Does a code-switching dialogue system help users learn conversational fluency in Choctaw?

Jacqueline Brixey
USC Institute for Creative Technologies
brixey@usc

David Traum
USC Institute for Creative Technologies
traum@ict.usc.edu

Abstract

We investigate the learning outcomes and user response to a chatbot for practicing conversational Choctaw, an endangered American Indigenous language. Conversational fluency is a goal for many language learners, however, for learners of endangered languages in North America, access to fluent speakers may be limited. Chatbots are potentially ideal dialogue partners as this kind of dialogue system fulfills a non-authoritative role by focusing on carrying on a conversation as an equal conversational partner. The goal of the chatbot investigated in this work is to serve as a conversational partner in the absence of a fluent Choctaw-speaking human interlocutor. We investigate the impact of code-switching in the interaction, comparing a bilingual chatbot against a monolingual Choctaw version. We evaluate the systems for user engagement and enjoyment, as well as gains in conversational fluency from interacting with the system.

1 Introduction and Motivation

Conversational fluency is a goal for many language learners. However, for learners of endangered languages like Choctaw, access to fluent speakers may be limited. This lack of access may be due to geographical features, such as not living on or near tribal lands, or because there are few remaining fluent speakers of the language. It is unclear how many Indigenous languages are still spoken today in the United States; one source (Moseley, 2010) estimated there were 256 in 2010, while the 2010 US census estimated 165[1]. At the time of writing, no similar summary could be found for the results of the 2020 census. However, it is anticipated that the number of speakers has declined over time (Simons and Fennig, 2018), particularly after the devastating effects of the COVID-19 pandemic (Healy and Blue; Rogers), thus support for learning these languages is time critical.

The goal of the chatbot investigated in this work is to serve as a conversational partner in the absence of a fluent Choctaw-speaking human interlocutor. The goal of the interaction is for the user to gain conversational fluency in Choctaw, such as through increased vocabulary or greater sense of ease, by interacting with the system. We compare a monolingual version of the chatbot against a code-switching one.

This work builds on our previous work on Masheli, a simplified Choctaw-English code-switching chatbot (Brixey and Traum, 2021). However, we address several new questions, such as: Will code-switching lead to a better user experience? Will users show a higher preference for the code-switching chatbot? Will code-switching improve the learning outcomes? Will Indigenous language learners want to use this technology? Our results indicate that interactions with the code-switching chatbot suggest a slight improvement in user experience but did not find significant learning benefits compared to the monolingual chatbot.

2 Relevant Literature

Technology developed for learning purposes, especially language learning, is a well-established area of research. Technology, particularly dialogue systems, has been implemented in this sphere for several reasons. While traditional classroom settings may attempt to create conversational opportunities, many student factors, such as shyness or fear of making errors, can prevent learners from engaging fully in conversation with a human partner (Shawar and Atwell, 2007). Chatbots are well suited for language learning environments since they can serve as an equal conversational partner without expectations of explicit correction on errors (Chou et al., 2003), and learners have reported

[1] https://www2.census.gov/library/publications/2011/acs/acsbr10-10.pdf

Proceedings of the Fifth Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP), pages 8–17

feeling more comfortable chatting with a dialogue system than with a human interlocutor (Fryer and Carpenter, 2006).

2.1 Second language acquisition literature

Second language acquisition is learning a second language other than the first after the first language has been acquired (Ortega, 2014). Theories, frameworks, and descriptions for second language acquisition abound (For a more detailed overview, see Ortega (2014); Mitchell et al. (2013); Lightbown and Spada (2013)). This section is thus related to Indigenous language learning and systems and emerging bilingual conversational behaviors and pedagogy that supports these behaviors.

2.1.1 Indigenous language pedagogy and language learning systems

Learning an Indigenous language differs from other second languages due to factors like a limited number of fluent speakers, often dispersed geographically, and the dominance of English in many Indigenous communities (2015). Language suppression and forced cultural assimilation have contributed to these challenges, along with a lack of published literature and media in the language. Moreover, the scarcity of learning opportunities and spaces to practice the language, even on reservations, further complicates revitalization efforts (White, 2006). Additionally, Indigenous language teachers may not always have formal training in pedagogy, and there may not be enough instructors to meet the growing demand for learners of all ages (Lukaniec and Palakurthy, 2022).

Technology has become a significant tool for overcoming some of these challenges, providing new opportunities for language learning and connecting speakers across geographical distances (Cassels and Farr, 2019). While technology alone cannot revitalize a language, it can supplement the efforts of motivated learners and serve as one of many tools for language revitalization (Cassels and Farr, 2019). The Choctaw Nation of Oklahoma has long utilized technology for language teaching, from early telecourses to more recent Zoom classes, which became especially popular during the pandemic and continue to thrive today [2]. As Mark Turin, former chair of the First Nations and Endangered Languages Program at the University of British Columbia, states, "tools and technology don't save language — speakers do" (Karstens-Smith).

2.1.2 Emerging Bilingual Conversational Behaviors and Translanguaging

Emerging bilinguals often code-switch, combining elements from different languages to communicate, even in non-grammatical ways, but they still co-construct meaning with interlocutors (Cenoz and Gortegaorter, 2017; Canagarajah, 2011). This is common in casual conversations where interlocutors share multiple languages and the language choice is not fixed (Auer, 1995). While learning a language is ultimately an individual endeavor, supportive pedagogy can enhance the process. Traditional immersion pedagogy required learners to interact only in the target language, but translanguaging—intentionally using multiple languages in a learning environment—has become more widely accepted in second-language pedagogy, especially for teaching endangered languages (Cenoz and Gortegaorter, 2017).

The literature differentiates code-switching from translanguaging. Code-switching involves shifting between languages in any conversational setting, while translanguaging encourages emerging bilinguals to use all their languages purposefully in a learning setting, with the instructor gradually reducing support as learners progress (Cenoz and Gortegaorter, 2017; Makalela, 2015). Originating in bilingual English-Welsh education, translanguaging emphasizes interaction and participation, even if not entirely in the target language, allowing learners to use other languages to fill gaps in their knowledge (Makalela, 2015; García, 2009). This contrasts with immersion-style teaching, which often discourages or ignores the use of the non-target language.

In monolingual settings, emerging bilinguals often avoid addressing their language confusion, hoping that future encounters or additional context in the same conversation will provide clarification (Canagarajah, 2011). This is known as the "let it pass" principle (Firth, 1996), the act of not addressing misunderstandings, which can hinder comprehension if additional examples do not occur. However, classrooms using translanguaging have seen better outcomes for second-language learners, as fewer "let it pass" instances happen (Champlin, 2016). While translanguaging is frequently considered a verbal act (Canagarajah, 2011), the literature supports translanguaging in text form. For example,

[2] https://www.choctawnation.com/about/language/classes/

Māori literacy improved when students used English to process Māori texts (Lowman et al., 2007). Translanguaging has also been shown to be psychologically beneficial for emerging bilinguals. It is suggested to legitimize a student's relationship with both languages and foster self-identification as a speaker of both languages (Makalela, 2015), while encouraging the use of all linguistic resources, rather than suppressing specific repertoires, can enhance students' self-confidence.

There are strategies to use translanguaging in a learning environment effectively. The most common approaches emphasize linking translanguaging to content in lessons, such as important vocabulary, and that the instructor should utilize translanguaging and encourage its use by individual students and within groups (Cenoz and Gortegaorter, 2017; Dougherty, 2021; Seals and Olsen-Reeder, 2020).

2.2 Hypotheses

To summarize the prior research, translanguaging and using an already known language can enhance a learner's learning gains and sense of comfort in a classroom setting with human-human interactions (Butzkamm and Caldwell, 2009). The literature also shows that code-switching can lessen the feeling of distance between conversational human interlocutors. Based on the literature, the hypotheses for this experiment are as follows.

1. H1: Code-switching bilingual chatbots that use translanguaging techniques and code-switching frameworks lead to a better learning experience, possibly through learning gains or a greater sense of rapport, comfort, or enjoyment for language learning users.
2. Users will demonstrate the highest learning gains with a code-switching system.
3. Users will have a lower user experience with the monolingual system than with the code-switching bilingual system.

3 System Design

For this work, we implemented two chatbots: a monolingual Choctaw version, and an English-Choctaw code-switching one. The backend of the chatbots is NPCEditor, a response classifier and dialogue management system (Leuski and Traum, 2011). NPCEditor uses a statistical classifier that is trained on linked questions and responses. The classifier is trained on a question-answer (QA) corpus.

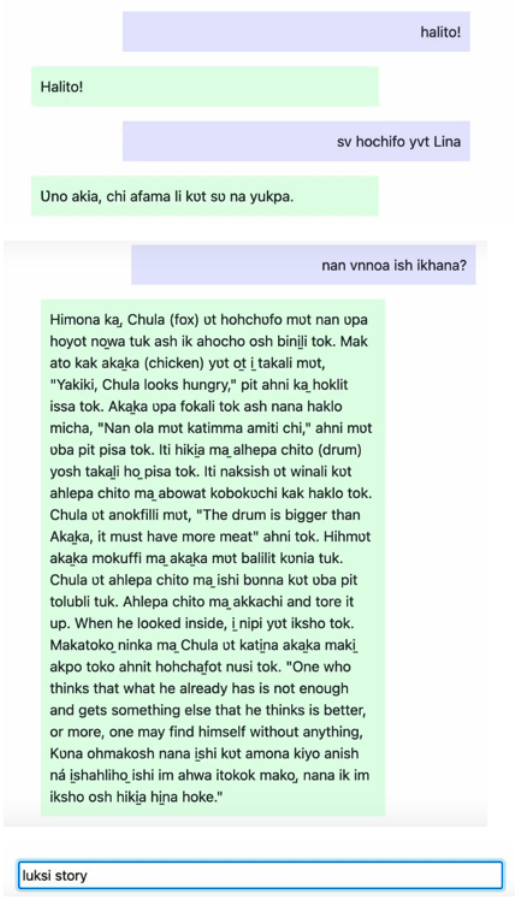

Figure 1: Example conversation with the chatbot.

For each user input, the classifier ranks all the available responses. NPCEditor also contains a dialogue manager, which selects an appropriate response from the ranked responses. Previous applications of NPCEditor have been used for interactive characters in multiple domains, such as interviews with Holocaust survivors (Traum et al., 2015). This was also the backend for an earlier version of Masheli (Brixey and Traum, 2021).

An example dialogue with the code-switching chatbot is in Figure 1, demonstrating some greetings (the first two complete turns) and then telling a story about a fox in Choctaw.

We elected to use NPCEditor and handcrafted utterances over LLMs or other approaches for two primary reasons. First, we wanted to implement a consistent strategy for code-switching, which we found LLMs struggled to produce reliably. Second, through experimentation, we discovered that LLMs often failed to generate syntactically correct Choctaw utterances. Since one of the chatbot's main goals is to help learners improve their language fluency, providing incorrect Choctaw would contradict that objective.

3.1 QA Corpus

Each question in the QA corpus is matched to at least one appropriate answer that serves as a response for the chatbot. There is no explicit module for recognizing the language in which the user is communicating. The knowledge base of the chatbot is sharing stories about animals. We made this selection because pedagogy literature, especially for American Indigenous languages (Cantoni, 1999), indicates that story-based instruction is beneficial in language learning environments (Kickham, 2015; Andrews et al., 2009).

3.1.1 Questions

We implemented a Python script to generate questions for the question portion of the QA corpus. The script included several sentences with predominantly English syntax, such as "Can I have a story about ..." or "Tell me about ..." and the list of animals from the stories in Choctaw to be added at the end of the sentence. The result produced a sentence like "Can I have a story about shawi?" (Can I have a story about raccoons?)

The monolingual chatbot version was intended to mimic an immersion-style pedagogy, so we only added a handful of English and code-switched sentences. Most of these were mapped to an off-topic response encouraging the user to speak in Choctaw. This type of response aligns with the immersion-style curriculum, which will ignore or discourage statements made in the non-target language.

3.1.2 Answers

To form the chatbot's domain knowledge, ten animal stories were selected from ChoCo (Brixey et al., 2018), a Choctaw language corpus. All stories are originally in Choctaw and have English translations. We created handcrafted responses for the two chatbots. To incorporate translanguaging strategies in the code-switching chatbot, we repeated key vocabulary to understand the story in English in parentheses. Repetition was one non-spontaneous strategy for effective translanguaging (Seals and Olsen-Reeder, 2020). The examples in Table 1 show how code-switching and translanguaging were incorporated into a given line in a story.

Code-switching was generated in two options, insertional and switching at clauses, which follows the linguistic literature on code-switching and the model described in Ahn et al. (2020). There were two options for the matrix language, either Choctaw or English. Not every sentence in a story includes code-switching. Instead, we aimed for roughly 75% of a given Choctaw story to have code-switching.

3.2 Dialogue manager

The dialogue manager can choose a lower-ranked response to avoid repetition. If the score of the top-ranked response is below the threshold that was selected during training, the dialogue manager will instead select a response that indicates non-understanding or that aims to end a conversation topic. For example, the expression "Mihacha?" ("It really is, isn't it?") might be selected as a response when no other response scores above the threshold.

3.3 Orthographic considerations

One challenge to support Choctaw is that the language does not have a fully standardized written form. Each training example in the question portion of the QA corpus was written in multiple formats to support many different possible orthographic presentations. For example, the sentence "Do you know a story about a woodpecker?" could be written with different formats of nasalized characters a̲ and i̲:

1. Biskinik am anumpa nan anoli ish i̲shi?
2. Biskinik a anumpa nan anoli ish i̲shi?
3. Biskinik an anumpa nan anoli ish i̲shi?
4. Biskinik a̲ anumpa nan anoli ish i̲shi?
5. Biskinik a̲ anumpa nan anoli ish inshi?

4 Methods

This section discusses consultations with the Choctaw Nation of Oklahoma, the IRB review process, and how we ensured tribal data rights. We also describe methods to assess the user experience: a language test to evaluate the user's learning and a survey to gauge their sense of rapport, comfort, and enjoyment.

4.1 Tribal review

Several steps are required to conduct research on the Oklahoma Choctaw language or with Choctaw tribal members. First, a sponsor must review and support the work. A sponsor must be someone who works for the tribal nation. The sponsors for this work evaluated the proposal for sensitivity to the community, adequate protection of tribal members, and alignment with tribal initiatives. Following a sponsor's approval and support, we then applied to

English	One day a man riding in a boat came to the end of the water.
Monolingual Choctaw	Mak atokọ nittak himona kạ hattak mvt oka peni fokka osh ont aivhli mạ ona tok.
Insertional-Cho matrix	Mak atokọ nittak himona kạ hattak mvt **a boat** fokka osh ont aivhli mạ ona tok.
Clausal-Cho matrix	**One day**, hattak mvt oka peni fokka osh ont aivhli mạ ona tok.
Insertional-Eng matrix	**One day a man riding in** oka peni **came to the end of the water**.
Clausal-Eng matrix	Mak atokọ nittak himona kạ **a man riding in a boat came to the end of the water**.
Repetition	**One day a man riding in** oka peni **(a boat) came to the end of the water**.

Table 1: Framework-based utterances examples. English portions are bolded in code-switched utterances.

Choctaw Nation's IRB. Our university's IRB then reviewed and approved the protocol.

4.2 Language Test

We created a 15-question language test to be administered before and after the interaction. The test determines whether learners gained any new vocabulary ("What is the word for deer in Choctaw?", 12 questions) or any new syntax ("How would you say, 'Do you know a story about deer?' in Choctaw?", three questions).

The language test also served to inform all participants about the chatbot's domain knowledge of animal stories, a fact given in the instructions read to each participant, so that participants would have more consistent experiences and not have to spend time discovering which stories the chatbot knows.

4.3 Survey design

The survey was designed to evaluate the user's sense of rapport, the naturalness of the code-switching, and the feeling of connection because of language identity.

The survey consisted of twelve 5-point Likert scale questions, and the answers were scored from 1 strongly disagree to 5 strongly agree. Many questions came from previous research on rapport (Novick and Gris, 2014; Gratch et al., 2007). Questions 7 and 10 are novel and tailored to this experiment. All survey questions were optional, and participants could choose to skip any questions.

1. The system understood me.
2. The system seemed unengaged.
3. The system was friendly.
4. The system and I worked towards a common goal.
5. The system and I did not seem to connect.
6. I didn't understand the system.
7. The system knows the Choctaw language.
8. The interaction was interesting.
9. The interaction felt natural.
10. I felt the system and I were in the same social group.
11. I would be willing to continue the conversation with the system for longer.
12. I would recommend interacting with this system to a friend.
13. Was there anything else that you wanted to talk to the system about? (open-ended)
14. Do you have any other comments to share about your experience? (open-ended)

Questions were selected to determine levels of rapport (1, 2, 4, 5, 6, 9) and engagement and connection (3, 8, 10, 11, 12). We hypothesized that the code-switching cohort would score the chatbot higher on these questions. The survey also measured people's perception of the chatbot's knowledge of the Choctaw language (7) to gauge how users perceived the fluency of the chatbot's code-switching.

4.3.1 Experiment session

Participants began by reading and signing a consent form, followed by an oral explanation. The experiment started with the language test, after which participants interacted with the chatbot for 15 minutes. They then completed the language test again and finished with a post-interaction survey to rate their experience and provide comments. Participants were encouraged to have a dictionary on hand; if not, they were given links to two online dictionaries, a 1915 publication (Byington, 1915) and a 2016 publication (The Choctaw Nation of Oklahoma Dictionary Committee, 2016).

4.3.2 Inclusion and exclusion criteria

Inclusion and exclusion criteria, in this case, specify which individuals from the participant population are eligible or ineligible to be included in the research study. The inclusion criteria required participants to follow instructions, engage meaningfully with tasks, and provide on-topic interactions

with the chatbot. They were instructed to communicate with the session leader only for questions or technical issues. Participants were expected to complete all tasks, including language proficiency tests, surveys, and structured interactions with the chatbot, ensuring data integrity. While participants were not required to spend the full 15 minutes interacting with the chatbot and would not be excluded for finishing early, they were encouraged to take time referencing the dictionary. No specific number of chatbot interactions was required, but at least one turn was necessary to demonstrate participation. Exclusion criteria included multiple off-topic utterances, inappropriate comments, or off-topic survey responses. Non-engagement was identified as discussing unrelated topics with the session leader, except for technical issues or clarifications.

5 Results

In total, 23 participants completed the experiment. Twelve participants interacted with the monolingual Choctaw chatbot, while eleven participants interacted with the framework-based code-switching chatbot. One participant from the monolingual chatbot met the exclusion criteria, so their survey and language test responses were omitted.

Two participants requested to finish the chat portion early. Their data was retained as they followed all protocols and engaged with the chatbot, albeit for less time. One participant using the monolingual chatbot ended the session after 6 minutes due to frustration, while another using the code-switching chatbot ended it after 13 minutes, citing frustration and disinterest. Many participants asked the chatbot for definitions and translations despite having a dictionary, suggesting future work could include providing these directly.

5.1 Language Test

All language tests (pre- and post-test) were scored for two factors. The first factor was how many questions were attempted, regardless of correctness. The second factor was correctness. A correct answer was one point; thus, a perfect score on the quiz would be 15.

For the first 12 questions on the language test, we applied a rubric for grading the questions. Since Choctaw is not standardized and can require a keyboard with the unique characters, we made allowances for differences in spelling. Half a point was deducted if an extra syllable was added, a vowel was sufficiently incorrect to impact the pronunciation, or a consonant was substantially incorrect. Likewise, half of a point was deducted for the syntax questions if the words were correct, but the ordering was off, or the pronoun was incorrect.

Next, we evaluated the average change for attempted and correct responses. The average change in the number of vocabulary questions attempted was 1.18 for the monolingual group and 1.36 for the code-switching group. This indicates that the code-switching group was slightly more inclined to try more questions after interacting with the chatbot. The average change in correct answers for vocabulary questions for the monolingual group was 1.5, while the code-switching group was 1.36. This indicates that all groups benefited from the interaction, with the monolingual group improving slightly more. No participants had decreased test scores. Several participants in both groups showed no improvement via the language test. The participant with the greatest improvement was in the code-switching group, with a gain of 4.5 points between pre- and post-interaction tests, and this participant also showed the greatest change in the number of questions attempted. For the grammar questions of the language test, the monolingual group showed an average increase of 0.1 in the number of grammar questions attempted. In contrast, the code-switching group had an average increase between pre- and post-grammar questions attempted of 0.09. The monolingual group improved in correct responses on average by 0.2 points, while the code-switching group improved by 0.18.

An overall positive finding is that learning occurred with both chatbots. Additionally, we did not observe a significantly higher level of learning with the monolingual chatbot, the "immersion" style of learning, over the translanguaging, code-switching chatbot.

5.2 Responses to survey

The results of comparing the two groups' survey responses using a one-tailed T-test are shown in Table 2. The table also shows the average score for each group, with the standard deviation given in parentheses next to the mean value.

We observed two $p<0.10$ values: (1) The code-switching group scored their chatbot as friendlier than the monolingual group, and (2) the code-switching group reported that they would be more likely to recommend the system to others.

	Question	T-test result	Mean Mono (std dev)	Mean CSW (std dev)
1	The system understood me.	0.25	3(1)	3.54(1.03)
2	The system seemed unengaged.	0.73	2.54(1.50)	2.36(1.36)
3	The system was friendly.	*0.07	3.18 (1.47)	4.36 (0.80)
4	The system and I worked towards a common goal.	0.65	3.36 (1.36)	3.54 (0.82)
5	The system and I did not seem to connect.	0.34	2.90 (1.51)	2.45 (1.12)
6	I didn't understand the system.	1	2.54 (1.29)	2.54 (1.29)
7	The system knows the Choctaw language.	0.64	3.9 (0.87)	4 (0.89)
8	The interaction was interesting.	0.16	3.63 (1.36)	4.36 (0.92)
9	The interaction felt natural.	0.19	2.81 (1.32)	3.45 (0.82)
10	The system and I were in the same social group.	0.16	2.45 (1.21)	3.18 (1.16)
11	I would be willing to continue the conversation with the system for longer.	0.13	3.63 (1.74)	4.45 (0.52)
12	I would recommend interacting with this system to a friend.	*0.06	3.63 (1.62)	4.54 (0.52)

Table 2: The results of comparing survey responses between the monolingual and code-switching interactions. p<0.10 results are marked with one asterisk. Standard deviations are given in parentheses next to the average in the final two columns.

We then analyzed the survey responses by clustering the questions by rapport (1, 2, 4, 5, 6, 9) and engagement and connection (3, 8, 10, 11, 12). We then summed the scores for each participant in the given cluster. We reversed the polarity for negatively phrased questions (2, 5, 6). The p-value for the clustered questions on rapport was 0.24. The p-value for engagement and connection was 0.04, a significant value.

6 Discussion

First, we will review the findings for the main research questions.

- *Will code-switching lead to a better user experience? Will users show a higher preference for the code-switching chatbot?*
  The survey results indicate that users had a better, more satisfying experience with the code-switching, translanguaging chatbot.
- *Will code-switching lead to an increase in learning?*
  The language tests indicate that participants learned new vocabulary while interacting with the code-switching chatbot. However, they did not learn significantly more than the monolingual group, indicating that interacting with any chatbot will lead to a learning experience.
- *Will Indigenous language learners want to use this technology?*
  Some participants expressed interest in interacting with the chatbot again, we invited them to chat with it again during one weekend over the month that experiments were held. The conversations that day were recorded via a log, but no information was noted about who spoke to the chatbot at any given time.

Now, we will review the findings in relation to the hypotheses.

H1: Code-switching bilingual chatbots that use translanguaging techniques and code-switching frameworks lead to a better learning experience, possibly through learning gains or a greater sense of rapport, comfort, or enjoyment for language learning users.

The code-switching chatbot followed translanguaging principles in its code-switching but also followed linguistic frameworks that produced insertional and clause switches. The survey results suggest that modeling code-switching aspects using linguistic frameworks leads to higher levels of reported rapport and enjoyment. The final question of the survey indicates that the code-switching cohort would be more likely to interact with the chatbot again; thus, it is possible that learning gains could be achieved over multiple interactions.

Survey results also show that participants found the code-switching chatbot more enjoyable and better suited as a language partner for Choctaw learners, with many describing it as friendlier. This aligns with the literature on face, suggesting that participants felt their face was threatened when the chatbot didn't understand their Choctaw attempts. Face is the image one has of oneself and emerges during interactions (Haugh, 2009). Face is important in any conversation as humans want to be liked and respected by others, but face is a key factor in learning scenarios (Wang et al., 2008) and particularly in second language conversations (Piirainen-Marsh, 1995; Ahvenainen, 2021).

H2: Users will demonstrate the highest learning gains with a code-switching system.

The language test results show no significant learning gains with immersion (interacting with the

monolingual chatbot), either in the number of correct answers or the number of attempted questions. Participants interacting with the code-switching chatbot had slightly lower correct answers but were more likely to attempt vocabulary questions, possibly due to increased confidence. These results suggest that both chatbots lead to positive learning gains.

H3: Users will have a lower user experience with the monolingual system than with the code-switching bilingual chatbot.

Based on the $p<0.10$ results for questions 3 and 12 on the user survey, a preference was observed for the code-switching system.

Literature on face and face-work explains why participants preferred the bilingual chatbot. As face is tied to emotional reactions, it can be threatened in language learning, leading to frustration, shame, or anger (Spencer-Oatey, 2007; Holtgraves, 2009; Ting-Toomey, 2009). Interlocutors are expected to protect each other's face (Holtgraves, 2009), thus users interacting with the monolingual chatbot may have felt rejected by its seeming disapproval of their English or non-standard Choctaw.

7 Conclusion and Future Directions

In this work, we tested a novel code-switching Choctaw language chatbot and the impact of code-switching on learning. As the language is endangered, effective revitalization efforts are time-critical. The results of our study indicate that users prefer the code-switching chatbot over the monolingual one based on the survey responses, which could have implications for maintaining long-term learning motivation and interest. Both cohorts demonstrated learning gains from the interaction in the form of a vocabulary and grammar quiz, with the monolingual cohort learning just slightly more but not significantly more. Our contributions include novel insights into the user experience of interacting with a code-switching dialogue system, a chatbot capable of responding to code-switched user input, a schema for chatbot responses using linguistic frameworks and translanguaging techniques, and a corpus of learning users' conversations with the chatbot. Choctaw learners have received little study, and the conversation logs could serve as a meaningful resource for language instructors and linguists.

One possibility for future work is to evaluate the learning gains over a longer period to determine if additional time spent interacting with the chatbot or over several sessions could produce strongly significant results, either on the survey or language test. It is possible that retention would be higher with the group paired with the code-switching chatbot, given the higher survey scores, and with higher retention, the possibility of higher learning. A final consideration is that replication is needed for other language communities to confirm that the results found here are not unique to the Choctaw language.

8 Limitations

One computational limitation of Masheli was that the system could not process some of the unique characters of participants' input. The system was trained using specific ASCII characters but had not been trained on some of the other possible ASCII variations. Additionally, some of the characters did not render correctly for unclear reasons, such as a̲ sometimes presented as å. Participants were encouraged during the experiment to use alternative spellings if the system could not process their original statement; however, this may have impacted user satisfaction.

9 Ethics

This work was completed with consultation and review from the Choctaw Nation of Oklahoma (see Section 4.1 for more details). All of the collected data from this research was requested to be archived at the Choctaw Nation's Cultural Center archives to ensure that the tribe would continue to benefit from this effort.

References

Emily Ahn, Cecilia Jimenez, Yulia Tsvetkov, and Alan W Black. 2020. What code-switching strategies are effective in dialog systems? In *Proceedings of the Society for Computation in Linguistics 2020*, pages 213–222.

Tarmo Ahvenainen. 2021. Language proficiency facework and perceptions of language proficiency face in l2 interaction. *JYU dissertations*.

Dee H Andrews, Thomas D Hull, and Jennifer A Donahue. 2009. Storytelling as an instructional method: Descriptions and research questions. *The Interdisciplinary Journal of Problem-based Learning• volume*, 3(1):6–28.

Peter Auer. 1995. The pragmatics of code-switching: a sequential approach. In Leslie Milroy and Pieter Muysken, editors, *One speaker, two languages*, chapter 6, pages 115–135. University of Cambridge.

Jacqueline Brixey, Eli Pincus, and Ron Artstein. 2018. Chahta anumpa: A multimodal corpus of the Choctaw language. In *Proceedings of LREC 2018*, Miyazaki, Japan.

Jacqueline Brixey and David Traum. 2021. Masheli: A Choctaw-English bilingual chatbot. In *Conversational Dialogue Systems for the Next Decade*, pages 41–50. Springer.

Wolfgang Butzkamm and John AW Caldwell. 2009. *The bilingual reform: A paradigm shift in foreign language teaching*. Narr Francke Attempto Verlag.

Cyrus Byington. 1915. *A Dictionary of the Choctaw Language*. US Government Printing Office. Edited by John R. Swanton and Henry S. Halbert. Smithsonian Institution Bureau of American Ethnology Bulletin 46.

Suresh Canagarajah. 2011. Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied linguistics review*, 2(1):1–28.

Gina P Cantoni. 1999. Using tpr-storytelling to develop fluency and literacy in native american languages. *Revitalizing Indigenous Languages*, page 53.

Morgan Cassels and Chloë Farr. 2019. Mobile applications for indigenous language learning: Literature review and app survey. *Working Papers of the Linguistics Circle*, 29(1):1–24.

Jasone Cenoz and Durk Gortegaorter. 2017. Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development*, 38(10):901–912.

Molly J Champlin. 2016. *Translanguaging and bilingual learners: A study of how translanguaging promotes literacy skills in bilingual students*. Ms thesis, St. John Fisher College.

Chih-Yueh Chou, Tak-Wai Chan, and Chi-Jen Lin. 2003. Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40(3):255–269.

Tyne Crow and David Parsons. 2015. A mobile game world for māori language learning. In *International Conference on Mobile and Contextual Learning*, pages 84–98. Springer.

Jessica Dougherty. 2021. Translanguaging in action: Pedagogy that elevates. *ORTESOL Journal*, 38:19–32.

Alan Firth. 1996. The discursive accomplishment of normality: On 'lingua franca'english and conversation analysis. *Journal of pragmatics*, 26(2):237–259.

Luke Fryer and Rollo Carpenter. 2006. Bots as language learning tools. *Language Learning & Technology*, 10(3):8–14.

Ofelia García. 2009. Education, multilingualism and translanguaging in the 21st century. In *Social justice through multilingual education*, pages 140–158. Multilingual Matters.

Jonathan Gratch, Ning Wang, Jillian Gerten, Edward Fast, and Robin Duffy. 2007. Creating rapport with virtual agents. In *Intelligent Virtual Agents: 7th International Conference, IVA 2007 Paris, France, September 17-19, 2007 Proceedings 7*, pages 125–138. Springer.

Michael Haugh. 2009. Face and interaction. In Michael Haugh and Francesca Bargiela-Chiappini, editors, *Face, communication and social interaction*, chapter 1, pages 1–30. Equinox London.

Jack Healy and Victor J. Blue. Tribal elders are dying from the pandemic, causing a cultural crisis for American Indians. *The New York Times*.

Thomas Holtgraves. 2009. Face, politeness and interpersonal variables: implications for language production and comprehension. In Michael Haugh and Francesca Bargiela-Chiappini, editors, *Face, communication and social interaction*, chapter 10, pages 192–207. Equinox London.

Gemma Karstens-Smith. B.C. Teen creates app to help revive fading Indigenous language. *Toronto Star*.

Elizabeth A Kickham. 2015. *Purism, Prescriptivism, and privilege: Choctaw language ideologies and their impact on teaching and learning*. Ph.D. thesis, UNIVERSITY OF OKLAHOMA.

Anton Leuski and David Traum. 2011. Npceditor: Creating virtual human dialogue using information retrieval techniques. *Ai Magazine*, 32(2):42–56.

Patsy M Lightbown and Nina Spada. 2013. *How Languages are Learned 4th edition-Oxford Handbooks for Language Teachers*. Oxford University Press.

Chris Lowman, Tangimai Fitzgerald, Patsy Rapira, and Rahera Clark. 2007. First language literacy skill transfer in a second language learning environment: Strategies for biliteracy. *Set*, 2:24–28.

Megan Lukaniec and Kayla Palakurthy. 2022. Additional language learning in the context of indigenous language reclamation. In *The Routledge handbook of second language acquisition and sociolinguistics*, pages 341–355. Routledge.

Leketi Makalela. 2015. Translanguaging as a vehicle for epistemic access: Cases for reading comprehension and multilingual interactions. *Per Linguam: a Journal of Language Learning= Per Linguam: Tydskrif vir Taalaanleer*, 31(1):15–29.

Rosamond Mitchell, Florence Myles, and Emma Marsden. 2013. *Second Language Learning Theories*. Routledge.

Christopher Moseley. 2010. *Atlas of the World's Languages in Danger*. Unesco.

David Novick and Iván Gris. 2014. Building rapport between human and eca: A pilot study. In *Human-Computer Interaction. Advanced Interaction Modalities and Techniques: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part II 16*, pages 472–480. Springer.

Lourdes Ortega. 2014. *Understanding second language acquisition*. Routledge.

Arja Piirainen-Marsh. 1995. Face in second language conversation. *JYU dissertations*.

Mike Rogers. Choctaw Nation members talk about impact of losing native speakers to COVID-19. *KXII.com*.

Corinne A Seals and Vincent Olsen-Reeder. 2020. Translanguaging in conjunction with language revitalization. *System*, 92:102277.

Bayan Abu Shawar and Eric Atwell. 2007. Fostering language learner autonomy through adaptive conversation tutors. In *Proceedings of the The fourth Corpus Linguistics conference*.

Gary F. Simons and Charles D. Fennig, editors. 2018. *Ethnologue: Languages of the World*, twenty-first edition. SIL International, Dallas, Texas.

Helen Spencer-Oatey. 2007. Theories of identity and the analysis of face. *Journal of pragmatics*, 39(4):639–656.

The Choctaw Nation of Oklahoma Dictionary Committee. 2016. *Chahta Anumpa Tosholi Himona: New Choctaw Dictionary*, 1st edition. Choctaw Print Services.

Stella Ting-Toomey. 2009. Facework collision in intercultural communication. In Michael Haugh and Francesca Bargiela-Chiappini, editors, *Face, communication and social interaction*, chapter 12, pages 227–249. Equinox London.

David Traum, Andrew Jones, Kia Hays, Heather Maio, Oleg Alexander, Ron Artstein, Paul Debevec, Alesia Gainer, Kallirroi Georgila, Kathleen Haase, and 1 others. 2015. New dimensions in testimony: Digitally preserving a holocaust survivor's interactive storytelling. In *International Conference on Interactive Digital Storytelling*, pages 269–281. Springer.

Ning Wang, W Lewis Johnson, Richard E Mayer, Paola Rizzo, Erin Shaw, and Heather Collins. 2008. The politeness effect: Pedagogical agents and learning outcomes. *International journal of human-computer studies*, 66(2):98–112.

Frederick White. 2006. Rethinking native american language revitalization. *American Indian Quarterly*, 30(1/2):91–109.