

Developing a spoken dialogue system for the Choctaw language

Jacqueline Brixey and David Traum

Abstract We present the first spoken dialogue system for the Choctaw language in this paper. Choctaw is an endangered American indigenous language spoken by the Choctaw tribe. Previous work in this area created a text-based English-Choctaw bilingual chatbot, named Masheli, that primarily shared stories about animals. Additional work developed an automatic speech recognizer (ASR) to process spoken Choctaw. In this paper, we demo the Choctaw ASR together with the Masheli chatbot to form a dialogue system that allows the user to speak, rather than type, to the system. As the language is endangered, a spoken dialogue system would assist revitalization efforts by promoting oral fluency in language learners.

1 Motivation

Many world languages face the threat of becoming extinct, and with each language that dies, we lose unique culture, heritage, and history. However, there is currently a greater emphasis on developing technological resources for low resource and indigenous languages, for example UNESCO's International Decade of Indigenous Languages¹. In our work, we implement a dialogue system that recognizes spoken Choctaw, an American indigenous language. In previous work we developed a text-based form of this dialogue system, named Masheli (meaning "fair sky" in Choctaw).

Dialogue systems for endangered languages can fill an important role as a surrogate conversation partner. For one, while traditional classroom settings may attempt to create conversational opportunities, factors, such as feelings of shyness or fear of making errors [25], can prevent learners from engaging fully in conversation with a human partner. Learners reported feeling more comfortable chatting with a dialogue

USC Institute for Creative Technologies
e-mail: [brixey,traum]@usc.ict.edu

¹ <https://en.unesco.org/idi12022-2032>

system than with a human interlocutor in previous research [12]. In addition, access to fluent speakers may be limited, or even nonexistent, for endangered languages. A dialogue system could be a surrogate conversational partner.

The Choctaw people have ongoing efforts to revitalize and reclaim the language including language courses at local schools and online classes². Developing a spo-