bilingual household environment. The experiment was done with a flash-card naming task. The subject of the study was tested on her realization of words containing three different Spanish and English vowels categories. The phonetic productions from each target sound realization for each linguistic code were compared using an acoustic description of F1 and F2 values.

The results show that the test subject has different vowel categories for English and Spanish languages, realizing their vowels in different places, hence producing different categories of the target vowels for each of the spoken codes.
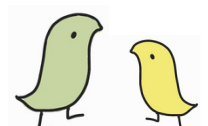
Metalinguistic Awareness in Young Children and Toddlers
Presenter(s): Julianne Bittante (SFU)

Keywords: metalinguistic awareness, children, online interview, bilingual, English

Prior research has demonstrated that older bilingual infants have an implicit awareness that informs their choice of language use with their caregivers and family members (Montanari, 2009; Nicoladis, 1998; Petitto et al., 2001) and that this awareness is consistent as these infants develop (Afshordi et al., 2018; Comeau et al., 2007; Genesee et al., 1996). This awareness can be detected in several ways, including by observing children using the correct language with their caregivers, a process defined as pragmatic differentiation (Nicoladis, 1998). However, studies have found that multilingually exposed 3- to 5-year-olds cannot consistently label the languages they speak or hear from their caregivers', demonstrating low metalinguistic awareness of the languages they speak, as only about half of these children respond with accuracy (Ahktar et al., 2012; Atagi & Sandhofer, 2020). Using a larger age range, our research examines at what age children can accurately label languages, as well as when children can accurately answer questions about their caregiver's frequency of language use, which as far as we are aware, has yet to be studied.

We examined the interview responses from a structured, online format, of 44 bilingual children. The children's age range was 2.5-6 years and they resided in an English-majority area of Canada or the USA, with their reported English language exposure being anywhere from 10- 85%. The study was conducted in English and consisted of a cartoon that described to child participants the concepts of knowing and hearing languages, before asking them what languages they know and how often their family members speak to them in each language. Children indicated their answer using provided visual supports of language exposure and their answers were recorded on a 5-point scale. Each child's parent completed a questionnaire, indicating on a
7-point scale, how often their child was exposed to each language and which language family members use with the child and how often.

During the interview, children were first tested on their accuracy in naming the languages they knew, and we found that children's answers became increasingly more accurate with age (Figure 1). Then we analyzed the accuracy of children's responses to being asked about their family member's language use and found that the answers from the youngest children were above chance and stayed constant through development (Figure 2). We found that, consistent with past research, bilingual children between the ages of 3 and 4 are unable to accurately label the languages they know (Akhtar et al., 2012). However, we also discovered that children begin to acquire this ability around age 5-6 years. Additionally we found children's understanding of which language each caregiver uses to be moderate but developmentally stable between the ages of 2.5-6 years. The results of this study help us to understand metalinguistic language development and language experience of young bilingual children, but more research is necessary to determine why there is so much variance in children's metalinguistic awareness.
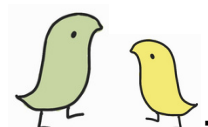
Phonetic variation and perceptual processing of code-switching
### *Presenter(s): Khushi Nilesh Patil (UBC)*

Keywords: code-switch, switch cost, accents, bilingualism, phonetic cues

Code-switching is the act of switching from one language to another mid-utterance, a common behaviour in bilingual or multilingual speakers. Research suggests there is an associated "switch cost", or negative impact on processing, recognition, and comprehension when code-switching occurs (e.g., Thomas & Allport, 2000; Costa & Santesteban, 2004; Soares and Grosjean, 1984; Olson, 2017). At the same time, recent work suggests that subtle phonetic cues prior to the code-switch may allow listeners to anticipate the coming code-switch. Shen, Gahl and Johnson (2020) demonstrated that anticipatory phonetic cues mitigate switch costs in Mandarin-English bilinguals. In this study, we build on this work and ask whether a speaker's accent, not solely their anticipatory switch cues, affect the comprehension of a code switch. Given that the processing cost in code switched speech may be due to expectation violations and not a switch between lexicons (Yacovone et al., 2021), we test whether English to Mandarin code switches are processed faster in Mandarin-accented English compared to locally-accented English.

The methods closely follow the concept monitoring task (Exp. 1) in Shen et al. (2020), but with the inclusion of a second voice. Two female first language speakers of Mandarin of similar ages were recorded for stimuli. One of these speakers also had English as a first language (local English speaker), while the other learned English as a second language (Mandarin-accented English speaker). Speakers were recorded producing items that either did or did not contain a code switch from English to Mandarin — e.g., We got a qì qiú (balloon) for her birthday or We got a balloon for her birthday. These stimuli are used in a concept monitoring task where listeners are asked to press a button when they hear the word corresponding to the image on the screen (e.g., a balloon). The task also contains catch trials where the image is not present in the sentence. We analyze data from any participant who was exposed to Mandarin before the age of 5. Data collection is ongoing.

Response time data to the concept monitoring will be analyzed with Speaker (local accent, Mandarin-accented English) and Switch (unilingual, code-switch) as factors. We predict that participants will identify the heard concepts faster in the code-switched condition when the speaker produces Mandarin-accented English, since accented speech globally increases listeners' activation of Mandarin, thereby mitigating switch cost. In the code-switched manipulations we expect to see faster reaction times in the Mandarin-accented sentences as compared to the locally accented condition. Because of the intuitive nature of a unilingual sentence, we also expect that participants will react fastest to the unilingual, locally accented stimuli. However, response time to locally accented vs. Mandarin-accented stimuli may differ based on participant language backgrounds. Results will be analyzed to see if any such patterns emerge.

This research contributes to the existing knowledge on phonetic cues and processing costs in code-switching. In particular, we look at accented speech, which is a little studied factor in this area. This work contributes to our knowledge of bilinguals' processing of languages.
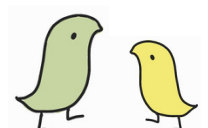
Language Learning in a Global Babywearing Community of Practice
## Presenter(s): Katherine Brand (UBC-O)

Keywords: Babywearing, Community of Practice, Communication, Online Communities, and Registers

With the support from our community partners at the Center for Babywearing Studies (CBWS), myself (undergraduate student Brand) and supervisor, Dr. Christine Schreyer, are contributing to scholarly research on the language of babywearing. !Babywearing" is a global term that embraces different culturally appropriate methods one can use to carry a child on their body. A range of materials including a cloth, sling, or soft structured carrier are used to babywear (Rus sell 2014). Our research began with a preliminary study in 2021 when Brand collected and analyzed data generated from a search of community hashtags including: #WearAllTheBabies, #BlackBabywearingWeek, and #Babywearing (Brand 2022). Our current project is community based research that draws on and contributes to the fields of linguistic anthropology, linguistics, gender and women's studies, and identity theory.

Our research is focused on in-group language of the babywearing community of practice. We examine how members of this community self-identify, how they learn terminology, and how they talk about babywearing. The concept of community of practice was developed by Jean Lave and Etienne Wenger in their book Situated Learning (Lave and Wenger 1991). These are sites of community built around shared practices, values, and worldview. A community of practice has shared registers for in-group language and in the case of babywearing, language is central to group knowledge and membership in online and off-line spaces.

We have just completed the first phase of this project, our transnational survey. The survey was created with the program Qualtrics and included both demographic questions to determine the identities of our participants, as well as open-ended questions that allowed participants to share the babywearing language important to them. Participants were required to read and write in English but there were opportunities to share babywearing knowledge and terminology in other languages throughout the survey. The survey ran from April 3rd to May 15th, 2022 and 396 individuals participated from a range of countries, including: the United States, Canada, Hungary, United Kingdom, and Italy.

From our analysis we learned that the most common primary languages of participants were English, Hungarian and Italian. Most participants, 97.7 %, identified as women. The survey data illustrated that the many participants learned about babywearing practices through baby wearing groups. However, this category could include online or offline groups as the babywear ing community is global, local, and virtual. Additionally, the data showed that babywearers most commonly learned terminology via social media and babywearing educators/education courses. Lastly, the data showed that the community is a rich source of in-group language as there was 1566 words for babywearing carriers provided by respondents, as well as 1439 words for carrying techniques, 712 slang terms, and 540 acronyms and abbreviations. It is evident from our research that this community of practice has a rich, unique, lexicon and in this presentation, we discuss more details of this, how it is learned, as well as why this knowledge matters to our community partner– the Center for Babywearing Studies.
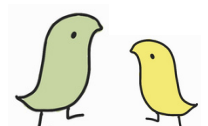
Part of Speech Tagging for Levantine Arabic
## *Presenter(s): Saughmon Boujkian (UBC)*

Keywords: CRF, POS, zero-morpheme, Arabic, dialects

Due to the poor performance of POS taggers in Levantine Arabic (LA), this paper provides a way of assigning part-of-speech (POS) tags to sentences in Levantine Arabic. To train this model, we are using Modern Standard Arabic (MSA) sentences. POS refers to categorizing words in the corpus such as nouns, and verbs. Modern Standard Arabic is the language used in formal written form, such as in news, academic papers, and newspapers whereas the language used in informal speech and informal texts, like social media posts, is Dialectal Arabic (DA). In this case, the targeting dialect is Levantine Arabic (Monirabbassi, 2008). LA and MSA sentence formation is quite different, making their POS tags different. To overcome this, we followed linguistic methods to make the POS tag more similar. Then, we trained a model using 500 sentences from the Penn Arabic Treebank's MSA data. Additionally, we extracted 100 sentences in Levantine Arabic from Whatsapp messages and Facebook posts. The words in each sentence were segmented into morphemes and finally, POS tags were added creating a Levantine Arabic dataset for the evaluation purpose. Using the Conditional Random Field algorithm (Lavergne, Capp´e, and Yvon, 2010) with 1-bfgs approximation and a learning rate of 0.01, the MSA data were used for training, then the trained model is evaluated on Levantine Arabic data. In order to further understand the way the LA dataset was made, I present an example below;

(1) raw sentence: هي هون بالغرفة ما بعرف شو عم تعمل
gloss: she here in-the-room don't know-I what is do-she

(2) morpheme by morpheme: هي هون ب الغرفة ما بعرف شو عم تعمل
gloss: she here in the-room don't know-I what is do-she
translation: 'she is here in the room, I don't know what she is doing'

(3) POS tags: PRON ADV IN NOUN RP VERB WP VERB

Firstly, The dataset was divided into sentences. Then, the sentences were split into morphemes. After that, Penn Arabic Treebank POS tags were added to each sentence. Finally, the POS tags of Levantine Arabic and Modern Standard Arabic were compared, showing differences in the sentences that have the exact same meaning. Therefore, we used linguistic concepts like zero morphemes and linguistic environment analysis to make the tags in both dialects more similar. For example, LA drops some articles before the main verbs in the sentence. Therefore, we used zero morphemes to add missing articles in LA to make the POS tags more similar to MSA that does not drop the articles. After making the POS tags more similar, the model was trained and tested on Levantine Arabic. The results and evaluation will be discussed during the conference since they are not ready yet.

In conclusion, this research uses a dialect, MSA, that has much more data available in order to study a low-resource dialect, LA (Darwish, Mubarak, Abdelali, Eldesouki, Samih, Alharbi, Attia, Magdy, and Kallmeyer, 2018). Nevertheless, this causes a reduction in the accuracy of the POS tags in LA. Improving the accuracy of POS tags in LA using more algorithmic-driven techniques is the next step beyond this research.
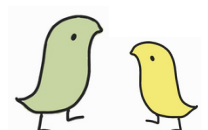
My Pronunciation is Éxcusable: A Diachronic Look at Stress Placement with English Derivational Suffixes

Presenter(s): Charys Russell (UCalgary)

Modern English stress in suffixed words, while standardized, is not static. Stress placement in English words has changed over time; for example, the words excusable and prosperity had initial stress in the 16th century, while today stress is placed on the second syllable, and the word demonstrate had stress on the second syllable while today it has initial stress. Lass (1999) attributes stress placement to a word's etymology; words of Germanic origin have stress on the first syllable of the base, while words of Romance origin follow rules based on syllable weight. This does not hold true of Modern English; therefore, one asks if English has systematically changed how it deals with stress in suffixed words.

I collected dictionary entries from the era of standardization of English (Levins 1570; Dyche 1725; Johnson 1755, The Oxford English Dictionary 2021); I then compared the described stress placements and observed any changes. I had predicted that English stress placement has been systematically changing over time. I found instead that while most suffixed words have retained their stress placements, there have been two waves of stress shifts in English history; the first saw stress moving leftwards (demónstrate > démonstrate) and the second rightwards (prósperity > prospérity). These shifts only occurred in some words and were not systemic. I propose that the first change in stress placement was due to natural language change, in which foreign words assimilated to the native stress pattern, while the second was due to the social pressures of language standardization.
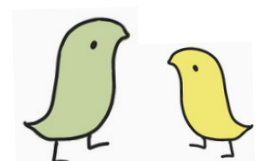
Indigenous Language Identification
Presenter(s): Yee Siong Pang (UBC)

With the advancement of information technology, machine translation such as Google Translate and Amazon Web Service is very much accessible to users of almost all of the most widely-spoken languages in the world, as well as many others. Through this kind of software, both written and spoken languages can be identified, understood, and translated to other languages for further processing. However, for under-resourced languages such as the various First Nations languages spoken in what we now call British Columbia, these services are currently not available – neither translation nor even identification. If given access to identification services (and later machine translation into other more widely-spoken languages such as English), these languages' revitalisation efforts could potentially be enhanced by allowing for quick translation at various morphosyntactic levels not necessarily found in dictionaries. For this reason, we propose to develop a text-based Indigenous language identification system for languages spoken in BC, geared towards firstly the eventual development of machine translation software, and secondly expansion beyond the colonial borders of BC. Further development with more advanced technology could also focus more heavily on spoken language, to account for many Indigenous languages traditionally being purely oral and not communicated through writing.

We describe here a project which has been implemented in a Personal Computer using freely available Machine Learning Tools and Natural Language Processing Algorithms and the program can identify successfully the following First Nations languages which have been approved for the UBC admission requirements:

 We describe here a project which has been implemented and can identify successfully the fol lowing First Nations languages which have been approved for the UBC admission requirements:

Halq'eméylem, Heiltsuk, Hul'q'umi'num, Kwak'wala, Liqwala/Kwakwala, nsíylxcn, Nte?kepmxcin, Nuu-chah-nulth, Secwepemctsin, Shashishalhem, Sim'algaxhl Nisga'a, Sm'algyax, Tsek'ene, and Upper St'at'imcets.

For proof of concept, a modest dataset of about forty sentences/phrases for each language was collected. Due to the use of different orthographies, identifying the above languages can easily be implemented. We anticipate we are currently not able to predict sister dialects from a single language. With a more comprehensive dataset, the accuracy of prediction will improve. We believe the expansion of the dataset is an on-going process that could be in the form of requesting users for inputs. As mentioned above, various types of orthographies are being used for Indigenous languages.

Linguists have introduced the Americanist variant of the International Phonetic Alphabet (IPA) and some communities have retained its use. The alternatives are recently introduced Practical Orthography which uses alphabets based on an English type-writer. The other method is syllabaries which have been used for a long time.
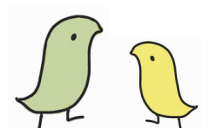
1. Americanist form:

   uł kʷul'ɬxʷus uł ikĺíʔ ʔácqaʔ ʔitx. ˌNsilxín

2. Practical orthography:

   Letsˈáxw kwˈeselh olu plhát te máqa. ˌHalqˈeméylem

3. Blackfoot (a First Nation) language: ᓄᐧᔭ ᗫ ᗫᐠᒍᑊᕁᖎᗫ

Whereas we do not discuss here the advantages and disadvantages of each type of orthography, this gives us a layer of complexity for the purpose of web-crawling and corpus building. It requires us to consider the problems of encoding and decoding of Indigenous languages, for reading and processing the dataset beyond the identification level. Currently we are using "which form" and we leave other encoding problems as future work.
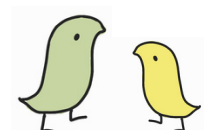
Measuring Brain Response When Exposed to Infant-Directed Speech: A Meta-Analysis
Presenter(s): Jaimie Muller (UBC)

Keywords: Infancy; infant-directed speech; fNIRS; brain analysis; meta-analysis

Adults speak to babies at a slower rate, with a higher pitch and exaggerated word contours (Naoi et al., 2012) – specifically, adults use infant-directed speech. Existing studies have found that the inferior frontal (Altvater-Mackensen & Grossmann, 2016) and superior temporal (Bortfeld et al., 2007; Cristìa et al., 2014) regions, commonly associated with language perception in adults, are activated in infants exposed to infant-directed speech. However, these studies' findings are inconsistent, with either both hemispheres (Fava et al., 2014), or primarily the left hemisphere (Altvater-Mackensen & Grossman, 2016; Bortfeld et al., 2007; Cristìa et al., 2014) activating. This creates controversy regarding the spatial organization of early language systems.

This pre-registered meta-analysis of twenty studies, with over 600 three- to twelve-month-old infants measured, investigates whether the inferior frontal and superior temporal lobes of both hemispheres activate when infants are exposed to infant-directed speech.

Functional near-infrared spectroscopy (fNIRS) is a quiet, non-invasive neuroimaging technique that has been widely and increasingly used in the past 20 years, allowing researchers to measure the brain with high spatial resolution in awake and mobile infants. fNIRS measures the change in relative blood oxygenation in the brain when participants are exposed to stimuli. As such, researchers can see which brain regions are activated when babies hear infant-directed speech. This meta-analysis will be the first to synthesize findings of infant-directed speech brain identification from studies involving fNIRS.

We collect studies by searching keywords through databases (UBC Library, PubMed, and Web of Science). Inclusion criteria involve healthy monolingual or bilingual babies, between 3-12 months, presented with any form of infant-directed speech stimuli (e.g., syllables, prosody). Once abstracts have been screened, we will complete full-text reviews on included papers to summarize the brain regions significantly activated in their results. We will code for a) whether our hypothesized brain regions were activated, b) whether audiovisual or auditory stimulus was presented, and c) the type of speech presented: native/non-native language, prosody, continuous speech, vowels, syllables, and syntax patterns. We follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol's (PRISMA-P) guidelines, which extensively review our inclusion criteria to ensure reliable findings (Quintana, 2015; Selçuk, 2019). We will use the PRISMA-P's flowchart to summarize included and excluded papers.

While currently collecting data, we hope to clarify and summarize the spatial brain organization for the network that processes infant-directed speech during infancy, as well as to inform the scientific community of the ongoing debate regarding the development of language lateralization (ie., one hemisphere processing stimuli more than the other).

Our field has now produced enough fNIRS studies to complete meta-analyses. A more comprehensive view of the identification of brain regions involved in processing infant-directed speech will further our understanding of how infants perceive human speech. Furthermore, researchers can build on these findings and study brain responses to speech in more naturalistic and generalizable settings.

Environmental Entropy as a Means to Further Explore Child Bilingualism.
Presenter(s): Farzana Ali (SFU)

Research on bilingualism is often reduced to a simplified categorization of monolinguals and bilinguals, which erases the linguistic variation that both monolingual (Bice et al. 2019) and bilingual populations experience (Gullifer & Titone, 2020). To better capture the entirety of bilingual experiences, some researchers have begun to consider more than just amount of language exposure and use. For example, Gullifer and Titone (2020) recently introduced a metric called language entropy, a calculation of diversity in language use (i.e., whether one or multiple languages is spoken) across contexts.

We apply this concept of entropy to child language development. Here, we focus on entropy in children's exposure to language (what we call environmental entropy) from individual speakers. That is, we asked how much one language was habitually or consistently used by different individuals in a child's life, which is distinct from code-switching, or mixing languages within a speech utterance (Byers-Heinlein, 2013). While prior work has suggested that a number of child language questionnaires ask about diversity in language exposure (Kašćelan et al., 2022), few studies, to our knowledge, have explored this concept of environmental entropy at the individual level, or determined how it may be distinct from other factors (like age, or the overall balance of language exposure).

In this study, we collected data through a parent questionnaire from 56 North American parents living in an English-majority context who had bilingual children between the ages of 2;3-6;6 years. Likert scale measurements of environmental entropy were provided by each child's parent for caretakers, additional family members, and non-family interlocutors (e.g., teachers) on a scale of 1 (habitually uses only English) to 7 (habitually uses only non-English).

We weighted the self-reported measures of language input by the amount of time each person spent talking to the child, and so, unpredictable language input resulted in high entropy scores, and predictable language input resulted in low entropy scores. Additionally, parents reported the child's overall exposure to English, as well as the age and gender of their child.

Results showed that environmental entropy did not significantly correlate with overall language exposure (Figure 1) or age (Figure 2). This suggests that environmental entropy is a construct that is distinct from how much English one hears, that does not change with age. These findings contribute to the literature in describing the variability in who provides English input to the child in immigrant contexts from English-speaking North America (Lee et al. 2015). For example, in some cases excessive amounts of English input may be from one person, whereas, others may receive the same amount of English input from multiple people who often code-switch (Kremin et al. 2022). Through measuring environmental entropy our study reveals the growing consensus that bilingual children have diverse language environments, and therefore, "bilingualism" in research should be treated in a continuous manner. By doing so, further investigation of bilingualism using entropy may reveal effects on children's linguistic and cognitive development.

POSTER PRESENTATIONS

The Social and Regional Distribution of Variable "Rock-paper-scissors": Monolingualism at Play
Presenter(s): Amy Ng & Lucienne Chang (UBC)

Digital without Q&A

Keywords: variation, multilingualism, written questionnaire, Canadian English, diversity

This report analyses data sets from our Written Questionnaire (WQ) on English usage to explore linguistic variants with a specific focus on Canada. The linguistic survey, designed in the WQ style in accordance with methodology guidelines from the textbook (Dollinger 2015), was distributed by the entirety of our ENGL323 class from October 26th to November 17th 2022 to a variety of initial research participants, including a mutual network of friends and family, social media sites, and any other opportunities that may have arisen accordingly. A TinyURL link and a QR code were generated for easier access and accessibility purposes to the complete survey; when given to our primary data sources and prompting them to nominate other potential data sources that would be able to participate in the survey, resulted in the combined sampling method of convenience sampling alongside snowball sampling. Overall, the survey gathered 1600 responses from a variety of different answers, including Canadian and international responses as seen in Table 1.

Variant	N
Rock - paper - scissors	1216
Ching - chang - push	7
Other	44
Rock - paper - scissors shoot	32
Ro - cham - beau	10
Paper - scissors - rock	11
Scissors - paper - stone	6
Scissors - paper - rock	3
Sum	1329

Table 1: All responses regardless of the status of residence

In focusing on a specific age range, data from a grouping of younger participants, in this case, confined to a range of 15-30 year-olds: split into focus groups of 15-19, 20-24, and 25-29. This range presented a pattern of high percentages for the popularity of rock-paper-scissors, with all results having percentages in the 90s: 93.8%, 91.9%, and 90.8% respectively. Interestingly, the second largest variant was other followed by rock-paper-scissors shoot. The second highest variable, other, may come as a result of participants inputting their own uniquely specific ways to refer to this style of game; this disparity will be discussed further in the paper with regard to EUI calculations and the possibility of this collected data showing a connection between other results and multilingual participants.
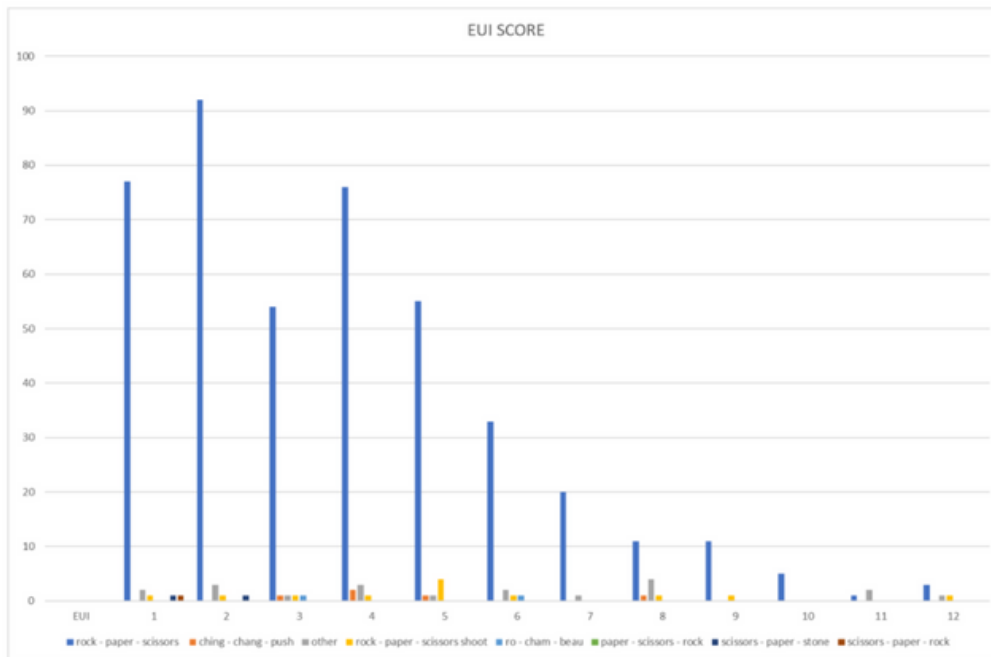
Figure 1: English usage index for Canada (regardless of region) Note: Did not include data >1.

Finally, we found that the findings of this study conclude that monolinguals in Canada overwhelmingly favour rock-paper-scissors. This may be due to the domination of the rock-paper-scissors variable both locally and internationally, as popular culture and media from the USA and other similar commonwealth countries also favour the variant of rock-paper-scissors. While the majority of answers by participants remain the long-standing, dominant answer of rock-paper-scissors, as confirmed by our monolingual speaker results, the existence of our second largest variable, other, suggests that various multilingual speakers of Canadian English both embrace the English language but retain aspects of their other languages rather than adhering strictly to a standard of "proper" English. The rising popularity of other and the inclusion of answers in different languages may suggest a lessening of the pressure to conform to English as the majority Canadian language for multilingual speakers, as multiculturalism and diversity movements become increasingly popularised in the Canadian social intelligence.
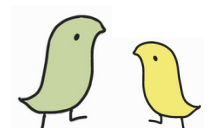
The Role of Affective Prosody in the Perception of Emotional States (Digital)

Presenter(s): Madeline MacLean (UCalgary)

Digital with Q&A

As humans, the ability to infer the emotional states of others is essential to our daily lives [1] and to do so, we process and integrate multiple cues that are available: previous and/or concurrent knowledge about our conversational partner, knowledge about the world, and various linguistic cues such as prosody [2]. The present study investigates two dominant sources of emotional cues: situational context and affective prosody, however, not all cues are weighted equally. For detecting the emotional states of others, some studies have found that people rely more on prosodic cues over situational context cues, but other studies have found the opposite [3]. The present study investigates whether emotional prosody or situational context is the driving factor when inferring other's emotions. Ten native speakers of English participated in a forced choice task. Participants were asked to infer a happy or sad emotional state of a character based on the given contextual information and the prosody used in an utterance spoken by the character. The utterances were recorded by male and female speakers and were paired with a visual scene. In the congruent condition, both context and prosody indicated the same emotional state, while in the incongruent condition, context and prosody indicated opposite states. If prosodic cues are more important than context, then the participants will pick the emotion that matches the utterance in the incongruent condition. Conversely, if contextual cues are more important than prosodic cues, then participants will pick the emotion that matches the story in the incongruent condition. In the congruent trials, as expected, participants accurately identified the emotional state of the character (happy = 100% correct; sad = 80%). In incongruent trials, participants chose the emotional state expressed by the prosodic cues more often than the emotion expressed by the context (happy prosody = 70%; sad prosody = 50%). When the data is broken down into whether a male or female delivered the stimulus item, participants relied on prosodic cues 70% of the time and sad prosody 50% of the time with the female voice, but 50% in happy and 20% in sad prosody 20% with the male voice. With these findings in mind, there may be an interaction between gender of the voice and the prosody expressed. Additionally, a plausible explanation for these findings could be due to an idea that when a prosodic cue isn't clear, participants may rely on the context, which may suggest that the two cues intertwine in language perception.

The Language Network in Task-Based fMRI
Presenter(s): John Shahki (UBC Alumni)

In-Person with Q&A

Keywords: Brain networks, Broca's and Wernicke's areas, linguistic stimuli, emotion recognition, fMRI-CPCA.

Introduction

Task-based functional magnetic resonance imaging (fMRI) reliably detects a brain network that simultaneously engages Broca's and Wernicke's areas. This network is left-lateralized during language tasks (Percival et al., 2020; Lavigne & Woodward, 2018), but is also elicited by a number of other tasks. This study aims to further elucidate the function associated with this network configuration by observing its activation over a range of cognitive tasks carried out during fMRI.

Methods

2-7 networks were extracted for each of 61 fMRI tasks using constrained principal component analysis for fMRI (fMRI-CPCA). Anatomical patterns were used to classify these networks into 11 prototype templates developed by the Cognitive Neuroscience of Schizophrenia Lab (Percival et al., 2020). The classifications were carried out using a MATLAB algorithm, which performs correlations between critical slices from these prototype templates and the parallel critical slices on the to-be-classified networks. The algorithm assigns 11 Fisher Z values to each to-be-classified network, and these can be inspected to determine if the to-be-classified network is simultaneously engaging Broca's and Wernicke's areas. Four networks from five tasks with the highest Fisher Z values for engaging Broca's and Wernicke's areas were selected for this study, and their activation and deactivation over these five tasks was inspected to determine the function of this network pattern. Descriptions of the fMRI tasks that produced the best matching networks are provided in Table 1.

fMRI Task	Description
Metrical Stress:	Participants (n=79) were shown 2-syllable words and were asked about the word's syllable stress and connotation.
Thought-Generation:	Participants (n=49) were presented with a word and its image, and were asked to either listen to, or generate, a definition for the word.
Semantic Association:	Participants (n=57) were shown three match options for a word prompt and were asked to select an option that most resembles the prompt.
Lexical Decision:	Participants (n=59) decided if visually presented letter strings represented a real word.
Emotion Discrimination:	Participants (n=70) evaluated whether a face reflected a particular emotion.

Table 1. fMRI task descriptions.

Results

All four networks exhibited activation in Broca's and Wernicke's areas, along with other characteristic activation patterns. The results for each network are provided in Table 2.

Table 2. Associated tasks, images, and Fisher Z values of the four extracted networks. Yellow depicts activation. Montreal Neurological Institute slice numbers are depicted above each slice. Characteristic activation patterns are highlighted in light blue. A Fisher Z value ≥ 0.80 indicates a significant correlation between the classified network, and prototype template.

Associated fMRI Task(s)	Images	Fisher Z Value	Observed Hemodynamic Effects
Merged Study: Metrical Stress, Thought-Generation, Semantic Association	Positive component loadings: [0.00, 0.34]	1.80	Increased activity with more linguistic demand, equivalent activation for phonological and semantic processing, more activation for volitional vs. passive extraction of meaning
Lexical Decision	Positive component loadings: [0.00, 0.35]	1.68	Suppressed when linguistic information is required to be ignored
Merged Study: Metrical Stress, Thought-Generation	Positive component loadings: [0.00, 0.34]	1.35	More activation for phonological vs. semantic processing, more activation for volitional vs. passive extraction of meaning
Emotion Discrimination	Positive component loadings: [0.00, 0.27]	0.95	No activation when no face, more for target emotion than foil, least activation for happy and most for fear

The characteristic activation patterns observed in all four networks, which are highlighted in light blue in the Table 2 images, are described in Table 3.

Orientation	Slice Number(s)	Activation Pattern Description
Coronal	20	Left-lateralized inferior frontal and frontal orbital activations
	-58	Temporal occipital and lateral occipital cortex activations
Axial	0	Left-lateralized middle and superior temporal gyrus, inferior frontal gyrus, and frontal pole activations
	-26, -22, -18	Lateral occipital cortex (inferior division) activations

Table 3. Descriptions of the characteristic activation patterns of this network configuration observed in all five tasks. Slice numbers are provided using the Montreal Neurological Institute convention.
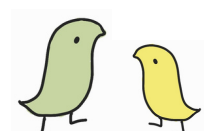
For the Metrical Stress, Thought-Generation, Semantic Association, and Lexical Decision tasks, predominantly left-lateralized activations were observed. Increased activations were observed during high linguistic demand and volitional interpretation of meaning. Suppression was observed when linguistic information was to be ignored. For the Emotion Discrimination task, bilateral activations were observed. Increased activations were observed when the target face was presented (Table 2).

Conclusion

A left-lateralized brain network that simultaneously engages Broca's and Wernicke's areas was observable during language tasks, but the same network configuration with a bilateral representation was observed during emotion recognition (Goghari et al., 2017). Based on the need of participants to acquire some form of meaning during the fMRI tasks, we conclude that this network's function is related to the extraction of linguistic- and emotion-based meaning. It should be noted that the four extracted networks overall represent a common network pattern of activated neurons, which was estimated using blood-oxygen-level-dependent signals while participants completed both language and emotion recognition tasks. Future work with a wider range of tasks will allow us to determine a wider range of other aspects of extraction of meaning with which this network is associated.

_____

QR CODE FOR LINKS

Please refer to the QR code below for access to Conference Zoom Link and digital posters.
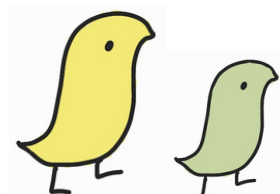
PLENARY SPEAKER PRESENTATIONS

Word Order or World Knowledge? The Role of Early Language On Resolving Conflicting Cues by Deaf Individuals

Presenter(s): Dr. Qi Cheng
Neuroplasticity & Language Lab) | Department of Linguistics, UW

Deaf individuals are more likely to have restricted early language environments due to the inaccessibility of spoken languages and the unavailability of a sign language. Lacking sufficient early language can result in deficits in both brain and language development. In this talk, I will look into the language processing patterns of deaf individuals with various early language exposure, with a focus on the ability to resolve syntactic and semantic cue conflicts. Behavioral studies using sentence-picture matching tasks in American Sign Language (ASL) and self-paced reading tasks in Chinese showed that even basic linguistic structures such as word order are less robust when early language is insufficient.

Biography
Dr. Qi Cheng is an Assistant Professor in Linguistics at the University of Washington, where she also directs the Neuroplasticity & Language Lab (NLL). Dr. Cheng's research mainly focuses on the role of early language experience on language and brain development, with special interests in sign languages, deaf language processing, and the neural foundations of human language. She uses both behavioral and neuroimaging methods to study the psycholinguistic and neurolinguistic aspects of human language learning and processing.
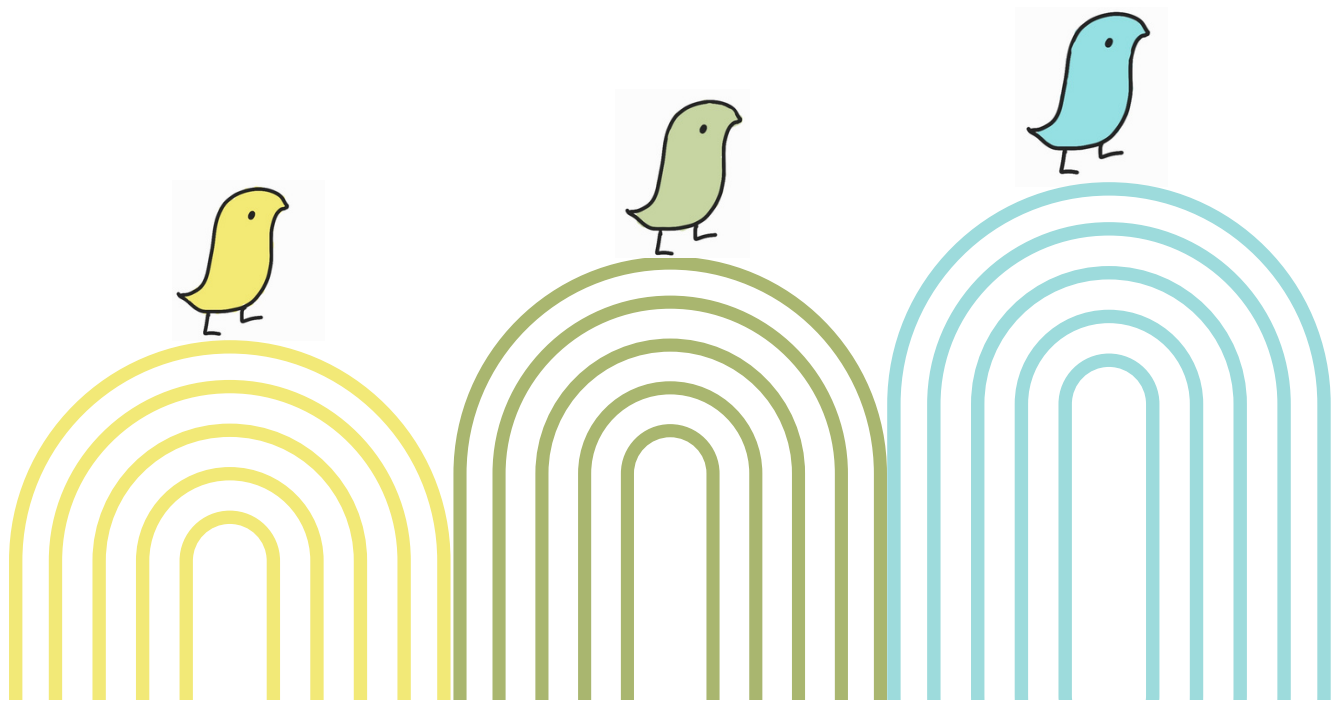
Aikamotsiipohtoo'p: Blackfoot Language Revitalization With Families and in The Home

Presenter(s): Dr. Heather Bliss (Department of Linguistics, SFU) & Annette Fox-BruisedHead

In this talk we tell the story of our collaborative work to inspire Blackfoot families with school-aged children to learn, use, and celebrate their ancestral language together.

Like many other Indigenous people across Canada, Blackfoot communities in Southern Alberta have experienced a significant disruption to patterns of intergenerational language transmission, due to damaging and genocidal strategies imposed by colonial governments and other institutions. The consequence for today's young Blackfoot families is that many parents do not identify as proficient speakers, although some may be silent speakers, who understand the language but do not speak it. Their children receive some language instruction at school, and our aim with this project is to bring the language beyond the school walls and back into the home where it can be shared and valued within families.

Towards this end, we have hosted a series of family nights that provide teachings and strategies to encourage language learning and use. We are inspired by the late Trish Rosborough's 'beautiful words' theory[1] of Indigenous language learning, which advocates for linguistically informed instruction that recognizes and teaches learners about the polysynthetic nature of many Indigenous languages in Canada. When learners can understand the complex morphological structures of the words they are learning, they can connect with their deeper meanings.

A relevant example comes from the title of our talk (and our project) - Aikamotsiipohtoo'p. This word refers to the process of language revitalization – but a morphological analysis reveals that it contains a root ikamo- 'give birth, escape danger, survive, recover,' which is also used often in prayer, as well as morphemes a- 'ongoing, continuous' and -o'p 'inclusive.' Our family nights include short lessons about the morphological make-up of words like aikamotsiipohtoo'p, and we invite families to reflect on the deeper meanings they can find within these words. The lessons are combined with language games that give learners opportunities to practice building words and recognizing patterns using the morphology they have learned.

The lessons and vocabulary introduced at each family night are uploaded to the Mioohpokoiiks website, a living resource for Blackfoot language teachers and learners, and we are also creating a family language learning guide that will be given to each family at the end of the project to inspire ongoing opportunities for language learning and creativity in the home.

1 Rosborough, Trish, chuutsqa Layla Rorick, and Suzanne Urbanczyk. 2017. Beautiful Words: Enriching and Indigenizing Kwak'wala Revitalization through Understandings of Linguistic Structure. The Canadian Modern Language Review 73 (4): 425-437.
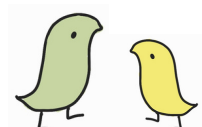
Biography – Heather Bliss

Oki, nitááníko Otskaapinaaki. Hello, my name is Heather Bliss. My Blackfoot name was given to me by my late friend and teacher Tootsinam Beatrice Bullshields. I am a non-Indigenous linguist living and working in the traditional and unceded territories of the hənq́ə́minə́m̓-speaking peoples. I was born and raised in Mohkinsstsis (Calgary, Alberta). I am an alumnus of the University of Calgary (BA Honours, 2003; MA 2005) and the University of British Columbia (PhD, 2013), and a former postdoctoral fellow at the University of Victoria (SSHRC 2014-16; Banting 2016-18). Today I am a half-time Lecturer at Simon Fraser University. My interest in Indigenous language revitalization stems from long-standing collaborative relationships with members of the Blackfoot-speaking Siksika and Kainai Nations in Southern Alberta. I am honoured to support Indigenous people, communities, and organizations through my research, writing, and educating on topics in language documentation, analysis, and revitalization.

Biography – Annette Bruised Head Fox

Annette Fox-BruisedHead is from Kainai First Nation. She began her formal teaching in 2002, as a classroom teacher teaching Blackfoot Language and Culture. After a decade working in leadership for the on-reserve education system, she returned to her initial school division where she began her teaching career. She took on the role as Division Principal, supporting the Board's goals for Indigenous education. It was at this time Annette was successful in securing an Indigenous Language in Education grant in 2019 for the Holy Spirit Catholic School Division in Southern Alberta, Canada. It was in this work that she met Dr. Heather Bliss and they completed the work for the grant developing the Mioohpokoiiksi site. In 2021 Annette successfully secured another grant, but this time as a board member for the Changing Horses NFP Society. She worked with Heather and the same school division to secure a Canadian Heritage Language grant. The work continues to develop the language resource within the site, working with Elders and families. In her spare time Annette spends time with her husband, Clark, their children, grandchildren, and in ceremony.
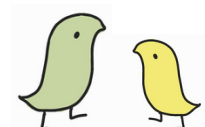
Social Perception and Evolution in the Use of Signed Languages

Presenter(s): Professor Nigel Howard
## *Department of Linguistics, UBC*

The presentation will focus on the change of social perception in the use of Signed Languages during pre- and post-COVID world. What may have or have not influenced a greater sense of awareness and acceptance of signed languages and Deaf Culture? The role of media/medium will also be discussed and how it influences society.

Biography

Nigel Howard is Adjunct Professor at the University of British Columbia in the Department of Linguistics. He also teaches at the University of Victoria in the Department of Linguistics. At Douglas College, he had worked in the Program of Sign Language Interpretation (INTR), and Child, Family and Community Studies (CFCS) for 25 years.

Nigel has been an interpreter for nearly 30 years in various setting such as medical, mental health, legal, theatre, community, and conference/workshop settings locally, nationally, and internationally. He is an Accredited International Sign Interpreter by the World Federation of the Deaf (WFD) and World Association of Sign Language Interpreters (WASLI). He is part of the United Nations team of International Sign interpreters and does interpret on media. Broadcasting, and translating are his forte.

Advocating for Deaf communities locally in British Columbia, nationally across Canada and globally around the world in terms of human rights, accessibility, and linguistic (signed languages) and Deaf culture.   Some activities are being an executive board member of the World Association of Sign Language Interpreters (WASLI) and part of the Editorial board for the International Journal of Interpreter Education (over 12 years).   Providing consultations, professional/educational workshops/trainings and presentations to various governments and organisations/associations has been an area to enhance better understanding, appreciation and awareness of Deaf culture and its signed language. It is a big part of his role as advocate, educator, and motivator. They are, for instance, "Global Diversity in Deaf Culture", and "Enculturation and Acculturation: Deaf and Hearing Worlds/Spaces".
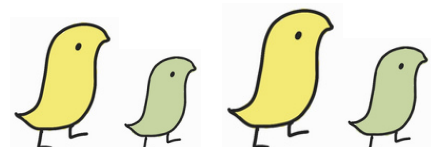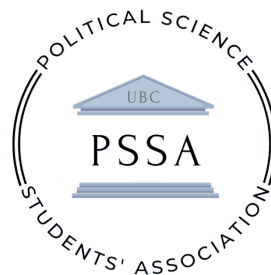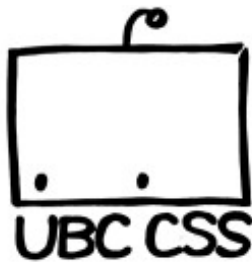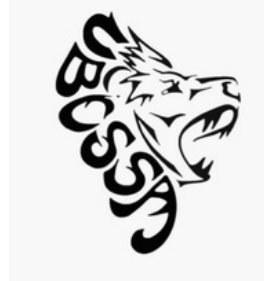
ACKNOWLEDGEMENTS

Prelude Talks by:

- Dr Antone Minard
- Dr Willis Monroe
- Dr Lisa Sundstrom
- Dr Julian Diekes
- Rabbi Kylynn Cohen
- Clinton Macadam

LSURC
Language Sciences Undergraduate
Research Conference

SPONSORS

Chun Fun How

Steve's POKÉ BAR

UBC Global Lounge

BUCHA BREW
MADE IN VANCOUVER, B.C.

WIZE TEA

THANK YOU

SPECIALS THANKS TO OUR ATTENDEES, SPONSORS, AFFILIATED GROUPS, & VOLUNTEERS FOR SUPPORTING US AND MAKING LSURC 2023

PLANNING COMMITTEE

CO-CHAIRS

Sherouk Shehata

Jonathan Elmer

TREASURERS

Flora Zhi

Vanessa Lee

HR CONSULTANTS

Mina Han

Zoe Chung

ADMINISTRATIVE COORDINATORS

Camille Savalli

Jody Yuen

Max Hong

Sarah Ong

VICE-CHAIR & SECRETARY

Flora Zhi

EXECUTIVE ASSISTANTS

Hasti Nozadi

Maria-Elena Mishchenko

SOCIAL MEDIA

Hayley Hau

Tia Wong

OUTREACH COORDINATORS

Ben Chinnik

CC Liang

Rachelle Chen

SALSA LIAISON

Hayley Hau