Singing Synthesizers: Musical Language Revitalization through UTAUloid

Morgan Sleeper
Macalester College

Abstract

Music plays many important roles in language revitalization, from attracting learners and fostering speech communities to supporting language learning. These effects, however, are largely independent from the skills which linguists bring to language revitalization. This study introduces one concrete way in which applied linguistics can directly support musical language revitalization with UTAUloids – speech-and-music software synthesizers – illustrated through the creation of a Cherokee UTAUloid as part of ancestral language reclamation by a learner-linguist Cherokee Nation citizen.

Through their focus on "massive collaboration," low-resource music production, and youth involvement, UTAUloids are uniquely situated to serve as instruments for language revitalization. Even the act of creating an UTAUloid itself allows speakers and learners who may not consider themselves "musical" to contribute to musical language revitalization, and this study provides a step-by-step methodology to make creating an UTAUloid as accessible as possible for anyone interested in incorporating music into their own language revitalization practice.

Résumé

La musique joue un rôle important pour la revitalisation des langues : attirer des apprenants, créer des communautés de locuteurs et soutenir l'apprentissage des langues. Or, ces effets ne font normalement pas partie des compétences que les linguistes apportent à la revitalisation des langues. Cette étude présente une façon dont la linguistique appliquée peut aider la revitalisation musicale des langues avec UTAUloids – des synthétiseurs de parole et de musique – à travers la création d'un UTAUloid Cherokee, un projet de récupération de la langue ancestrale par un apprenant-linguiste et citoyen de la Cherokee Nation.

Mettant l'accent sur la « collaboration massive, » la production de musique à faibles coûts et l'implication des jeunes, les UTAUloids sont particulièrement bien placés pour servir d'instruments de la revitalisation des langues. Même le simple fait de créer un UTAUloid permet aux locuteurs et aux apprenants qui ne se considèrent pas comme des « musiciens » de contribuer à la revitalisation musicale des langues. Cette étude propose une méthodologie pas à pas pour rendre la création d'un UTAUloid aussi accessible que possible à toute personne souhaitant intégrer la musique dans sa propre pratique de revitalisation des langues.

Singing Synthesizers: Musical Language Revitalization through UTAUloid

Introduction

Language revitalization is inherently interdisciplinary, and one intersection that has seen a growing amount of scholarly and community attention in recent years is the confluence of language revitalization and music. Music can play many roles in language revitalization, from helping learners progress in terms of grammar, pronunciation, and vocabulary (Bracknell et al., 2021; Tuttle, 2012; Vallejo, 2019), to reinforcing identity (Barrett, 2016; Dołowy-Rybińska, 2020; Johnson, 2012; Llewellyn, 2000; Lucas, 2021; Sparling et al., 2022), expanding a language's domains of use (Bracknell et al., 2021; Cotter, 2001; Cru, 2018; Lucas, 2021; Sometimes & Kelly, 2010), bringing speakers and learners together (Ashton, 2020; Bracknell, 2020; Nummelin, 2020), and infusing language learning with a powerful sense of joy (Przybylski, 2018; Sparling et al., 2022; Vallejo, 2019).

Notably, however, the labour which enables these positive effects of music is largely independent from the specific skillset linguists bring to language revitalization, and music has thus far played little role in the productive relationship between applied linguistics and Indigenous language revitalization (outlined in, e.g., Daniels & Sterzuk, 2022; McIvor, 2020). To that end, this study introduces one concrete way in which these two disciplines can come together to support musical language revival, through UTAUloids: free and open-source speech-and-music software synthesizers used for collaborative vocal songwriting. This study explores their potential for language revitalization through the creation of an UTAUloid in Cherokee – a revitalizing Iroquoian language of Oklahoma, the US Southeast, and the Cherokee people dispersed around the United States and beyond in a digitally connected "at-large" community – as part of ancestral language reclamation by a learner-linguist Cherokee Nation citizen.

The paper begins with an overview of the Cherokee language and revitalization, before exploring the place of music in language revitalization efforts more broadly. It then introduces UTAUloids (along with their inspiration, Vocaloids) before illustrating a step-by-step methodology for creating an UTAUloid using Cherokee as an example, followed by a discussion of UTAUloids in language revitalization and the ways in which these unique musical tools can expand how we conceptualize the relationship between applied linguistics and Indigenous language revitalization.

Cherokee

Cherokee (ᏣᎳᎩ *tsalagi*) is an Iroquoian language, part of a family that also includes Mohawk, Wendat/Wyandot, Seneca, Tuscarora, and Cayuga (Mithun, 1999). As of the 2010 census, Cherokee counted 12,300 speakers, including approximately 10,000 in and around the Cherokee Nation in Oklahoma, approximately 1,000 in North Carolina (where the Eastern Band of Cherokee Indians are located), and an undetermined number of members of the United Keetoowah Band of Oklahoma and Arkansas (Golla, 2010). The present-day geographical distribution of Cherokee speakers is a result of a prolonged campaign of ethnic cleansing by the US government, and particularly of the 1838 forced removal of the Cherokee people from their homelands in the US Southeast to eastern Oklahoma, commonly known as the Trail of Tears. The majority of surviving Cherokee

settled in Oklahoma, in what is now the Cherokee Nation reservation, while groups who escaped the initial removal and took refuge in the Appalachian Mountains eventually settled in western North Carolina, the territory of today's Eastern Band of Cherokee Indians. Beyond these two communities, there is a sizable diasporic population of Cherokee people across the United States; out of the Cherokee Nation's 450,000 citizens, around 180,000 reside outside of Oklahoma (ᎠᎾᎫᏍᎪᎢ/Anadisgoi, 2023), with significant populations in California, Washington State, Texas, Kansas, and Florida.

Typologically, Cherokee is a polysynthetic language and distinguishes itself from other Iroquoian languages as the only variety in the family with lexical tone. Culturally, the Cherokee language is strongly associated with its syllabic orthography (Cushman, 2012), developed by ᏍᎪᏲᏙᎲ (Sequoyah) in the 1820s. The syllabary consists of 85 symbols, each representing a CV or V syllable of Cherokee; the sole exception is <ᏍᎠ>, which represents the single segment /s/. The syllabary is shown with romanized orthographic equivalents[1] in Table 1:

Table 1
Cherokee Syllabary

Ꭰ 'a'		Ꭱ 'e'	Ꭲ 'i'	Ꭳ 'o'	Ꭴ 'u'	ᵢ 'v'
Ꭶ 'ga'	Ꭷ 'ka'	Ꭸ 'ge'	Ꭹ 'gi'	Ꭺ 'go'	Ꭻ 'gu'	Ꭼ 'gv'
Ꭽ 'ha'		Ꭾ 'he'	Ꭿ 'hi'	Ꮀ 'ho'	Ꮁ 'hu'	Ꮂ 'hv'
Ꮃ 'la'		Ꮄ 'le'	Ꮅ 'li'	Ꮆ 'lo'	Ꮇ 'lu'	Ꮈ 'lv'
Ꮉ 'ma'		Ꮊ 'me'	Ꮋ 'mi'	Ꮌ 'mo'	Ꮍ 'mu'	
Ꮎ 'na'	Ꮏ 'hna'	Ꮐ 'nah'	Ꮑ 'ne'	Ꮒ 'ni'	Ꮓ 'no'	Ꮔ 'nu'	Ꮕ 'nv'
Ꮖ 'qua'		Ꮗ 'que'	Ꮘ 'qui'	Ꮙ 'quo'	Ꮚ 'quu'	Ꮛ 'quv'
Ꮜ 's'	Ꮝ 'sa'	Ꮞ 'se'	Ꮟ 'si'	Ꮠ 'so'	Ꮡ 'su'	Ꮢ 'sv'
Ꮣ 'da'	Ꮤ 'ta'	Ꮥ 'de'	Ꮦ 'te'	Ꮧ 'di'	Ꮨ 'ti'	Ꮩ 'do'	Ꮪ 'du'	Ꮫ 'dv'
Ꮬ 'dla'	Ꮭ 'tla'	Ꮮ 'tle'	Ꮯ 'tli'	Ꮰ 'tlo'	Ꮱ 'tlu'	Ꮲ 'tlv'
Ꮳ 'tsa'		Ꮴ 'tse'	Ꮵ 'tsi'	Ꮶ 'tso'	Ꮷ 'tsu'	Ꮸ 'tsv'
Ꮹ 'wa'		Ꮺ 'we'	Ꮻ 'wi'	Ꮼ 'wo'	Ꮽ 'wu'	Ꮾ 'wv'
Ꮿ 'ya'		Ᏸ 'ye'	Ᏹ 'yi'	Ᏺ 'yo'	Ᏻ 'yu'	Ᏼ 'yv'

The syllabary was quickly adopted and remains widely used among speakers and learners today; it has also been incorporated into Unicode since 1999 and is available as a default input method on Apple's macOS and iOS and Microsoft's Windows operating systems.

In terms of vitality, Cherokee is considered "definitely endangered" by UNESCO (Moseley, 2010), and though usage by speakers under 40 is low, with most children no longer learning it as a home language (Cherokee Nation, 2003, as cited in Uchihara, 2016) the Cherokee people and the Cherokee Nation are committed to revitalizing the language. Revitalization efforts for Cherokee are multifaceted and wide-reaching, including the

popular ᏣᎳᎩ ᏧᎾᏕᎶᏆᏍᏗ (*Tsalagi Tsunadeloquasdi)* immersion schooling from pre-school through sixth grade (Peter et al., 2017), master/apprentice programs, and community and university language classes. Media also plays an important role, with radio broadcasts from the Cherokee Nation and the monthly bilingual Cherokee Phoenix newspaper providing news and features in the language.

The Cherokee Nation has in recent years also been at the forefront of innovative, Indigenous approaches which conceptualize language revitalization expansively, in terms of the lifeways of current and future speakers. The Speaker Services program launched in 2022, for instance, aims to help Cherokee speakers with everyday basic needs, including home repairs, accessing healthcare and medical devices, and installing new appliances and accessibility aids in speakers' homes. This program is housed in the Cherokee Nation's Language Department on the belief that "if speakers aren't worried about their roofs leaking, they can worry about their grandchildren speaking Cherokee" (Cherokee Phoenix Staff, 2022). Other recent expansive language revitalization efforts have included installing a state-of-the-art cellular tower in Kenwood – a remote reservation community of around 1,000 people with a high proportion of fluent, daily speakers – so that young people can take advantage of remote work to stay in the community, Cherokee speakers can more readily use the language with friends and family across the Nation, and online language classes can be taught from the town (Caldwell, 2023).

Another notable aspect of Cherokee language revitalization in the Cherokee Nation is that there has been a focus on reaching heritage speakers and tribal members who are physically removed from the reservation and its speech community. The Cherokee Nation offers several online language classes (at three skill levels) taught by a native speaker throughout the year, for instance, as well as self-paced pre-recorded classes, and the option for remote learners to call in to talk with native speakers. A monthly livestream series called "ᏣᎳᎩ: Wherever we are" was also started in 2021 to connect diasporic Cherokee Nation citizens with cultural programming on topics like Cherokee spirituality, music, games, and history, alongside updates on important current tribal issues, including language revitalization. In addition, the Cherokee Nation organizes annual in-person outreach events to more than 20 "At-Large" communities with significant populations of tribal members across the United States, from the Puget Sound and California to Texas, Kansas, and Florida. Along with voter registration, healthcare information, games, storytelling, and dance, these annual events also prominently include Cherokee language resources, materials, and taster classes intended to garner interest and participation in language preservation from at-large members.

It is within this wider at-large community that I situate both myself and my relationship to the Cherokee language. A citizen of the Cherokee Nation, I grew up in Florida, removed from the reservation but continually hearing the importance of our language, stories, and culture from my father and grandfather, spending time in the Cherokee Nation while visiting family over summers, and later attending At-Large gatherings when living in California and Washington State. Neither my father nor my grandfather spoke Cherokee fluently, but they passed down individual words and phrases – ᎣᏏᏲ, ᏩᏙ, ᎣᏍᏓ, ᏙᎾᏓᎪᎲᎢ – and a reverence for the language that inspired me to try to learn, and to pass on words to my son in turn. I have been slowly working towards reclaiming Cherokee over the past 10 years, through online classes and At-Large meetups. Now living in Minnesota, the At-Large gatherings are less accessible, and I find myself

relying on digital means to stay connected to the language and the Cherokee Nation, including the "ᏩᎳᏯ: Wherever we are" programming, social media, and online language courses. I am also a lifelong musician – a fact which guides my work as a linguist, and shapes how I interact with language learning and reclamation; I love to learn through song, and I have been particularly inspired by other Native people learning their languages through songwriting (e.g., Przybylski, 2018; Tuttle & Lundström, 2015). This current project sits at the confluence of these aspects of my identity, and of my relationships with music, linguistics, and the Cherokee language – UTAUloid gives me a way to use music in my personal language reclamation, while also leveraging my linguistic skills to help with wider Cherokee language revitalization efforts, within the context of our digitally-connected diasporic Cherokee community.

Music in Language Revitalization

This section will provide a brief overview of some of the many ways music can aid in language revitalization efforts, focusing on three specific areas: attracting learners and retaining speakers, creating opportunities for speech community, and supporting language learning.

Attracting learners and engaging interest in the language movement among speakers is a particularly important contribution music can make in language revitalization. The Cherokee Nation's Cherokee National Youth Choir, for instance, sings an all-Cherokee-language repertoire specifically to both introduce young learners to the language, and to keep immersion school students engaged with the language movement outside of classes (Cherokee Nation, 2022), while the Oklahoma Native American Youth Language Fair (ONAYLF) offers a venue for young learners and speakers to share both traditional and newly composed music in Native languages. For Cornish – a revitalizing language spoken in Cornwall – music is routinely cited as one of the main ways new learners first encounter the language, from a folk music revival in the 1970s (MacKinnon, 2005) to the modern electronic-pop of Gwenno Saunders, who the Cornish Language Board directly credits for the record number of learners registering for Cornish exams after her 2018 album *Le Kov* (BBC News, 2018). Music can also help spread awareness of both revitalizing languages and Indigenous language revitalization in general to the wider public, as when a translated cover of The Beatles' *Blackbird* in Mi'kmaq by Cape Breton teenager Emma Stevens went viral in 2019. The video was uploaded to mark the United Nation's International Year of Indigenous Languages and has seen over 1.8 million views on YouTube, introducing viewers to the UN's campaign for language revitalization while also "[showing] non-Mi'kmaw people the beauty of our language" (Goodyear, 2019).

Another important function of music in language revitalization is that it can serve to change people's attitudes about language, with important benefits for language use. In Aotearoa New Zealand, for instance, the 1982 song *Poi E* – a fusion of Māori culture and hip-hop music co-written in Te Reo Māori by linguist Ngoingoi Pēwhairangi and Maui Dalvanius Prime – was conceptualized as a way to inspire young Māori listeners to use Te Reo in their everyday lives (Archer, 2002). The song was a breakout hit and instrumental part of the Māori language and cultural revival that followed, and helped shape the image of Māori as a modern, vitally relevant language for a new generation (Sheehan, 2016, p. 78). Elsewhere, in Minnesota, Anishinaabe artist Tall Paul similarly uses bilingual hip-hop

music to change attitudes about Anishinaabemowin and inspire more young people to take up the language, saying specifically: "If I incorporate the language into hip hop, it'll make the language cool for those kids. Maybe they'll be interested in learning it at a young age" (Przybylski, 2018, p. 388).

Music can also be instrumental in language revitalization by creating opportunities for speech community. In many minoritized language contexts, a prevailing issue is that another language (or set of languages) has become the medium of everyday communication; even with a sizable number of fluent speakers, this can make it difficult for speakers to find natural domains of use for the language in daily life (Hinton, 2001). Music can help provide this space – both physical and conceptual – for speakers to come together in speech community.

MacKinnon (2005, p. 249), for instance, notes how Cornish music festivals are particularly important as "opportunities for Cornish speakers and learners to come together and use the language", and concerts, festivals, and participatory musical gatherings fulfill this purposes in revitalizing language communities around the globe, from the *Eisteddfodau* festivals in Wales to Breton *Fest-Noz* ("night-festivals") (Dołowy-Rybińska, 2020), the multi-Nation sākihiwē indigenous music festival in Winnipeg (Przybylski, 2021), Guernésiais choir rehearsals on Guernsey (Johnson, 2012), and the concerts of the Ainu band Marewrew. Marewrew perform entirely in Ainu – a critically endangered language indigenous to Hokkaido – and explicitly teach the language behind their repertoire as part of their performances, explaining translations and linguistic concepts while teaching the audience to sing along. This approach creates a space where "the audience become active participants in a performance that uses the Ainu language… almost delivered as mini-workshops" (Nummelin, 2020, p. 291).

In addition to physical gatherings of speech community, music can also help to link speech community through distributed media, such as radio and podcasts. *Radyo an Gernewegva* (literally "radio of the Cornish-language area"), for instance, is a Cornish-language, podcast-format radio show produced weekly since 2007, with the goals of providing Cornish immersion, showcasing Cornish-language music, culture, and talent, and giving Cornish musicians exposure. Its listeners are often language learners, and an active, online community forum encourages participation and engagement with the music in Cornish, even when participants are geographically isolated from other speakers. Traditional over-the-air radio in revitalizing languages fulfills a similar function, and is particularly effective at "strengthening, sustaining, and revitalizing cultural and linguistic traditions" (Danos & Turin, 2021, p. 76), by creating asynchronous "speech communities" that can be tuned into even as listeners go about their daily lives.

Finally, perhaps one of the most immediately obvious ways in which music can contribute to language revitalization is in supporting language learning. Of course, this benefit is not restricted to revitalizing languages, and much of the research in this area is based on commonly taught languages (Davis, 2017; Engh, 2013; Good et al., 2015; Tegge, 2018); but the central place of language teaching and learning in revitalization makes this affordance of music especially relevant here.

A large portion of previous research on music in language teaching focuses on the use of songs in classroom pedagogy and describes how music can be particularly helpful for teaching natural pronunciation, new grammatical structures, and idioms (Jolly, 1975); capturing students' attention and contextualizing colloquial uses of language (Abrate, 1983); and introducing vocabulary domains and language issues not commonly brought up

in textbooks (Schmidt, 2003). More recently, Vallejo (2019) explores how teachers in Kanien'ke:ha (Mohawk) language immersion programs use both traditional and contemporary songs to provide wholistic, culturally-grounded language education, with music as a "linchpin pedagogical tool that promotes intergenerational interactions, builds social relationships, and facilitates the daily use of language in and outside the classroom." Notably, the musical repertoire used includes translations of Western nursery rhymes, popular hits from Elvis and Johnny Cash, and Christmas carols alongside traditional Kanien'ke:ha music, and while these may on the surface seem incompatible with the idea of a culturally-grounded Kanien'ke:ha language education, Vallejo (2019) argues that English-language songs essentially "take on a new form" when translated into Kanien'ke:ha (p. 106), and that this mix is both pedagogically beneficial and reflective of the broad musical tastes and experiences of the Kanien'ke:ha community.

Of course, many revitalizing languages are not taught in classroom settings, and music can help support language learning in these situations as well, especially in terms of indigenous methodologies. Antoine (2015) shows how Lakota songs can teach language and culture "from the Native perspective, reinforce culturally appropriate ways of behavior, and teach the tribe's social structure as well as history and spirituality" (p. 17). Tuttle and Lundström (2015) present a case study of a young composer learning potlatch singing in three endangered Interior Athabaskan languages of Alaska directly from elders and point specifically to how "gaining proficiency in song, through Indigenous channels, can further proficiency in language" (p. 38)

One important thread running through many of these musical approaches is that they are not exclusively – or even primarily – concerned with a narrow view of language revitalization as "increasing speaker numbers." And in fact, many of these musical practices might be more accurately thought of as forms of language reclamation, defined by linguist Wesley Leonard as "a larger effort by a community to claim its right to speak a language and set associated goals in response to community needs and perspectives" (Leonard, 2012, p. 359). Rather than revitalizing a "language" in the abstract, language reclamation is a process of "personal and communal agency and the expression of Indigenous identities, belonging, and responsibility to self and community" (McCarty et al., 2018, p. 160), which includes claiming and creating new practices around language and music and (re-)defining culture (Leonard, 2012).

Each of the three areas above, then – attracting and engaging learners and speakers, fostering speech communities, and supporting language learning – are important ways in which language revitalization and reclamation can happen through music. Notably, however, the necessary work which enables this – composing, producing, and performing music, organizing concerts and festivals, documenting song repertoires – is not something that linguists are trained to help with. To that end, the remainder of this study introduces a practical overlap between applied linguistics and Indigenous language revitalization in which these two fields can collaboratively create a powerful tool for musical language revitalization: vocal music production tools known as UTAUloids.

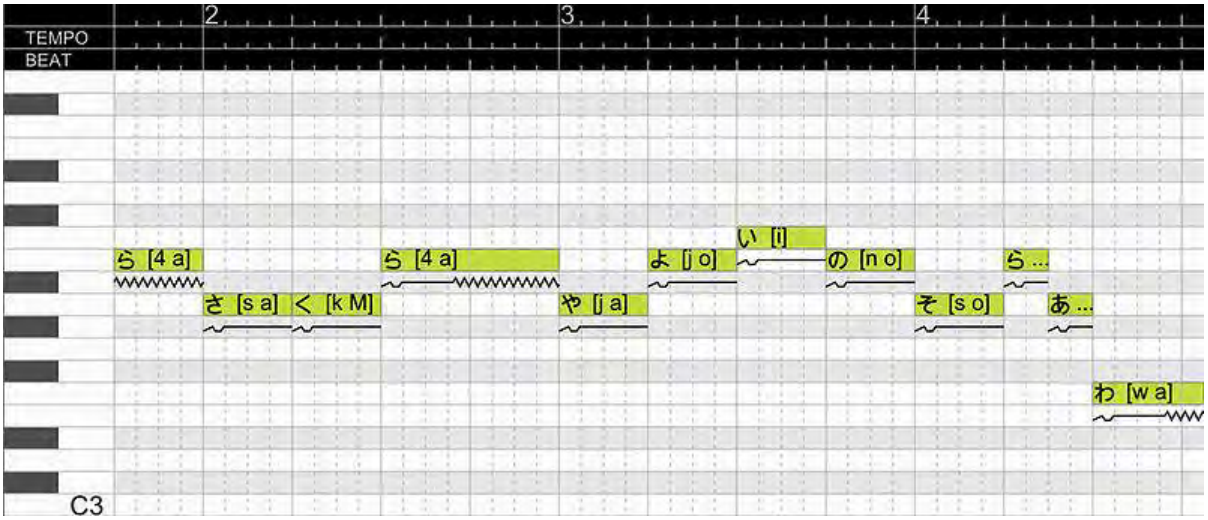
Vocaloid

To fully explain UTAUloids, it is first necessary to briefly introduce their inspiration, Vocaloids. Often described as "anthropomorphized singing synthesizers," Vocaloids were developed by Yamaha and the Music Technology Group at Universitari

Pompeu Fabra, Barcelona in 2004 (Kenmochi & Ohshita, 2007). Vocaloid software uses a library of human speech samples (called a “voicebank”) and combines them with musical information in order to output digitally “sung” melodies. Voicebanks are created by recording a voice donor (usually a singer or voice actor/actress) singing all possible syllables (or morae) in a language at several different pitches, so that each individual sample can be arranged and manipulated by the Vocaloid synthesizer to approximate a singing voice.[2] The first Vocaloids were developed to sing in English and Japanese, but subsequent releases have expanded to include languages like Mandarin, Korean, and Spanish.

Users of the Vocaloid software input both notes and syllables using a piano roll-style score editing interface shown in Figure 1, which features an excerpt from the Japanese folk song *Sakura, Sakura*:

Figure 1
Vocaloid Score Editor (Clusternote, 2014)

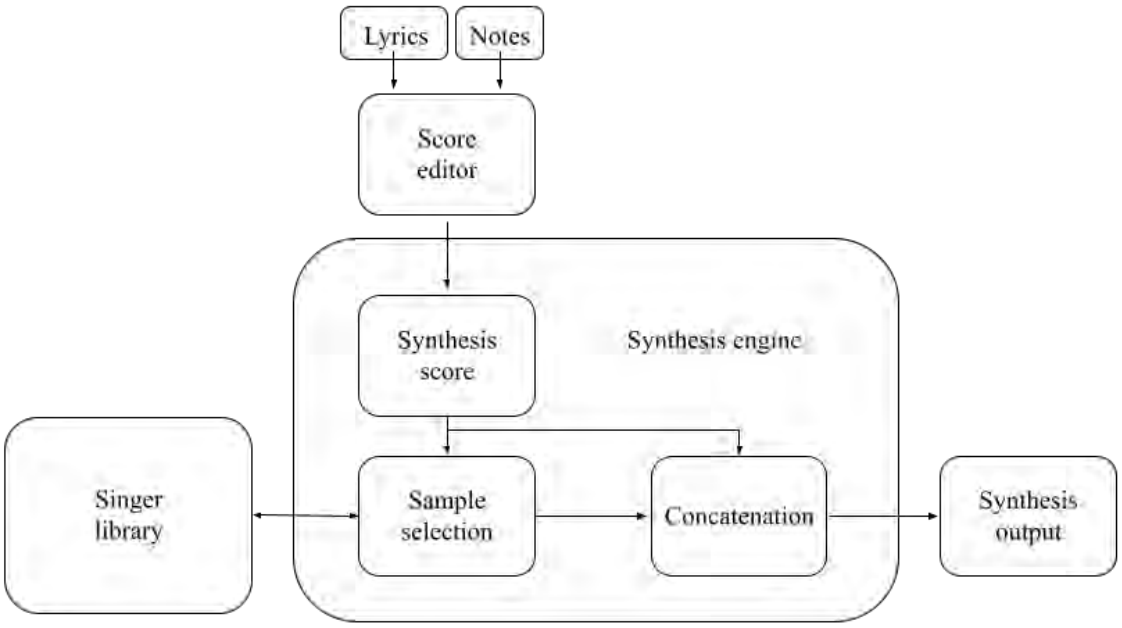


The Vocaloid software then takes the resulting score as input into its synthesis engine, which chooses appropriate samples from its voicebank (i.e. the syllables to be sung, at the nearest pitch to the target note) and concatenates and alters these samples based on musical information in the score (including pitch, dynamics, timbre, attack, decay, reverb, and vibrato) to produce the synthesized output: a sung melody. This synthesis engine is represented graphically in Figure 2, adapted from Kenmochi and Ohshita (2007).

This is the basic synthesis behind the sounds the Vocaloid software produces, but an equally important component of the Vocaloid concept is that the individual packages are anthropomorphized; they are not just 'singing synthesizers', but “synthesized singers.” There have been dozens of different Vocaloids introduced since the technology's debut in 2004, but by far the most popular has been Hatsune Miku, designed by Crypton Future Media in Japan.

Hatsune Miku (初音ミク, meaning “first sound of the future”) made her debut in 2007, and is stylized as a “virtual idol,” with long turquoise twintails and design features on her costume which recall the user interface of her software.

Figure 2
Vocaloid System Diagram (adapted from Kenmochi & Ohshita, 2007)



She has proved enormously successful; over 100,000 songs have been created using her voicebank, and she appears in over 170,000 YouTube videos, dozens of which have more than a million views each[3] (Crypton Future Media, n.d.). She has starred in video games and car commercials, topped album charts in Japan, and performed (via hologram) as an opener for Lady Gaga and in her own headlining sold-out concerts around the world. Her rise to become something of an international icon illustrates how the creative culture which has arisen around Vocaloids makes them much more than just instruments for the producers, musicians, and fans who use them.

One of the key pieces of Vocaloid's appeal is the fact that Vocaloid encourages iterative creation by multiple users over the internet. The specific combination of format (exportable so that users can download and edit notes, lyrics, and other parameters on existing tracks) and licensing (which in general allows for the free personal and commercial use of Vocaloids' voices and likenesses in videos, art, and other adaptations) has made Vocaloid especially conducive to what has been called "massive collaboration" (Sousa, 2014).

This "massive collaboration" means that users can interact with and iterate on Vocaloid content in a myriad of ways. One person might upload a melody composed in the software; a second person could then download it, add drum and bass tracks behind it, and re-upload it. A third person could remix the music, change the lyrics, and then re-upload the result. A fourth person could record themselves singing the tune over the instrumental backing tracks, and a fifth person could create a video compiling different interpretations of the song.

This phenomenon is a central component of Vocaloid culture, and one result is that the boundary between "fan" and "creator" in Vocaloid is substantially blurred, if not altogether erased. Rather than an audience passively consuming musical texts, Vocaloid users are "a distributed group of fan-producers" (Condry, 2011), and Vocaloid has become

"a catalyst for collective, grassroots, and multidisciplinary creation" (Sousa, 2014) spanning music, lyrics, art, animation, costume design, choreography, writing, and more.

The potential for 'massive collaboration' is a large part of Vocaloid's appeal to its users and fans, along with the inherent creative control it provides; Vocaloid producers can fine-tune almost every part of a performance, from backdrops and costumes to vocal delivery, intonation, and microtiming. The only limitation to what can be customized in a Vocaloid performance comes from the voicebank itself – the voice donor's initial recordings. This is where UTAUloid comes in.

UTAUloid

UTAUloid[4] (From 歌う *utau* meaning "to sing" in Japanese), created by developer Ameya/Ayame in 2008, is a freeware implementation of the Vocaloid concept. Like Vocaloid, UTAUloid allows users to key in scores with notes and lyrics, which it in turn relays to a synthesis engine to produce synthesized singing output. The main difference between Vocaloid and UTAUloid is that while Vocaloid products come with a pre-recorded, non-modifiable voicebank, UTAUloid instead allows users to record and use their own voicebanks, which then act as a sample library for the synthesis. Importantly, this means that not only can UTAUloids sing in any voice, they can also sing in any language.

In addition to UTAUloids who sing in Japanese and English, users have created UTAUloids singing in many languages not available in any Vocaloid, including Catalan, Tagalog, Indonesian, Basque, Irish, and Esperanto. This is particularly impressive because creating an UTAUloid requires a significant amount of specialized linguistic knowledge about a given language, and this is reflected in the linguistic proficiency of the wider community. Fans on the UTAU Wikia, UTAU Wiki, and UtaForum.net websites create linguistic tutorials, curate threads on relevant phonetic and phonological issues for specific languages, recommend scholarly linguistic work as references, and synthesize academic articles into advice for creators to use when recording their "reclists" of potential phonemic combinations.

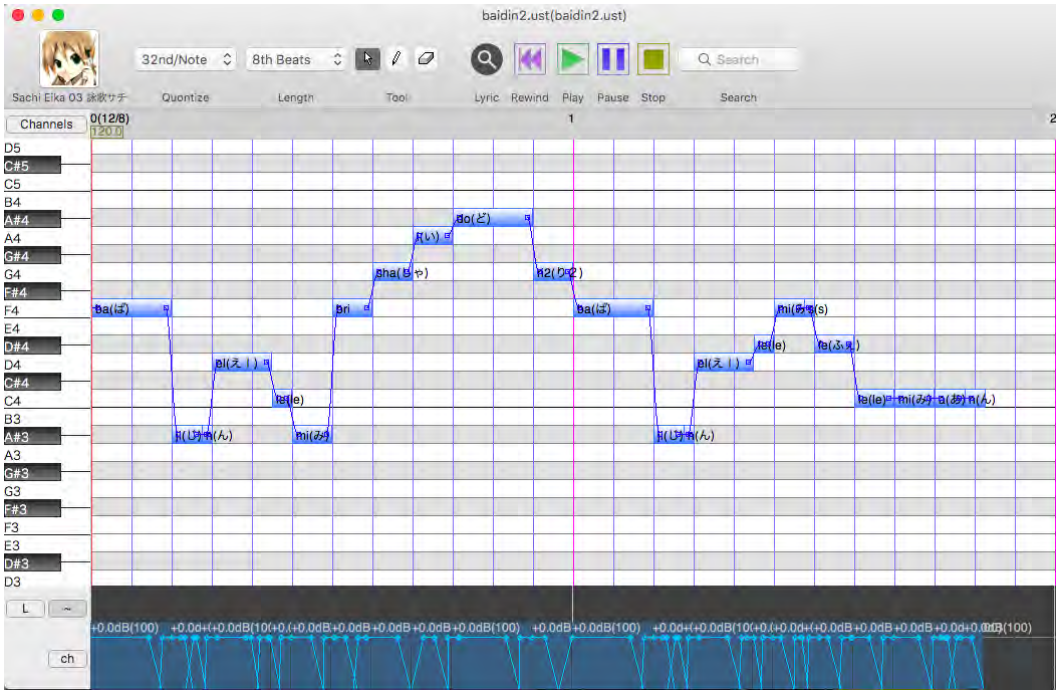
Along with this community linguistic knowledge, it is notable that the linguistic plurality of UTAUloids available includes several who sing in minoritized, marginalized, and revitalizing languages. One prominent example is Sachi Eika (詠歌サチ), an Irish-language UTAUloid created by user *Jadii* in 2009. Figure 3 shows Sachi Eika's design, in both the key art and a 3D model for use in music videos.

As with Vocaloids, UTAUloids like Sachi Eika exist within the framework of massive collaboration and collective creation (Le, 2014); once made available online, UTAUloid fans were able to write and produce music with Sachi Eika, re-record existing Vocaloid and UTAUloid songs with her voice, download and modify others' Sachi Eika songs, and create art, videos, and other examples of "Nth fanfiction" (Kenmochi, 2010) featuring the character. As of 2023, YouTube hosts over 140 videos tagged for Sachi Eika; the fan-art community website DeviantArt listed over 100 works of the character before its recent restructuring, and similar Japanese site Pixiv still hosts several dozen; Sachi Eika also appeared as a character in a printed Irish manga magazine in 2010 (UTAU Wikia contributors, n.d.).

Figure 3
Sachi Eika (Jadii, 2009)

Additionally, while most of these fan works were created in the years directly following Sachi Eika's debut, the nature of UTAUloid means that even over a decade later, it is still possible to load Sachi Eika's voicebank onto any computer and have her sing, within a matter of minutes, an Irish-language song, such as *Báidín Fheilimí,* shown in Figure 4 below.

Figure 4
Sachi Eika UTAUloid Score for Báidín Fheilimí



It is in this vein that UTAUloids can be used as literal and figurative instruments for language revitalization, and the remainder of this article provides a step-by-step guide for producing an UTAUloid for revitalization, illustrated through the creation of an example Cherokee UTAUloid called ᎧᏃᎩᏍᏗ (*kanogisdi* "singing").

Method: Building a Cherokee UTAUloid

This section will demonstrate the procedure for creating an UTAUloid, illustrating the process through Cherokee. An UTAUloid can be broadly conceptualized into three parts: the synthesis engine, the sample library ("voicebank"), and the set of instructions (called "tunings") unique to each UTAUloid that specify how the synthesis engine should interpret, blend, and concatenate individual samples into synthesized singing. The three main components of creating an UTAUloid, then, are to install the synthesis engine, to record a voicebank of samples, and to tune those samples for synthesis; and these final two steps can be directly conceptualized as applied linguistic work.

The UTAU synthesis engine is available in several different software packages, including the original UTAU application for Windows,[5] UTAU-Synth for macOS,[6] and more recently, the open-source project OpenUTAU,[7] compatible with both Windows and macOS. Voicebanks and files are fully cross-compatible between the different applications, and the user interfaces are also largely the same. The Cherokee UTAUloid and examples shown here were created in OpenUTAU, which is recommended for several reasons: it is in active development by an engaged, helpful, and welcoming community; its interface is fully localized into 17 languages; and its open-source license and design allows for all the benefits of using open tools in Indigenous language revitalization, including accessibility, longevity, and community customization (Berez-Kroeker et al., 2023; Brinklow et al., 2019; Salazar et al., 2021).

Recording a Voicebank

Recording a voicebank of samples from the UTAUloid's voice donor is a process that requires the application of a certain amount of linguistic knowledge about the target UTAUloid language, and closely resembles collecting wordlist data for phonetics research.

The first step is to assemble the list of sounds to be recorded, into what the UTAUloid community calls a "reclist." The reclist should consist of all possible combinations of phonemes needed to reproduce any syllable (or mora) in the target language. This requires a linguistic understanding of the phonetic inventory, phonotactics, and syllabic (or moraic) structure of the language, as well as an understanding of how UTAUloid synthesis works in order to know exactly what phoneme combinations need to be recorded.

The phonetic inventory of Cherokee (Montgomery-Anderson, 2015; Pulte & Feeling, 1975; Uchihara, 2016) is presented in Tables 2 and 3 below, based on Montgomery-Anderson (2015) and Uchihara (2016):

Table 2
Cherokee Consonant Inventory

		Bilabial	Alveolar	Palatal	Velar	Labio-velar	Glottal
Plosives			t		k	k^w	ʔ
Affricates	Central		ʧ				
	Lateral		tɬ				
Fricatives			s				h
Nasals		m	n				
Approximants	Central			j		w	
	Lateral		l				

Table 3
Cherokee Vowel Inventory

	Front		Central		Back	
High	i	i:			u	u:
Mid	e	e:	ə̃	ə̃:	o	o:
Low					a	a:

From this inventory, the next step is to determine how the phonemes pattern in terms of syllables (or moras), which serve as the unit of recording for UTAUloid. In Cherokee, the minimal syllable consists of a vowel, with optional onsets of single consonants or more complex clusters, and an optional single consonant coda, represented in Table 4 (adapted from Uchihara, 2016) below:

Table 4
Cherokee Syllable Structure

Onset			Nucleus	Coda
(s)	(t ʧ tɬ l k k^w)	(h)	V(V)	(l n j w ʔ h)
	(n j w)			
	(s m h ʔ)			

Even with Cherokee's relatively small phonetic inventory, this maximal syllable structure would make recording each possible potential syllable for sampling a considerable task. Thankfully, the architecture of UTAUloid allows for combining samples on a single note, so that a smaller number of recorded samples can approximate the entire syllabic inventory of a language. CVC syllables, for instance, can be produced by combining two appropriate CV and VC segments, and in languages with phonemic vowel length (such as Cherokee), it is not necessary to record separate samples for V and V: (or CV and CV:) syllables; vowels can be lengthened either by note length or with the addition of a

following sample of the same vowel quality. Likewise, complex onsets and codas can be 'cropped' via volume manipulation in the UTAU score, so that a C_aC_bV segment could also be used to render a C_bV syllable, and a C_xV syllable and C_yV syllable could be manipulated to form the complex onset of a C_xC_yV syllable.

Of course, because of the phonetic effects of co-articulation and gestural timing, the more an UTAUloid voicebank relies on 'shortcuts' such as syllable cropping, the less natural it will sound; in general, the more unique segments are recorded, the more ‘natural’ the UTAUloid will sound. It is worth noting, however, that “natural” is not necessarily the goal of UTAUloid, and many popular UTAUloids and Vocaloids are made without taking every phonetic process of a language into account. Hatsune Miku, for instance, does not produce the vowel devoicing between voiceless consonants which is saliently characteristic of much spoken Japanese, and deviation from “natural” phonetic production is often seen as a positive creative affordance of UTAUloids and Vocaloids.

Another consideration is that, particularly in Indigenous language revitalization contexts, it may be important to record samples that are not strictly necessary for forming words from the language’s phonemic inventory but are important for other reasons. In Cherokee, for instance, the syllables /hna/ and /nah/ would not be part of a minimal reclist, since they can be formed by combining other CV and VC samples (i.e. /na/ and /ah/). Both of these syllables are, however, also independent characters in the Cherokee syllabary (Ꮒ 'hna' and Ꮀ 'nah'), and because of the cultural importance of the syllabary, it was a goal of this UTAU project to include each character as a standalone sample, and so they were recorded here as well.

Taking the above factors into account, the reclist for the Cherokee UTAUloid consists of the following segments:

Table 5

Segments for Cherokee Reclist, by Syllable Type

	Category	Examples	Number
1	CV: all initial consonants + all vowels	/tɬo/ /se/ /ʔa/	78
2	CCV: /s/ + {/k kʷ t ʧ tɬ l/} + all vowels	/skʷa/ /stə̃/ /stɬi/	36
3	CCV: {/k kʷ t ʧ tɬ l n j w/} + /h/ + all vowels	/khi/ /tho/ /nhe/	54
4	V: all vowels	/a/ /e/ /i/	6
5	VC: all vowels + final {/l n j w ʔ h/}	/əl/ /aʔ / /uh/	36
6	Misc.: Ꮒ 'hna' and Ꮀ 'nah'	/hna/ /nah/	2
		Total:	**212**

Once the reclist is complete, it can then be recorded. The choice of voice to be recorded is an important one, and will have the most significant impact on the sound of the finished UTAUloid. In many language revitalization contexts, older speakers and/or skilled language users (as recognized within the community) may have the phonetic productions considered most representative or most desirable by speakers and learners. On the other hand, the act of recording for an UTAUloid voicebank is also a uniquely approachable

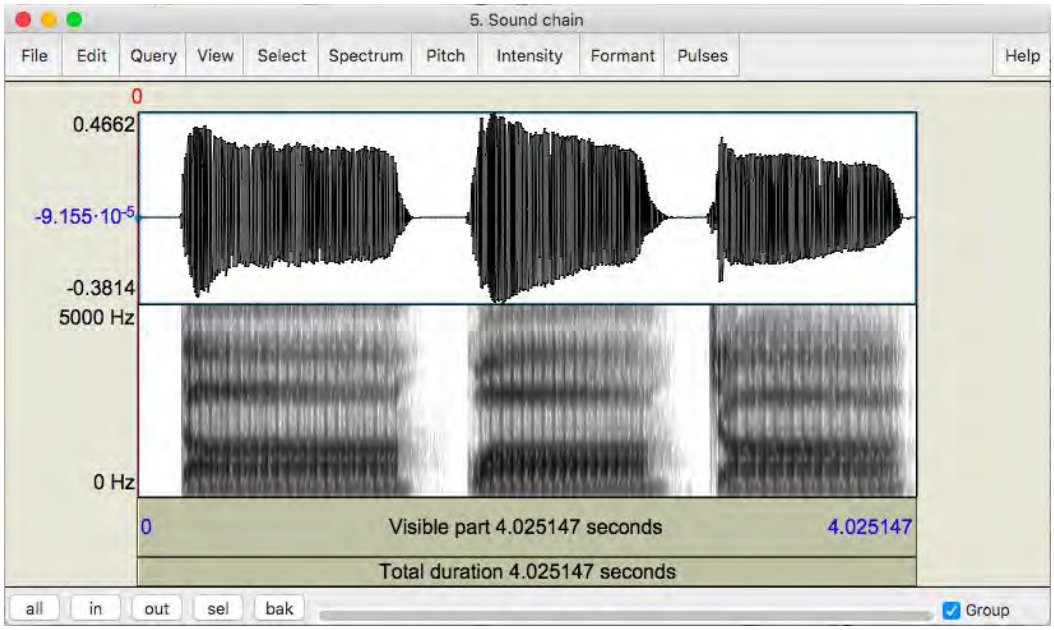
language project for heritage learners, who do not need to be even close to fluent or conversational in their language to create a fully-featured UTAUloid; all that is required is a willingness to produce the required sounds. Recording a voicebank can be a powerful act of language reclamation for heritage learners, then, both in terms of the practical pronunciation practice it entails, and in the ability to contribute to language revitalization in a significant way even early on in the language learning process. Ultimately, any speaker or learner of a language can voice an UTAUloid,[8] and recording multiple UTAUloids in a language is, of course, both feasible and beneficial. For this Cherokee UTAUloid example, I provided the voice recordings, as part of my own language reclamation process as a heritage learner of Cherokee.

For recording, UTAU requires samples to be in WAV format, but these can come from any source. They can be recorded directly onto a computer using free software like Praat (Boersma & Weenink, 2023) or Audacity (Audacity Team, 2022), or on a dedicated recording device and copied to the computer for use in UTAU. For this Cherokee UTAUloid, the reclist was recorded as a single WAV file in Praat, at 44.1 kHz in 16bit, using an Audio-Technica ATR2100-USB dynamic cardioid microphone, and then split into individual samples using Audacity.

Aside from a quiet environment and the use of a windscreen if feasible – to cut down on high frequencies which can make it difficult for the UTAU synthesis engine to blend sibilants smoothly – the main considerations when recording specifically for UTAUloid are that each syllable or mora should be as close to a steady pitch as is possible for the speaker (i.e. trying to avoid list intonation), and that the nucleus of the syllable or mora should be held long enough to ensure that the synthesis engine has enough steady state of the nucleus to adjust as needed. Figure 5 illustrates raw samples for the sequences /ta kʷa ʧa/ of the Cherokee UTAUloid:

Figure 5

Raw Samples for /ta kʷa ʧa/



After all raw audio files are recorded, they can then be 'tuned' for synthesis.

Tuning Samples for Synthesis

Once the raw audio has been recorded, the next step in turning these samples into a viable voicebank is to "tune" them for use in UTAU, by demarcating specific regions of time in each individual sound file that are relevant for manipulation by the synthesis engine – a process which directly applies linguistic skills in phonetics and spectrogram reading.

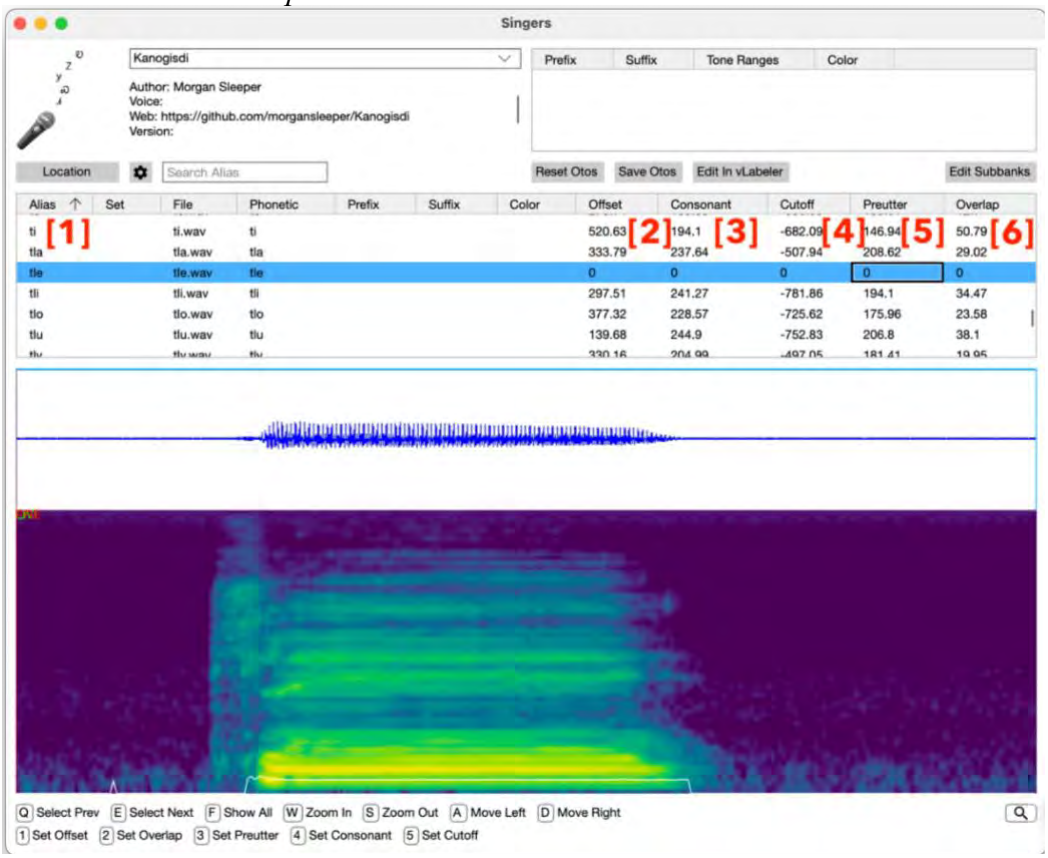
The first step to tuning is to create a plain-text file called "oto.ini"[9] in the voicebank directory listing each individual .wav file on its own line, followed by an "=," after which the tunings can be stored. For example, an excerpt from the Cherokee UTAUloid's blank "oto.ini" file before tuning would read:

o.wav=
go.wav=
yo.wav=

The tuning process then varies slightly by the UTAU software used,[10] but in OpenUTAU, it can be done through the voicebank editor in the "Singers" tool. The voicebank editor, shown in Figure 6, lists each individual .wav file present in the voicebank and displays a waveform and spectrogram of the selected sound, along with a series of editable attributes used in tuning: [1] alias, [2] offset, [3] consonant,[11] [4] cutoff,[12] [5] pre-utterance, and [6] overlap.

Figure 6

Voicebank Editor in OpenUTAU



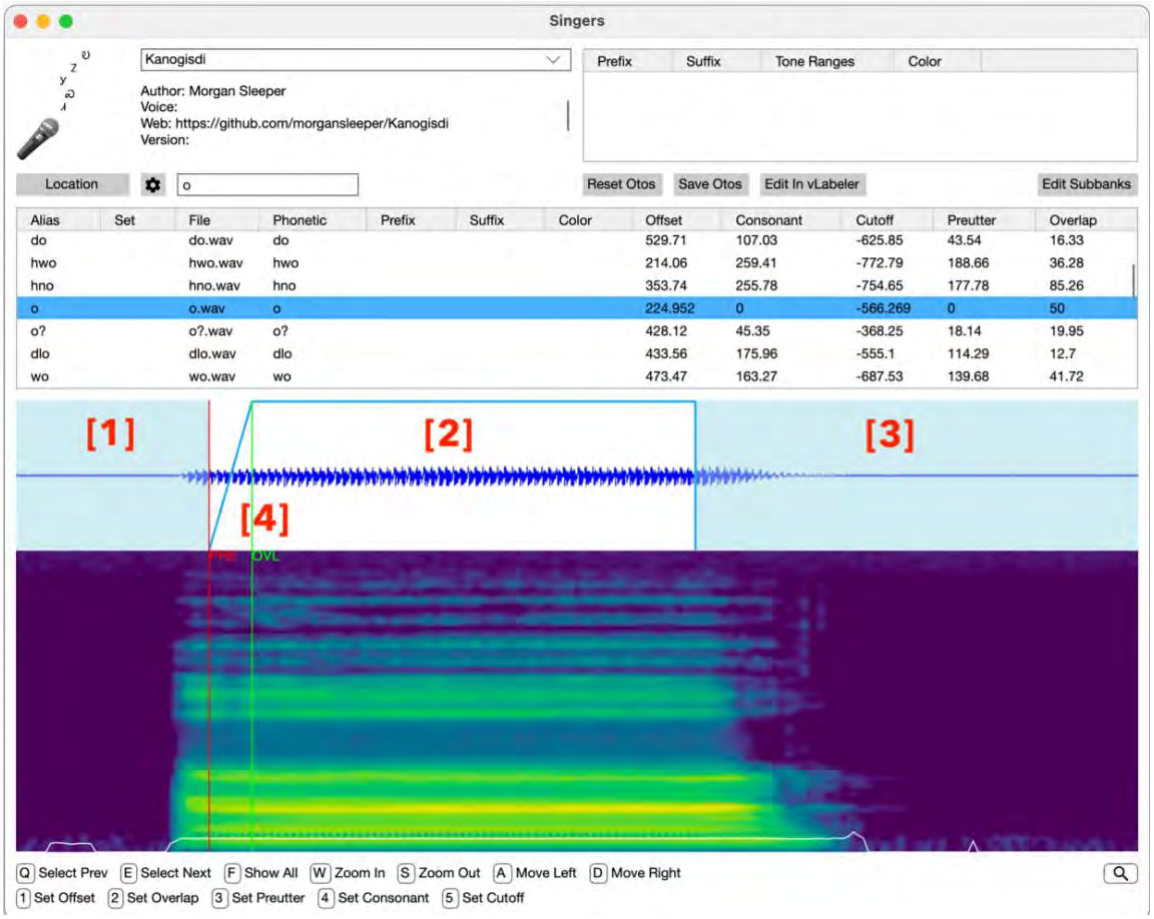
Alias ([1]) allows users to specify one or more alternate names for a given sound. This can be used to allow for samples to be referenced using either phonetic transcriptions or orthography, or for the use of more than one orthographic system in a single voicebank. In Cherokee, for instance, adding an alias could allow for a sample “o.wav” (/o/) to be called with either <o> or <ꭰ> (the syllabary character for /o/) in the score.

The rest of these attributes variously come into play while tuning, as the process varies slightly depending on the structure of the syllable or mora being tuned, with one method for syllables consisting only of vowels, and another for any syllable or mora containing any number of consonants. There are also different philosophies, personal preferences, and aesthetic considerations that can affect how a creator might tune their UTAUloid, and so rather than an exhaustive survey of UTAU tuning in general, the following subsections detail the particular tuning process used in this Cherokee UTAUloid, beginning first with vowels, and then syllables containing consonants.

Vowels

Figure 7 shows an example of tuning of a syllable consisting of only vowels (whether V, V:, VV, etc.), illustrated with the Cherokee syllable ꭰ /o/:

Figure 7
Tuning a V Syllable (ꭰ /o/)



For vowels like ꮎ /o/, the relevant attributes which can be edited in tuning are the offset (indicated by the blue-shaded region on the left [1]), the steady state (the central white region [2]), the cutoff (the blue-shaded region on the right [3]), and the overlap (indicated by the line in [4]). The offset [1] is measured in time in milliseconds relative to the start of the file and indicates where the sample will begin to be played from in synthesis in order to omit any preceding silence and often the initial, volatile state of the vowel. For vowels, the offset can be positioned at the beginning of a periodic cycle in the steady state, as seen in Figure 7 above.

The next modifiable attribute is the cutoff [3], measured in milliseconds from the end of the file, which indicates where the playback will end in synthesis, and is used to exclude the end or decay of the sound, as well as any following silence. For vowels, this can be placed at the end of a periodic cycle in the steady state, with the result that the steady state [2] – the region of the sample the synthesis engine will stretch or shrink to fit the required note length – should ideally represent a loopable, periodic series of cycles.

The final attribute for tuning vowels is the overlap [4]. Measured in milliseconds relative to the offset, it indicates how far the synthesis engine should cross-fade the previous note into a sample; the portion of the sample to the left of the overlap will be mixed with the previous sample, while the portion to the right will not. There is no specific point of the vowel that the overlap needs to be anchored to, but 50ms after the offset results in a blending sound consistent with popular UTAUloid voicebanks.

Consonants

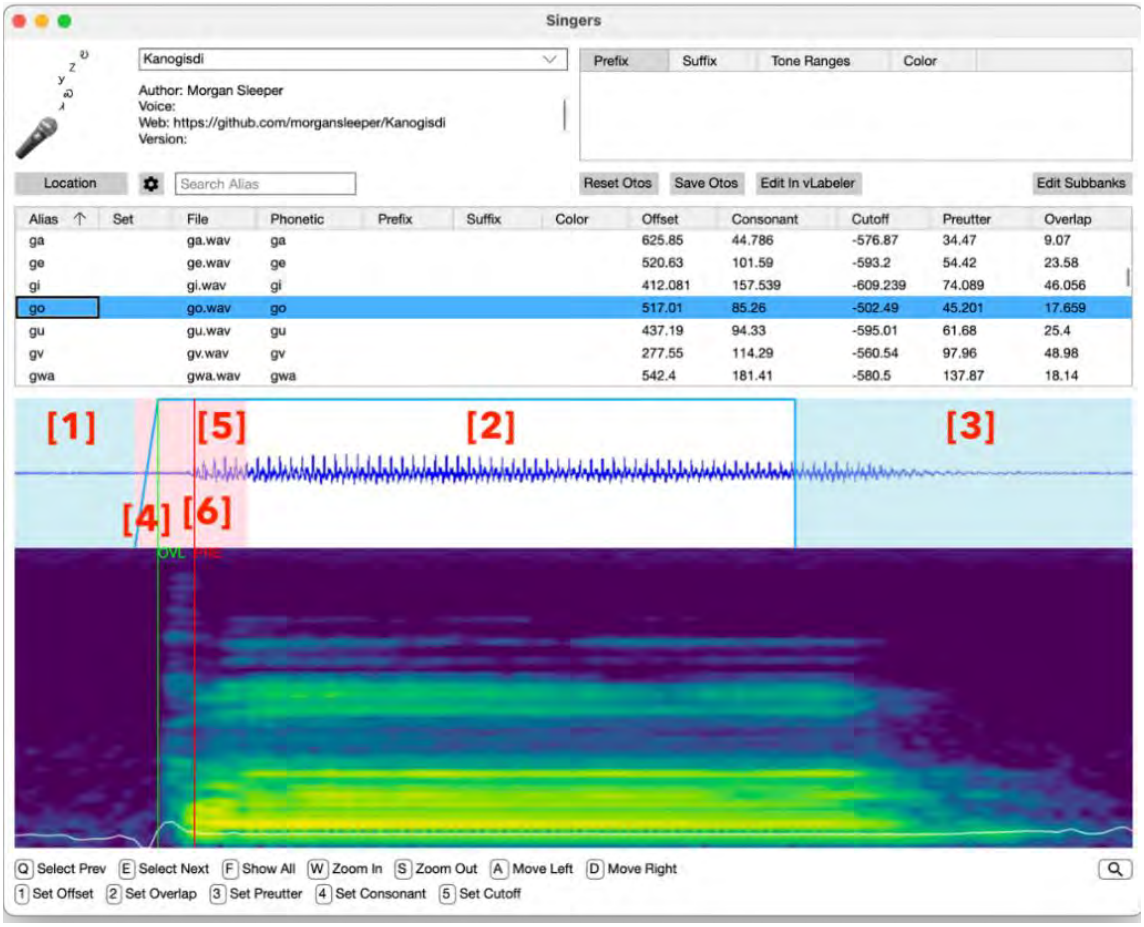
For segments containing any number of consonants (whether C_0V, VC_0, etc.), the same attributes apply, with the additional considerations of a “consonant” region and a pre-utterance point. Figure 8 shows the tuned Cherokee syllable Ꭺ /ko/.

The offset [1] performs the same function here as in the V example above: marking where the sample will begin to play, and excluding the region shaded in blue from synthesis. The cutoff [3] again does the same for marking the end of the sample, by excluding the blue region, and should extend to the end of a periodic cycle of the vowel, so that the steady state [2] represents a loopable periodic vowel sound.

The difference for syllables with consonants is that instead of being bound by the offset [1] and cutoff [3] points, the steady state [2] is here bound by the cutoff [3] and a “consonant” region [5]. The length of the consonant region [5] is measured in milliseconds relative to the offset, and this indicates the portion of the sound file that will be played in synthesis, but not manipulated in terms of length by default[13]. The entire consonant region will be played, without being either stretched or shrunk to account for note length as the steady state [2] is. In tuning consonant syllables, then, this region should include both the consonant itself and also the initial portion of the vowel affected by formant transitions – both perceptual cues which are most salient when unaltered by the synthesis engine. For the same reason, the overlap [4], which determines how far into a sample any cross-fade with a previous sample should extend to, should be placed before the consonant information, so that relevant consonant cues (such as the stop burst in Ꭺ /ko/ above) are not obscured or lost in blending with the previous sample.

Figure 8

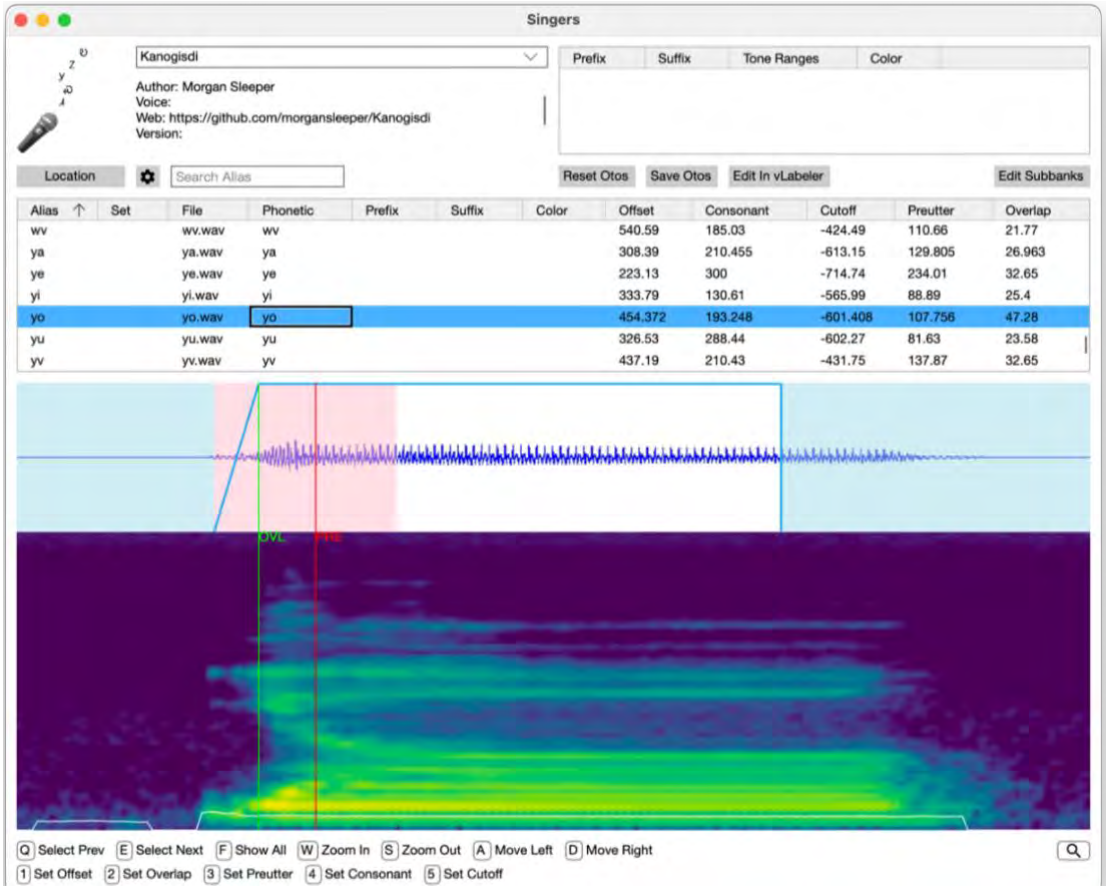
Tuning a CV Syllable (A /ko/)



The other new attribute for consonant syllables is the red pre-utterance line [6]. Measured in milliseconds relative to the offset, the pre-utterance indicates which point in the sample should be aligned with the beginning of the sung musical note in terms of rhythm. This is important because when speakers sing, rhythmic timing is organized around the nucleus of the syllable (or mora), rather than the onset; the pre-utterance allows users to align the start of the nucleus with the start of the note. In a CV syllable like Ꭺ /ko/ above, then, the pre-utterance should be positioned at the onset of the vowel. In some cases, such as when the preceding consonant is a liquid or glide, it can be difficult to determine the exact onset of a vowel from the waveform alone. To tune these samples more accurately, it is useful to refer to the spectrograms shown below the waveform, as seen in Figure 9 which shows the Cherokee syllable ꭻ /jo/:

After adjusting each of these parameters for a given sample, clicking “Save Otos” saves the tunings to the “oto.ini” file in the voicebank directory. This file will then contain the tuning information for all the sound files in the voicebank, in the format of [filename]=[alias],[offset],[consonant],[cutoff],[pre-utterance],[overlap], with values in milliseconds.

Figure 9
Tuning Ꭶ /jo/



The lines of the Cherokee UTAUloid oto.ini corresponding to the examples shown above after tuning, for example, are:

o.wav=o,224.9,0.0,-566.3,0.0,50.0
go.wav=go,517.0,85.3,-502.5,45.2,17.7
yo.wav=yo,454.4,193.2,-601.4,107.8,47.3

Since the “oto.ini” file is plain text and user-editable, this also means it is possible to tune the samples for an UTAUloid without using the OpenUTAU software, by determining the positions for each attribute in any audio editing program such as Praat or Audacity and manually entering the values into the “oto.ini.” This can be helpful when creating an UTAUloid collaboratively, as tunings from different sources (such as multiple collaborators working simultaneously on different computers) can be copied and pasted into the same “oto.ini” file without issue or created jointly through online platforms like Google Docs; it can also be useful if tuning a particular sample requires more detailed spectrogram manipulation (e.g. window length) than the view within OpenUTAU provides.

With the audio samples and a finished “oto.ini” file in the same directory, the UTAUloid is complete. Metadata about the UTAUloid (including the UTAUloid’s name, voice donors, tuners, illustrators, contributors, and contact information) can be specified in

text files called “readme.txt” and “character.txt,” which are read by UTAU software to provide information in the application, and a picture can optionally be set for the voicebank by including it as “image.png” in the same directory. The directory can then be compressed (into a .zip file, for example) and shared; anyone with the OpenUTAU application (or other UTAU software) installed can then easily load in the .zip file and use the enclosed UTAUloid.

Of course, while massive collaboration is a key part of the general UTAUloid experience, in many language revitalization contexts creators may want to restrict the usage of an UTAUloid in accordance with community norms and wishes, and it is important to note that the distribution of any UTAUloid is always completely at the discretion of its creator. An UTAUloid could be kept for use only within a given community, rather than distributed over the internet, or even not be shared at all – while UTAU allows for almost limitless collaboration, there is no requirement or expectation that a voicebank be shared or shared in any specific way.

Along with distribution, the acceptable usage of an UTAUloid can also be specified with a license provided with the voicebank. General UTAU guidelines already prohibit the use of any UTAUloid to advance racism or synthesize hate speech, among other stipulations, but many voicebank licenses also specify restrictions prohibiting commercial use, redistribution, or altering the tunings or voice samples, for example. In language revitalization contexts, licenses are especially important in that they allow UTAUloids to align with Indigenous practices of data sovereignty (Kukutai & Taylor, 2016). The open-source nature of UTAU and its community of massive collaboration has important potential for language revitalization, but as Brinklow et al. (2019) point out, “‘open source’ requires a more nuanced application in the Indigenous context, especially at the interface between ‘tool’ and ‘data’” which UTAUloids inhabit (p. 405). Crafting an appropriate license that grants access to the voicebank, the language data which it contains, and/or the music created from it in accordance with a community’s customs and protocols is a vital part of using UTAUloids in Indigenous language revitalization, and this process can in turn be a fruitful part of wider community conversations around Indigenous data sovereignty and licensing in general.[14]

For this Cherokee UTAUloid, I included a brief, bespoke license (reproduced below) based on the Cherokee concept of ᏍᏚᏯ (*gadugi*), an ethic of coming together in cooperative community labor to work towards shared goals, represented by the image of “building one fire.” As praxis, ᏍᏚᏯ is closely connected to language and cultural preservation (Cushman, 2010), and so drawing on the Cherokee Nation’s practice of making language resources – including in-person and online language classes, learning materials, and digital archives of stories and songs (Cushman, 2013) – available to anyone interested in learning Cherokee, including non-citizens, the license for this UTAUloid stipulates that it can be used freely by anyone for the purposes of furthering the Cherokee language:

> This UTAU voicebank is distributed under a ᏍᏚᏯ (*gadugi*) license. ᏍᏚᏯ is the Cherokee ethic of coming together to work towards a common goal.
>
> In the spirit of ᏍᏚᏯ, this UTAUloid can be used freely by anyone working towards our common goal of helping to revitalize the Cherokee language.

Please do not alter, modify, or redistribute this voicebank elsewhere.

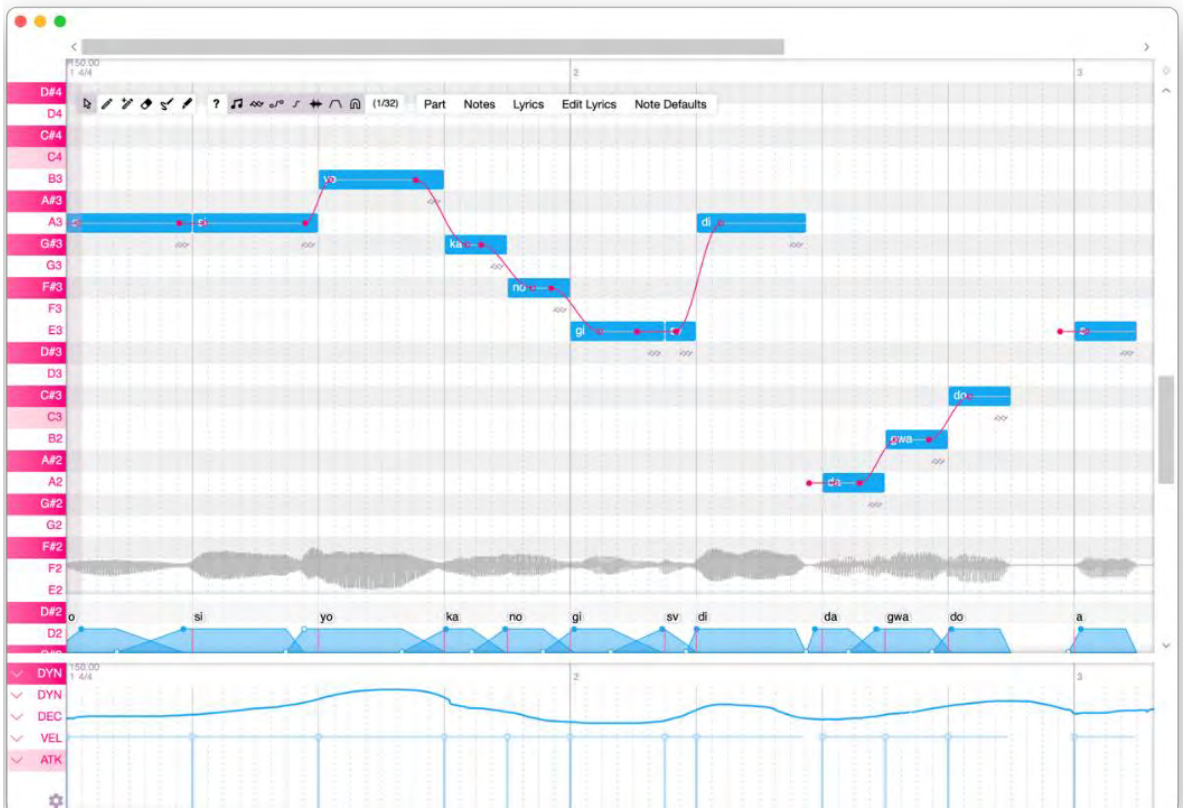
As with all UTAU voicebanks, it is prohibited to use ᏫᏃᎩᏅᏗ/Kanogisdi to advance racism, sexism, or bigotry of any kind.

ᏩᏙ!

Results and Discussion

The result of the above method is a complete Cherokee UTAUloid, ᏫᏃᎩᏅᏗ (*kanogisdi*, "singing"): an early release with much room for improvement, but which can already be used to simply and easily create vocal melodies in the language. Once the UTAUloid is loaded into OpenUTAU or other UTAU-compatible software, users place musical pitches on the piano roll-style score, with vertical height indicating pitch and width indicating length; clicking on a placed note allows users to specify the syllable to be sung, based on the file labels and aliases used in the voicebank. Figure 10 shows a simple melody composed and sung in Cherokee using ᏫᏃᎩᏅᏗ in OpenUTAU, which can be heard alongside other samples at the project's webpage:[15]

Figure 10
Using the Cherokee UTAUloid ᏫᏃᎩᏅᏗ in OpenUTAU



Initial goals for the ᎤZᏘᏫᏗ project include writing songs based on the lesson dialogues from the Cherokee Nation self-paced online classes – both as a learning exercise for my own language reclamation, and to help other learners through all the benefits music can provide to language learning – as well as collaborations with other Cherokee artists to eventually create a character model and key art for the UTAUloid.

As a tool for music creation, UTAUloids have several characteristics which make them appealing to music producers. Many of these same features also help make them uniquely useful as tools for language revitalization. The first of these features is that UTAUloid is a very versatile format for music-making; melodies composed in the framework are both portable and exportable in a variety of formats. The sung audio can be exported into a .wav file, for instance, and the resulting "vocal track" can be added to any digital audio workstation, including open-source systems such as Audacity or LMMS (LMMS Developers, 2020), and combined with additional tracks to create a complete song. UTAUloid scores can also be shared in their native UST format, which means that vocal melodies composed or started by one user can be edited, tweaked, and/or finished by any other user that they share the .ust file with. Finally, UTAUloid melodies can be exported natively to MIDI, providing for easy importing into open-source graphical music editors like MuseScore (MuseScore Contributors, 2023).

This (ex)portability makes UTAUloid particularly well-suited to collaborative work, and that, combined with the culture of "massive collaboration" around UTAUloid, creates unique potential in terms of Indigenous language revitalization. In the context of language revitalization, massive collaboration means that speakers, learners, community members – and depending on the context and license, potentially people outside of the community – of all different language and musical skill levels can all participate collaboratively in the same musical-linguistic project of creating and using UTAUloids. Speakers who may not consider themselves 'musical' could contribute lyrics or voice samples; learners who may not feel comfortable writing lyrical songs in the language could contribute melodies or take on lyric-writing as part of the language learning process, perhaps in collaboration with fluent speakers (see for example Ashton, 2020 on the Jersey Song Project); musicians could contribute melodies or backing tracks without any specific knowledge of or interest in the language; and other community members could contribute with any of the other elements of UTAUloid's inherently "multidisciplinary creation" (Sousa, 2014), including art, writing, and character and costume design. In the Cherokee case in particular, I see the concept of massive collaboration as a reflection of ᏕᏚᏯ *gadugi* in action in a musicolinguistic framework for language revitalization: coming together from different perspectives in the shared work of language and community, engaging a variety of skillsets, motivations, and lived experiences, and building "one fire from the flames and fuel of many fires" (Cushman, 2010, p. 95).

The distributed model of creation associated with massive collaboration can also be especially advantageous in languages with large, engaged digital diasporas, like Cherokee. With around 10,000 speakers in the Cherokee Nation, but over 450,000 registered tribal members spread out across the reservation, Oklahoma, the 23 recognized "At-Large" communities throughout the United States, and beyond – many of whom are strongly engaged with the Cherokee Nation and the language revival movement through the internet and social media – this digital diaspora represents tremendous creative potential for projects like UTAUloid. The massive collaboration inherent in making music with UTAUloid

represents a particularly participatory and accessible way to engage the benefits of music for language revival, as well as a unique medium to bring people together through musicking, whether in physical community, or in the shared virtual space of an Indigenous online creative community.

In addition to multidisciplinary collaboration, another of the primary reasons users turn to UTAUloid in any context is also a key advantage for language revitalization: low-resource music making. Provided someone has access to a given UTAUloid and a computer, they can make vocal music in that language for free. Combining the vocal tracks created by UTAUloid with free and open-source digital audio workstations (such as LMMS or Audacity) and their integrated instrument synthesizers then allows users to create full, commercial quality vocal songs without any monetary investment beyond a computer. UTAUloid is also an example of truly low-resource music making in terms of physical space; without the need for studio recording equipment or a low-noise environment, and with the ability to compose start-to-finish with headphones, UTAUloid songs can be composed in their entirety in any space with a computer – including quiet public spaces such as libraries, and louder shared spaces like community centers.

Perhaps most importantly in terms of language revival, UTAUloid also represents an excellent potential avenue to bolster youth involvement with language revitalization. The UTAUloid community is overwhelmingly a community of youth; many popular UTAUloids have been created by teenagers, as is the case for most of the minoritized language UTAUloids mentioned above. And crucially, rather than introducing a new concept to young people for the purposes of language revitalization, UTAUloid represents a concept, platform, and community (through both UTAUloid and Vocaloid) that is already incredibly popular with young people (Condry, 2011; Lam, 2016; Le, 2014; March, 2022; Yin, 2018). Furthermore, the countries in which UTAUloid and Vocaloid are most popular – Canada, the United States, Mexico, Japan, China, Taiwan, Malaysia, and Indonesia,[16] among others – are all countries in which colonialism and linguistic assimilationist policies have led to widespread endangerment of, and subsequent revitalization efforts in support of, Indigenous languages (Eberhard et al., 2022). As one example of this worldwide appeal, Figure 11 shows a sold-out 2016 live concert by Hatsune Miku (projected via hologram and accompanied by a live band) in San Francisco's 2,300-seat Warfield Theatre, packed completely full of young people swinging glowsticks in time to the music – music composed and created by their peers, fellow Vocaloid and UTAUloid fan-producers.

UTAUloid offers a way for young people to bring that energy to their own revitalizing languages, and to bring those languages into new, exciting domains. This can in turn be a powerful source of momentum in language revitalization; as the organizers of a youth workshop designed to create new music in Pitjantjatjara observe: “if media content created by young people is of the same high standard as other media with which they engage, then their own language content will always be more popular – we have observed this hands down” (Sometimes & Kelly, 2010, p. 88).

It is worth noting here that while Vocaloids and UTAUloids can be employed to create any type of vocal music – and users have, for example, composed operas and synthesized traditional folk songs with both – the electronic, computerized, and by design not-fully “natural” nature of the language sounds they create means that UTAUloids may not be culturally appropriate in all language or musical contexts, or compatible with protocols around musicking and song creation within a given community.

Figure 11
Hatsune Miku Concert in San Francisco, 2016 (photo by author)

Further, Indigenous language revitalization and music revitalization are interrelated (Brown et al., 2017; Grant, 2014; Marett & Barwick, 2003), and in some cases where communities are also working to protect or revitalize traditional musical practices, UTAUloids could potentially be seen as orthogonal – or even harmful – to a community's goals in cultural reclamation.

In Cherokee culture, however, there is a longstanding custom of adapting new and non-"traditional" musical genres into the Cherokee language, as part of a general cultural practice of "[making] things Cherokee simply by doing them" (Snyder, 2016, p. 62). This can be seen from the 19th century integration of hymn-singing, to more recent examples of Cherokee-language reggae, heavy metal, pop, and hip-hop on the **ᎠᏅᏛᏁᎵᏍᎩ** (*anvdvnelisgi* 'performers') compilation by Horton Records[17] (Eaton, 2022), young campers at the Snowbird Cherokee Traditions Language Camp in North Carolina performing a Cherokee translation of Lil NasX's 2019 country trap hit *Old Town Road*[18] (Knoepp, 2019), immersion school students playing with language in a Cherokee version of the *Ghostbusters* themesong (Snyder, 2016), and the experimental Cherokee-language fusion of futuristic disco, electronica, and modern dance of Elisa Harkin's **ᎦᏬᏂᏍᎩ ᏦᎢ** (*gawonisgi tsoi* 'Radio III').[19] In the Cherokee context, then, the electronic sounds and worlds afforded by UTAUloid represent an exciting new approach to explore, as well as a continuation of this tradition of drawing new musical aesthetics and practices into the Cherokee world.

Finally, while UTAUloids in this context are primarily a vehicle for language reclamation and revitalization, they also represent a unique point of fruitful collaboration between applied linguistics and Indigenous language revitalization. Because of the vital importance of language learning in Indigenous language revitalization, this collaboration is often rightfully conceptualized in terms of language education (whether in classrooms or in community) and second language acquisition. Here, applied linguistics can provide tools for language learning, teacher training, and creative and efficient ways which people can use to help learn, teach, and pass on ancestral and revitalizing languages, contextualized within the ontologies, experience, and decolonization-focused social justice lens of Indigenous language revitalization (e.g. Daniels & Sterzuk, 2022; McIvor, 2020).

What UTAUloids represent is an expansion of the applied linguistics knowledge base which can be fruitfully brought to this collaboration – in the phonetic, phonotactic, and practical recording skills required to create an UTAUloid. In this particular skillset, the UTAUloid project echoes applied linguistic approaches which leverage phonetic methods and computer technology for teaching L2 pronunciation (Chun, 1998; Chun & Jiang, 2022; Levis, 2007). But here, rather than offering frameworks for language learning, applied linguistics provides the background required to create a tool for musical language creation (the UTAUloid), which can then be taken up by Indigenous language revitalization in many different ways: as an individual project of language reclamation; as a community collaboration, potentially engaging contributors with diverse skills across space and time in one or many musical-linguistic projects; as an educational tool for songwriters or in classroom or community language pedagogies; or simply as an instrument for people to express themselves creatively in their ancestral language. The expansion of this relationship between applied linguistics and Indigenous language revitalization into the realm of music – with all of the distinctive benefits and affordances that that results in for learners and speakers – also points to the potential of bringing musical disciplines working towards language revitalization – like ethnomusicology (Ashton, 2022; Przybylski, 2018; Sparling et al., 2022) – into the conversation as well.

Daniels and Sterzuk (2022) describe the relationship between applied linguistics and Indigenous language revitalization in practical terms: “Applied linguistics exists to offer real-life solutions to language problems. We see threats to Indigenous languages as the most urgent language problem of our time and therefore understand that applied linguistics is called upon to offer solutions” (p. 15). Because of that urgency, and especially because of the demanding and potentially overwhelming nature of language revitalization work (Walsh, 2018), the intersection of applied linguistics and Indigenous language revitalization should ideally offer a variety of solutions that are innovative, exciting, meaningful for participants independent of larger-scale outcomes or metrics like speaker numbers, and – crucially – joyful; all of which, I believe, describes UTAUloid.

Conclusion

This study has aimed to introduce and explain the creation of a Cherokee UTAUloid, as an example of one way in which applied linguistics can contribute directly, using specialist skills, to the musical side of Indigenous language revitalization. Through their focus on “massive collaboration” – allowing speakers, learners, musicians, and others to work together on the same project – low-resource music production (in terms of both cost and space), and youth involvement (through existing popularity in global youth

culture), UTAUloids are uniquely well-situated to serve a variety of roles in language revitalization. And along with all the potential benefits from music made with UTAUloids – such as attracting learners and engaging speakers, creating opportunities for speech community, and supporting language learning – even the act of creating an UTAUloid itself allows speakers and learners who may not consider themselves "musical" to contribute to language revitalization through music, to engage in meaningful language reclamation resulting in tangible tools for others to take up, and to explore a unique and fruitful confluence of applied linguistics and Indigenous language revitalization.

Acknowledgements

My heartfelt thanks to Matt Gordon, Eric Campbell, Tim Cooley, Marianne Mithun, Heather Sparling, Kit Ashton, and all of the participants at the *A' Chànain Cheòlmhor: Language Revitalization through Music* workshop for their inspiring conversations and helpful feedback on earlier versions of this work. I'm also deeply grateful to Ed Fields for sharing his knowledge of the Cherokee language with me and countless others — ᏀᏙ! — and to the UTAUloid community for sharing their virtual voices and creative energy with the world.

Correspondence should be addressed to: Morgan Sleeper
Email: msleeper@macalester.edu

Notes:

[1] The symbol <v> in Cherokee romanization represents a nasalized schwa /ə̃/
[2] Samples of various Vocaloids can be heard at <https://soundcloud.com/cryptonfuturemeida>
[3] Examples of recent Hatsune Miku songs can be found at <https://soundcloud.com/hatsunemikuofficial> (for audio) and <https://www.youtube.com/@HatsuneMiku> (for videos)
[4] Formally, the software framework itself is known as 'UTAU', and the derived synthesizers are known as 'UTAUloids', but 'UTAUloid' is used here as a convenient term for the combined concept, on analogy with 'Vocaloid'.
[5] Available at <http://utau2008.xrea.jp/>
[6] Available at <http://utau-synth.com/>
[7] Available at <http://www.openutau.com>
[8] One important exception to this is that UTAU usage guidelines forbid the creation of a voicebank from a speaker without their express consent. This includes (specifically) deceased persons, which means that archival recordings of speakers who have passed away should not serve as voicebank sources.
[9] From *oto* (音), meaning 'sound' in Japanese
[10] There are also standalone programs designed specifically for UTAU tuning; one popular option is vLabeler (available at <https://vlabeler.com>), which will also create and update the 'oto.ini' file automatically.
[11] Also referred to as 'fixed' in other UTAU software
[12] Also referred to as 'blank'
[13] The length of the consonant region can optionally be manipulated in OpenUTAU for stylistic effect on a per-note basis with the 'VEL' (velocity) parameter.
[14] As one example, the Kaitiakitanga License developed by Māori media organization Te Hiku Media (2023) is based on the Māori concept of *kaitiakitanga* (loosely, 'guardianship') rather than ownership of data, and was created to protect against the digital colonization of Māori data and knowledge.
[15] <https://github.com/morgansleeper/Kanogisdi>
[16] These countries have all hosted the Hastune Miku Expo Vocaloid concert series – a concert tour presented by Crypton Future Media featuring their Vocaloids singing as holographs, accompanied by live band
[17] <https://hortonrecords.bandcamp.com/album/anvdvnelisgi-performers>

[18] <https://soundcloud.com/tsalagiseli/old-town-road-cherokee-version>
[19] <https://elisaharkins.bandcamp.com/album/radio-iii>

References

Abrate, J. H. (1983). Pedagogical applications of the French popular song in the foreign language classroom. *The Modern Language Journal*, *67*(1), 8–12. https://doi.org/10.1111/j.1540-4781.1983.tb02492.x

Antoine, J. (2015). The role of traditional songs in the maintenance and preservation of Lakota language and culture. In N. Ostler, & B. W. Lintinger (Eds.), *Proceedings of foundation for endangered languages* (Vol. 19, pp. 17–23).

Archer, J. (2002, October 4). *Poi-E*. New Zealand Folk Song. https://folksong.org.nz/poi_e/

Ashton, K. (2020). The Jersey song project. In J. Olko, & J. Sallabank (Eds.), *Revitalizing endangered languages: A practical guide* (pp. 289–290). Cambridge University Press.

Ashton, K. (2022). *Êcliaithe Man Tchoeu [Light Up My Heart]: Applied ethnomusicology and the revitalisation of the endangered language of Jèrriais* [PhD Thesis]. Goldsmiths, University of London, England, U.K.

Audacity Team. (2022). *Audacity®: Free Audio Editor and Recorder* (3.2.5) [Computer software]. https://www.audacityteam.org/

Barrett, R. (2016). Mayan language revitalization, hip hop, and ethnic identity in Guatemala. *Language & Communication*, *47*, 144–153. https://doi.org/10.1016/j.langcom.2015.08.005

BBC News. (2018, October 20). Gwenno "sparks record numbers" in Cornish exams. *BBC News*. https://www.bbc.com/news/uk-england-cornwall-45917661

Berez-Kroeker, A. L., Gabber, S., & Slayton, A. (2023). Recent Advances in Technologies for Resource Creation and Mobilization in Language Documentation. *Annual Review of Linguistics*, *9*,195–214. https://doi.org/10.1146/annurev-linguistics-031220-120504

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* (6.3.09) [Computer software]. http://www.praat.org/

Bracknell, C. (2020). Rebuilding as research: Noongar song, language and ways of knowing. *Journal of Australian Studies*, *44*(2), 210–223. https://doi.org/10.1080/14443058.2020.1746380

Bracknell, C., Bracknell, K., Fenty Studham, S., & Fereday, L. (2021). Supporting the performance of Noongar language in Hecate. *Theatre, Dance and Performance Training*, *12*(3), 377–395. https://doi.org/10.1080/19443927.2021.1943506

Brinklow, N. T., Littell, P., Lothian, D., Pine, A., & Souter, H. (2019). Indigenous language technologies & language reclamation in Canada. In European Language Resources Association (Ed.), *Language technology for all (LT4All): Enabling language diversity and multilingualism worldwide* (pp. 402–406). UNESCO Publishing.

Brown, R., Manmurulu, D., Manmurulu, J., O'Keeffe, I., & Singer, R. (2017). Maintaining song traditions and languages together at Warruwi (western Arnhem Land). In J. W. Wafer, & M. Turpin (Eds.), *Recirculating songs: Revitalising the singing practices of indigenous Australia*. Hunter Press.

Caldwell, E. (2023, March 27). Can a newly installed cellphone tower help preserve a language? NPR. https://www.npr.org/2023/03/27/1166180027/can-a-newly-installed-cell-phone-tower-help-preserve-a-language

Cherokee Nation. (2022). *Cherokee national youth choir*. Cherokee Nation Education Services. https://www.cherokee.org/all-services/education-services/youth-

leadership/cherokee-national-youth-choir/
Cherokee Phoenix Staff. (2022, August 9). *Cherokee Nation launches innovative speaker services program*. Cherokee Phoenix https://www.cherokeephoenix.org/services/cherokee-nation-launches-innovative-speaker-services-program/article_dcd76e94-17e7-11ed-9509-ffd3a6635251.html
Chun, D. M. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology*, *2*(1), 74–93. http://dx.doi.org/10125/25033
Chun, D. M., & Jiang, Y. (2022). Using technology to explore L2 pronunciation. In J. M. Levis, T. M. Derwing, & S. Sonsaat-Hegelheimer (Eds.), *Second language pronunciation: Bridging the gap between research and teaching* (pp. 129–150). John Wiley & Sons.
Clusternote. (2014). *Vocal synthesizer piano roll—Sakura Sakura.jpg* [Digital image]. https://commons.wikimedia.org/wiki/File:Vocal_synthesizer_piano_roll_-Sakura_Sakura.jpg
Condry, I. (2011, July 11). *Miku: Japan's virtual idol and media platform.* MIT Center for Civic Media. https://civic.mit.edu/index.html%3Fp=1749.html
Cotter, C. (2001). Continuity and vitality: Expanding domains through Irish-language radio. In L. Hinton, & K. Hale (Eds.), *The green book of language revitalization in practice* (pp. 301–316). Brill.
Cru, J. (2018). Micro-level language planning and YouTube comments: Destigmatising indigenous languages through rap music. *Current Issues in Language Planning*, *19*(4), 434–452. https://doi.org/10.1080/14664208.2018.1468960
Crypton Future Media. (n.d.). *Who is Hatsune Miku?* Retrieved March 31, 2023, from https://ec.crypton.co.jp/pages/prod/virtualsinger/cv01_us
Cushman, E. (2010). Gadugi: Where the fire burns (still). In S. Kahn, & J. Lee (Eds.), *Activism and rhetoric: Theories and contexts for political engagement* (pp. 56–61). Routledge.
Cushman, E. (2012). *The Cherokee syllabary: Writing the people's perseverance*. University of Oklahoma Press.
Cushman, E. (2013). Wampum, Sequoyan, and story: Decolonizing the digital archive. *College English*, *76*(2), 115–135. http://www.jstor.org/stable/24238145
Daniels, B., & Sterzuk, A. (2022). Indigenous language revitalization and applied linguistics: Conceptualizing an ethical space of engagement between academic fields. *Canadian Journal of Applied Linguistics*, *25*(1), 1–18. https://doi.org/10.37213/cjal.2022.31841
Danos, D., & Turin, M. (2021). Living language, resurgent radio: A survey of Indigenous language broadcasting initiatives. *Living Language*, *15*, 75-152 http://hdl.handle.net/10125/24971
Davis, G. M. (2017). Songs in the young learner classroom: A critical review of evidence. *ELT Journal, 71*(4), 445-455. https://doi.org/10.1093/elt/ccw097
Dołowy-Rybińska, N. (2020). Fest-noz and the revitalisation of the Breton language. In J. Olko, & J. Sallabank (Eds.), *Revitalizing endangered languages: A practical guide* (pp. 287–288). Cambridge University Press.
Eaton, K. (2022, April 19). *Cherokee artists are preserving their language through music*. Good Good Good. https://www.goodgoodgood.co/articles/cherokee-language-music-album
Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2022). *Ethnologue: Languages of*

the world (25th ed.). SIL International. http://www.ethnologue.com
Engh, D. (2013). Why use music in English language learning? A survey of the literature. *English Language Teaching*, *6*(2), 113-127. https://doi.org/10.5539/elt.v6n2p113
Golla, V. (2010). North America. In C. Moseley (Ed.), *Encyclopedia of the world's endangered languages* (pp. 1–96). Routledge.
Good, A. J., Russo, F. A., & Sullivan, J. (2015). The efficacy of singing in foreign-language learning. *Psychology of Music*, *43*(5), 627–640.
Goodyear, S. (Director). (2019, May 1). *Meet the N.S. teenager who sang Blackbird by The Beatles entirely in Mi'kmaw.* CBC Radio. https://www.cbc.ca/radio/asithappens/as-it-happens-wednesday-edition-1.5118294/meet-the-n-s-teenager-who-sang-blackbird-by-the-beatles-entirely-in-mi-kmaw-1.5118296
Grant, C. (2014). *Music endangerment: How language maintenance can help*. Oxford University Press.
Hinton, L. (2001). Language Revitalization: An Overview. In L. Hinton, & K. Hale (Eds.), *The green book of language revitalization in practice* (pp. 3–18). Academy Press.
Jadii. (2009). *Sachi Eika UTAUloid*. http://terraloid.tumblr.com/sachi_eika
Johnson, H. (2012). "The group from the west": Song, endangered language and sonic activism on Guernsey. *Journal of Marine and Island Cultures*, *1*(2), 99–112. https://doi.org/10.1016/j.imic.2012.11.006
Jolly, Y. S. (1975). The use of songs in teaching foreign languages. *The Modern Language Journal*, *59*(1/2), 11–14. https://doi.org/10.2307/325440
Kenmochi, H. (2010). VOCALOID and Hatsune Miku phenomenon in Japan. *Proceedings of InterSing 2010: First Interdisciplinary Workshop on Singing Voice*, 1–4. https://www.isca-archive.org/intersinging_2010/kenmochi10_intersinging.pdf
Kenmochi, H., & Ohshita, H. (2007). VOCALOID - Commercial singing synthesizer based on sample concatenation. *Proceedings of Interspeech 2007*, 4009–4010. https://www.isca-archive.org/interspeech_2007/kenmochi07_interspeech.pdf
Knoepp, L. (2019, September 24). *Heading down the 'old town road' to teach the Cherokee language*. Blue Ridge Public Radio. https://www.bpr.org/news/2019-09-24/heading-down-the-old-town-road-to-teach-the-cherokee-language
Kukutai, T., & Taylor, J. (2016). Data sovereignty for Indigenous peoples: Current practice and future needs. In T. Kukutai, & J. Taylor (Eds.), *Indigenous data sovereignty: Toward an agenda* (pp. 1–22). Australian National University Press.
Lam, K. Y. (2016). The Hatsune Miku phenomenon: More than a virtual J-pop diva. *The Journal of Popular Culture*, *49*(5), 1107–1124. https://doi.org/10.1111/jpcu.12455
Le, L. K. (2014). Examining the rise of Hatsune Miku: The first international virtual idol. *The UCI Undergraduate Research Journal*, 1-12.
Leonard, W. Y. (2012). Reframing language reclamation programmes for everybody's empowerment. *Gender and Language*, *6*(2), 339–367. https://doi.org/10.1558/genl.v6i2.339
Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, *27*, 184–202. https://doi.org/10.1017/S0267190508070098
Llewellyn, M. (2000). Popular music in the Welsh language and the affirmation of youth identities. *Popular Music*, *19*(3), 319–339. https://doi.org/10.1017/S0261143000000192

LMMS Developers. (2020). *LMMS* (1.2.2) [Computer software]. https://lmms.io/

Lucas, O. R. (2021). *Kaitiakitanga, Whai Wāhi* and alien weaponry: Indigenous frameworks for understanding language, identity and international success in the case of a Māori metal band. *Popular Music*, *40*(2), 263–280. https://doi.org/10.1017/S0261143021000131

MacKinnon, K. (2005). Cornish/Kernewek. In D. O'Néill (Ed.), *Rebuilding the Celtic languages: Reversing language shift in the Celtic countries* (pp. 214–274). Y Lolfa.

March, L. (2022). "Wrap you up in my blue hair": Vocaloid, Hyperpop, and Identity in "Ashnikko Feat. Hatsune Miku – Daisy 2.0." *Television & New Media*, *24*(8), 894-910. https://doi.org/10.1177/15274764221093599

Marett, A., & Barwick, L. (2003). Endangered songs and endangered languages. In J. Blythe, & R. M. Brown (Eds.), *Maintaining the links: Language identity and the land* (pp. 144–151). Foundation for Endangered Languages.

McCarty, T. L., Nicholas, S. E., Chew, K. A. B., Diaz, N. G., Leonard, W. Y., & White, L. (2018). Hear our languages, hear our voices: Storywork as theory and praxis in Indigenous-language reclamation. *Daedalus*, *147*(2), 160–172. https://doi.org/10.1162/DAED_a_00499

McIvor, O. (2020). Indigenous language revitalization and applied linguistics: Parallel histories, shared futures? *Annual Review of Applied Linguistics*, *40*, 78–96. https://doi.org/10.1017/S0267190520000094

Mithun, M. (1999). *The languages of native North America*. Cambridge University Press.

Montgomery-Anderson, B. (2015). *Cherokee reference grammar*. University of Oklahoma Press.

Moseley, C. (Ed.). (2010). *Atlas of the world's languages in danger*. UNESCO Publishing.

MuseScore Contributors. (2023). *MuseScore* (4.0.2) [Computer software]. https://musescore.org/

Nummelin, G. (2020). One song, many voices: Revitalising Ainu through music. In J. Olko & J. Sallabank (Eds.), *Revitalizing endangered languages: A practical guide* (pp. 291–293). Cambridge University Press.

Peter, L., Hirata-Edds, T., Feeling, D., Kirk, W., Mackey, R. "Wahde," & Duncan, P. T. (2017). The Cherokee Nation immersion school as a translanguaging space. *Journal of American Indian Education*, *56*(1), 5-31. https://doi.org/10.5749/jamerindieduc.56.1.0005

Przybylski, L. (2018). Bilingual hip hop from community to classroom and back: A study in decolonial applied ethnomusicology. *Ethnomusicology*, *62*(3), 375–402. https://doi.org/10.5406/ethnomusicology.62.3.0375

Przybylski, L. (2021). Indigenizing the mainstream: Music festivals and Indigenous popular music. *IASPM Journal*, *11*(2), 5–21. https://doi.org/10.5429/2079-3871(2021)v11i2.2en

Pulte, W., & Feeling, D. (1975). Outline of Cherokee grammar. In W. Pulte (Ed.), *Cherokee-English dictionary* (pp. 235–355). Cherokee Nation of Oklahoma.

Salazar, J., Belmar, G., Scanlon, C., Troiani, G., & Campbell, E. W. (2021). Bridging diaspora: Technology in the service of the revitalization of Sà'án Sàvĭ ñà Yukúnanĭ (Mixtec). *Endangered Languages and Diaspora - XXV Annual Conference Proceedings*, 176–185. https://www.linguistics.ucsb.edu/sites/default/files/sitefiles/people/campbell/Salazar

etal-bridgingdiaspora-2021_11_04.pdf
Schmidt, J. (2003). German rap music in the classroom. *Die Unterrichtspraxis/Teaching German*, *36*(1), 1-14. https://doi.org/10.2307/3531679
Sheehan, M. (2016). Mana Wahine: Māori women in music. *Te Kaharoa*, *9*(1), 76-90. https://doi.org/10.24135/tekaharoa.v9i1.12
Snyder, S. L. (2016). *Poetics, performance, and translation in Eastern Cherokee language revitalization* [PhD Thesis]. Columbia University, New York, New York, U.S.
Sometimes, B., & Kelly, A. (2010). Ngapartji Ngapartji: Indigenous language in the arts. In J. Hobson, K. Lowe, S. Poetsch, & M. Walsh (Eds.), *Re-awakening languages: Theory and practice in the revitalisation of Australia's Indigenous languages* (pp. 85–89). Sydney University Press.
Sousa, A. M. D. de. (2014). A colaboração massiva de Hatsune Miku: Software vocaloid como catalisador de criações colectivas, grassroots e multidisciplinares na subcultura otaku [The massive collaboration of Hatsune Miku: Vocaloid software as a catalyst for collective, grassroots and multidisciplinary creation in otaku subculture]. *Revista Croma, Estudos Artísticos*, *2*(3), 121–137. http://hdl.handle.net/10451/12237
Sparling, H., MacIntyre, P., & Baker, S. (2022). Motivating traditional musicians to learn a heritage language in Gaelic Nova Scotia. *Ethnomusicology*, *66*(1), 157–181. https://doi.org/10.5406/21567417.66.1.09
Te Hiku Media. (2023). *Kaitiakitanga License*. https://github.com/TeHikuMedia/Kaitiakitanga-License/blob/tumu/LICENSE.md
Tegge, F. (2018). Pop songs in the classroom: Time-filler or teaching tool? *ELT Journal*, *72*(3), 274–284. https://doi.org/10.1093/elt/ccx071
Tuttle, S. G. (2012). Language and music in the songs of Minto, Alaska. *Language Documentation and Description*, *10*, 82–112. https://doi.org/10.25894/ldd191
Tuttle, S. G., & Lundström, H. (2015). Taking charge: Learner agency in the transmission of song and speech traditions. In N. Ostler, & B. W. Lintinger (Eds.), *Proceedings of Foundation for Endangered Languages* (Vol. 19, pp. 38–44).
Uchihara, H. (2016). *Tone and accent in Oklahoma Cherokee* (1st ed.). Oxford University Press.
UTAU Wikia contributors. (n.d.). *Sachi Eika*. UTAU Wikia. Retrieved March 31, 2023, from http://utau.wikia.com/wiki/Sachi_Eika
Vallejo, J. M. (2019). Revitalising language through music: A case study of music and culturally grounded pedagogy in two Kanien'ke:ha (Mohawk) language immersion programmes. *Ethnomusicology Forum*, *28*(1), 89–117. https://doi.org/10.1080/17411912.2019.1641124
Walsh, M. (2018). "Language is like food..." Links between language revitalization and health and well-being. In L. Hinton, L. M. Huss, & G. Roche, *The Routledge handbook of language revitalization* (pp. 5–12). Routledge New York. https://doi.org/10.4324/9781315561271
Yin, Y. (2018). Vocaloid in China: Cosmopolitan music, cultural expression, and multilayer identity. *Global Media and China*, *3*(1), 51–66. https://doi.org/10.1177/2059436418778600
ᎠᎾᏁᎳᏗ/Anadisgoi. (2023, February 21). *Cherokee Nation reaches 450,000 Cherokee citizens*. ᎠᎾᏁᎳᏗ/Anadisgoi. https://anadisgoi.com/index.php/government-stories/cherokee-nation-reaches-450-000-cherokee-citizens