Creating an intelligent dictionary of Tsuut’ina one verb at a time

Christopher Cox
Tsuut'ina Gunaha Institute / Carleton University
christopher.cox@tsuutina.com

Bruce Starlight
Tsuut'ina Nation
spottedeagle1947@yahoo.com

Janelle Crane-Starlight
Tsuut'ina Nation
janelle.crane@tsuutina.com

Hanna Big Crow
Tsuut'ina Gunaha Institute
hanna.bigcrow@tsuutina.com

Antti Arppe
University of Alberta
arppe@ualberta.ca

Abstract

In this paper, we discuss the development of a long-term partnership between community and university-based language workers to create supportive language technologies for Tsuut'ina, a critically endangered Dene language spoken in southern Alberta, Canada. Initial development activities in this partnership sought to rapidly integrate existing language materials, with the aim of arriving at tools that would be effective and impactful for community use by virtue of their extensive lexical coverage. We describe how, as this partnership developed, this approach was gradually superseded by one that involved a more targeted, lexical-item-by-lexical-item review process that was directly informed by other community language priorities and connected to the work a local language authority. We describe how this shift in processes correlated with other changes in local language programs and priorities, noting how ongoing communication allowed this partnership to adapt to the evolving needs of local organizations.

1 Introduction

Tsuut’ina (ISO 639-3: srs, Glottocode: sars1236) is a Dene language spoken by members of the Tsuut'ina Nation, a signatory to Treaty 7 in present-day southern Alberta, Canada. Together with Plains Apache, Tsuut'ina is one of only two Dene languages spoken on the Great Plains, and is separated from other members of the Dene language family by surrounding Algonquian and Siouan-speaking Indigenous nations. As of October 2024, there are 18 first-language speakers of Tsuut'ina, all over the age of 75 and almost all residing at the Tsuut'ina Nation (Tsuut'ina Gunaha Institute, p.c.). While Tsuut'ina is thus critically endangered, strong connections between Tsuut’ina language and community identity and culture have fostered equally strong retention of Tsuut’ina language proficiency among present-day speakers. These same connections have also encouraged community-based language documentation, education, and revitalization initiatives, including those supported by collaborations with individuals and organizations outside of the Tsuut'ina Nation, as discussed in this paper.

From a linguistic perspective, Tsuut'ina closely resembles other Dene languages, with complex, prefixing, polysynthetic verbal morphology (cf. Cook, 1984; Rice, 2000; Rice, 2020). Tsuut'ina also relies heavily on tone to convey both lexical and grammatical distinctions, having one of the largest inventories of tone contrasts attested in the Dene language family (cf. Sapir, 1925; McDonough et al., 2013; Starlight & Cox, 2024). While previous research on the language conducted by both Tsuut'ina and non-Tsuut'ina linguists has resulted in notable collections of textual and grammatical documentation (e.g., Goddard, 1915; Onespot & Sapir, 1922; Cook, 1984; Starlight, Moore & Cox 2018; among others), linguistic research into aspects of Tsuut'ina grammar is ongoing, with many areas of grammatical organization still under active investigation. Both the presence of open questions concerning basic grammatical features of

the language (e.g., how many tone and vowel length contrasts should be recognized; Starlight & Cox, 2023) and the degree to which the overall profile of the language differs from neighbouring Indigenous languages and from English present particular challenges, both for current Tsuut'ina language learners and teachers who are aiming to acquire and convey the language effectively and for efforts to develop approaches and resources that support 'front-line' language revitalization work.

2 History – How the partnership came about

The partnership described in this paper has deep roots in language education, documentation, and revitalization initiatives at Tsuut'ina Nation. Since the early 1970s, members of the Tsuut'ina Nation, recognizing a significant shift in the number of first-language speakers, began implementing programs aimed at supporting Tsuut'ina language retention and intergenerational language transmission. Over several decades, these efforts resulted in the establishment of K–12 school-based language education programs, Tsuut'ina literacy programs for L1 speakers, and the adoption of a standard Tsuut'ina orthography (cf. Cook, 1984: 1–2), alongside concurrent work to develop classroom resource materials, documentation with Tsuut'ina Elders, and an initial language curriculum (Calgary Roman Catholic Separate School Division, 1996).

While the direction of these initiatives was determined and led by the Tsuut'ina Nation, on several occasions, members of the Tsuut'ina Nation also sought out partnerships with individuals and organizations outside of the Nation. The second author of this paper, Dr. Bruce Starlight, a linguist and fluent Tsuut'ina speaker who had been involved in language revitalization and documentation initiatives since 1972, worked extensively to develop such relationships, collaborating with non-Tsuut'ina colleagues to support local language programs and projects. This included extensive work with Gary Donovan at the University of Calgary on the creation of pedagogical resources for Tsuut'ina and Sally Rice at the University of Alberta on Tsuut'ina language documentation and revitalization programs. Bruce's involvement in university-based programs also extended to linguistic field methods courses at the University of Alberta in 2007 and 2009, where the first author of this paper, Christopher Cox, a linguist with an interest in community-based language work, became involved in Tsuut'ina language programs as a student volunteer during his graduate studies.

Relationships such as these continued to develop in parallel with language programs at Tsuut'ina Nation, where community interest in Tsuut'ina language revitalization continued to grow. In 2008, the Tsuut'ina Nation established the Tsuut'ina Gunaha Institute, the body within Tsuut'ina Nation tasked with supporting the full revitalization of the Tsuut'ina language. Bruce served as the Institute's founding director until 2012, when he was invested as the first Tsuut'ina Language Commissioner, a position that oversaw the development of Tsuut'ina language documentation, contributed to the visibility of the language (e.g., through the translation of public signage into Tsuut'ina), and ensured the continued integrity of the language. The creation of both of these offices was accompanied by a substantial expansion in the resources and positions available for local language revitalization programs, providing opportunities for many younger Tsuut'ina Nation members to engage with local language work on a full-time basis. It was during this period that the third and fourth authors of this paper, Janelle Crane-Starlight and Hanna Big Crow, joined the Tsuut'ina Gunaha Institute, eventually coming to serve as the Executive Director of Language and Culture for Tsuut'ina Nation (Janelle) and the Director of the Tsuut'ina Institute (Hanna).

As language programs continued to expand at Tsuut'ina Nation over the past decade, both the Office of the Tsuut'ina Language Commissioner and the Tsuut'ina Gunaha Institute noted an increased demand for resources that supported Tsuut'ina language education, documentation, and revitalization activities in digital contexts, particularly as activities in all of these areas moved increasingly into the digital realm. This shift not only resulted in more emphasis being placed on developing new Tsuut'ina language resources in digital formats, suitable for use in community-based programs, but also increased access to information found in existing, non-digital language materials; support for continued teacher training for Tsuut'ina language educators; and tools that could assist in creating such resources quickly and reliably, such as spell-checkers, predictive text systems, and text-to-speech applications. Through the network of relationships that had been

developed with the University of Alberta previously, colleagues at Tsuut'ina Nation were introduced to the fifth author of this paper, Antti Arppe, a linguist who had been involved in recent years in supporting the development of language technologies and morphologically aware online dictionaries for other Indigenous languages in North America, drawing in part on computational infrastructures developed for Indigenous language technologies in northern Eurasia (Trosterud, 2006).

Intelligent online dictionaries combine a lexical database, with entries organized under citation forms and their (English) translations, with a computational model of the word-structure of a language (Johnson et al., 2013). Firstly, this "intelligence" allows the online dictionary to recognize all inflected word-forms for the entries (which the model covers, of course), to provide linguistic analyses for these word-forms and to link those to their citation forms. Secondly, one can use the computational model in reverse and generate full inflectional paradigms for each of the citation forms; for verbs, such inflectional paradigms are often called *conjugations* (following French via Latin); for nouns, the corresponding paradigms would be known as *declinations*. For languages with a rich (inflectional) morphology, as is the case for many Indigenous languages spoken in North America, and in particular the Dene languages, such "intelligence" is indispensable, as any lexeme can have tens if not hundreds or thousands of inflected word-forms, which would be impossible to harvest from corpora of any size, and impractical to store exhaustively as their own dictionary entries.

Over a period of two years, the authors of this paper began to meet informally and discuss a potential collaboration to expand such tools to support Tsuut'ina. This began modestly by arranging initial, in-person meetings between all partners at Tsuut'ina Nation, which focused on becoming better acquainted with one another, sharing information about current priorities for language programs at Tsuut'ina Nation (for Bruce, Janelle, and Hanna), and what a partnership to develop digital tools could realistically contribute (for Antti and Chris). For Tsuut'ina partners, developing an intelligent dictionary had the potential to respond to several priorities for supporting second-language learners and local language revitalization programs. First, an online dictionary was seen as potentially improving access to Tsuut'ina documentation materials from previous and ongoing/future language projects for second-language learners. Such materials serve a crucial purpose for Tsuut'ina language learners and teachers as a resource for language education programs, self-study, and other language revitalization resource development initiatives. Second, it was also recognized that tools that could model and present inflectional patterns could be particularly valuable for Tsuut'ina second-language learning and teaching, since verbs and verb paradigms are critical to using and understanding the language.

While these initial meetings suggested that a collaboration might indeed be desirable, having concrete discussions around tools and technologies that did not yet exist for Tsuut'ina and that had few familiar precedents among other Indigenous languages sometimes proved challenging. While it was possible to discuss what already been accomplished for other, neighbouring Indigenous languages, it was with the preparation of still mock-ups of what Tsuut'ina-specific digital tools could look like (e.g., screenshots of a browser window showing an example paradigm from the intelligent online dictionary for another language, with all paradigm entries replaced with Tsuut'ina word-forms and the layout adapted to fit Tsuut'ina tense and aspect categories) that the group found a way to effectively conceptualize and discuss what these tools could accomplish. For example, in Figure 1, the Tsuut'ina verb form *nàguts'idáátlíł* is recognized and analyzed as the Progressive Fourth Person form of the Intransitive Verb *nàgudiitłod*, meaning roughly "he/she/it jumps down", for which an entry exists in the lexical database, and to which this inflected word-form is linked. If the user then would click on the entry, this would yield an inflectional paradigm, giving all the person forms in the various aspects, of which an exemplary sample is provided here. The mockup in Figure 1 was created by taking an earlier version of an intelligent dictionary for another, unrelated Indigenous language spoken in Canada, and replacing the content with correct Tsuut'ina elements.

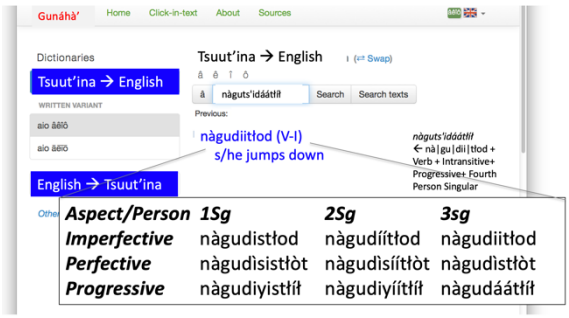

Figure 1: A mock-up of an intelligent dictionary entry for the Tsuut'ina lemma *nàgudiitłod* 'he/she/it jumps down'.

These initial meetings and co-design sessions quickly led to exchanging further ideas and information, with university-based partners drawing on the Giella infrastructure to prototype a preliminary computational model of Tsuut'ina verbal morphology and bootstrap working demos of a number of text-proofing tools and an online dictionary. Preparing and sharing presentations with Tsuut'ina Nation leadership of these tools in action—for example, with videos of spell-checking suggestions being offered for Tsuut'ina words while editing a document in LibreOffice, or of searching for morphologically complex Tsuut'ina words and being presented with their lemmas in a morphologically aware online Tsuut'ina dictionary—both teams found support for an application to the Social Sciences and Humanities Research Council of Canada (SSHRC) for a seven-year Partnership Grant, "21st Century Tools for Indigenous Languages", in which Tsuut'ina Nation would serve as one of two lead Indigenous partner nations. The awarding of this grant in 2019 allowed collaboration on these tools to move ahead on an expanded scale, with the promise of stable financial support for project activities until 2026.

3 Documentation, description, and deliberation

With support from the above grant in place, team members turned their attention to determining how best to expand the existing prototypes into applications that could meaningfully support access to Tsuut'ina language for local language programs. Since the Office of the Tsuut'ina Language Commissioner and its collaborators had recently been working on a number of substantial language resources for school-based programs that were available in digital format, it was recommended that these be prioritized for inclusion in an online dictionary and related tools. These resources included draft copies of a 'modernized' edition of the extensive list of elicited Tsuut'ina word-forms that linguist Edward Sapir and Tsuut'ina speaker John Whitney-Onespot developed together in 1922 (ca. 11,000 items; Whitney-Onespot & Sapir, 1922; Starlight et al., 2016), as well as two 100-page collections of Tsuut'ina verb paradigms (the Tsuut'ina Verb Phrase Dictionary, Books 1 and 2; cf. Starlight & Donovan, 2019). Through the contributions of Josh Holden during a postdoctoral fellowship at the University of Alberta, as well as Karoline Antonsen and Ruben Mögel, Master's students at the University of Alberta, work began to prepare to incorporate the information in these resources into a lexical database that would underlie all of these technical tools.

While this initial work on processing these materials got underway in the early days of the COVID-19 pandemic, regular video teleconference meetings were scheduled with all team members to discuss grammatical issues. For partners at Tsuut'ina Nation, these meetings often provided an opportunity to present and discuss issues around the interpretation of particular morphemes and constructions that had been encountered in recent language projects, with non-Tsuut'ina partners sharing comparisons with similar forms in other Dene languages and/or contributing to analysis together. For university-based partners, these sessions also provided opportunities to share regular updates about ongoing work on transferring information from the language resources into a database, as well as to seek advice on forms whose grammatical analysis or meaning seemed unclear or that needed to be confirmed by first-language Tsuut'ina speakers before being included in the computational morphological model. The mutual support and connection afforded by these meetings served important functions at the outset of the SSHRC Partnership Grant, contributing a sense of ongoing collaboration even when pandemic restrictions precluded any in-person gatherings.

As these regular meetings and efforts to incorporate existing language resources into a comprehensive lexical database moved forward, more of the characteristics of the latter materials' scope and coverage became apparent. By assigning each inflected Tsuut'ina verb word in these resources to a corresponding lemma, it became

possible to determine which lemmas contained information on all of the tenses/aspects/modes that are associated with regular Tsuut'ina verb phrases and which contained substantial gaps in documentation. It soon became apparent that, across the ca. 1,577 verbal lemmas attested in the Onespot-Sapir resource, the majority of lemmas showed at least one gap in a regular tense/aspect/mode form, with at least 700 having only a single tense/aspect/mode attested. A similar review of the verb phrase dictionary books revealed fewer missing paradigm forms, but brought attention to potential inconsistencies in tone marking in Tsuut'ina forms that eventually led the Office of the Tsuut'ina Language Commissioner to request that these preliminary books be set aside. While information from these analyzed resources would later prove valuable, regular project meetings underscored concerns over incomplete and potentially inaccurate information being circulated out of these provisional resources, as well as over the challenge of addressing such significant documentary gaps without systematic support from a much larger number of fluent Tsuut'ina speakers.

At the same time as these issues with the available language resources came to be discussed, other language initiatives at Tsuut'ina Nation continued to advance, including efforts to develop a new curriculum for use in core Tsuut'ina language programs at all age levels (i.e., Headstart, K–12, and adult education). This curriculum aimed to help reorient Tsuut'ina language learning and teaching from the noun-focused approaches that had generally been adopted in previous programs (e.g., beginning with teaching and learning lists of nouns at all age levels) to introducing and emphasizing verb phrases early on, recognizing how important verb-based patterns are in Dene languages like Tsuut'ina. With language learners and teachers being among the primary intended audiences for the tools being developed in this partnership, it was decided to set the previous language resources aside and attempt to align work on the online dictionary as closely as possible with the needs of curriculum users—that is, Tsuut'ina language educators, language learners, and curriculum team members. All partners recognized that the intelligent dictionary could be an essential resource to support this new curriculum, especially in its focus on verb paradigms. Encouragingly, this reorientation allowed members of the partnership team to draw on parts of the documentation analyzed in previous stages of this project to fill in portions of the vocabulary needed for the curriculum, thus saving time and effort. This work also brought attention to other gaps in existing documentation, this time in vocabulary related to both everyday activities and cultural practices that were either incompletely recorded in previous resources or entirely absent from prior documentation (e.g., specialized vocabulary related to hanging up meat on drying racks, pounding chokecherries, or other important cultural practices). Team members drew additional inspiration for curriculum vocabulary from several sources, including input from Tsuut'ina language teachers and advanced language learners and pedagogical resources developed for other Dene and non-Dene Indigenous languages. Connecting the development of the lexical contents of digital tools with the needs of Tsuut'ina language learners and teachers thus helped expose (and, in turn, contribute to addressing) significant gaps in the domain coverage of existing Tsuut'ina language materials in several high-priority areas.

Systematically addressing these gaps in lexical documentation—whether encountered in existing language resources or made apparent by requests from language teachers and members of curriculum development teams—and ensuring the accuracy of the information that would be represented in the tools developed in this partnership presented a standing challenge. This was addressed in part by the development of processes within Tsuut'ina Nation that sought to ensure that curriculum materials and other language resources reflected the understandings of fluent, first-language Tsuut'ina speakers. This involved the formation of the Tsuut'ina Language and Culture Committee, an advisory body consisting of six Tsuut'ina-speaking Elders that had within its mandate the review and approval of Tsuut'ina language resources prior to their use in the community. The establishment of a review process that supported nearly a third of the present-day first-language speakers of Tsuut'ina in gathering to offer constructive feedback on Tsuut'ina language matters proved important to addressing the above concerns over accuracy and coverage in language resources, with committee members often helping one another to recall less frequent Tsuut'ina terms and expressions that were previously in more active use. This new review

process required careful, item-by-item (or, in the case of sets of paradigmatically related forms, summarized paradigm by summarized paradigm; see Appendix A for an example) review to ensure that all members of the committee were in agreement that these language resources were acceptable for further use. This consensus-driven process of collectively reviewing lexical items one at a time, while requiring more time than previous attempts to incorporate existing language resources wholesale, helped not only to ensure that any inaccuracies or inadvertent gaps in Tsuut'ina forms or their English translations were systematically addressed, but also that the approved materials could be taken to reflect the collective understanding and priorities of the Tsuut'ina speech community, thereby fostering greater inclusion and a sense of collective ownership and investment in these collaboratively developed resources.[1]

4 Steps towards an intelligent online Tsuut'ina-English dictionary

As sets of verb phrases are identified for inclusion in the new Tsuut'ina language curriculum, members of the partnership team now have a comparatively straightforward workflow for ensuring that they are incorporated systematically in the online Tsuut'ina dictionary:

1. The second author and/or the Language and Culture Committee are consulted to recommend suitable Tsuut'ina equivalents, with the second author providing a brief overview of the regular aspectual forms.

2. These aspectual forms are added to a preliminary lexical database used to hold as-of-yet unapproved lexical items, then compiled as the lexical component of the current Tsuut'ina finite-state morphological model (`https://github.com/giellalt/lang-srs/`; Holden et al., 2022), producing a temporary finite-state transducer (FST) model (according to the Xerox-style specifications, cf. Beesley & Karttunen 2003; Hulden 2009).

3. The temporary FST is used to populate a Word document template, producing a condensed (1–2 page) overview of inflected forms for each of the regular tense/aspect/mood categories. A separate Python module developed by the partnership team also provides provisional English free translations for each Tsuut'ina word-form in this document, converting FST tag sequences and an English translation template sentence into contextually appropriate translations (e.g., `to_english("+V+I+Pfv+SbjSg1", "he/she/it will run") => "I ran"`, mapping the `+Pfv` perfective aspect tag to past inflection and the `+SbjSg1` first-person singular subject tag to "I" in English).

4. These automatically populated 'paradigm review sheets' are then reviewed and edited by the second author. Once that initial round of editing is complete, the first and second authors meet to review and record all of the Tsuut'ina word-forms together, producing high-quality WAV audio recordings of all entries in the paradigm review sheets that can later be incorporated as audio clips into the online dictionary.

5. On the basis of the paradigm review sheets emerging from this process, the preliminary lexical database is updated to reflect the corrected forms, then used again to produce an FST that is used to populate an overview of the recorded paradigms for the Language and Culture Committee to review. Any feedback from the committee members on this overview can then be incorporated and all lexical information moved into a permanent lexical database for approved material.

Importantly, the above review process is undertaken one verb lexeme at a time. That is, we ensure that the entire inflectional paradigm and all of its principal parts is fully validated for one

[1] The partnership has also provided financial support for the training of Tsuut'ina language instructors with Tsuut'ina-specific courses within the Community Linguist Certificate provided by the Canadian Indigenous Languages and Literary Institute (CILLDI) at the University of Alberta.

lexeme before we continue to the review of the next lexeme and its paradigm. Strict adherence to this ensures that no aspectual gaps are accidentally left in the paradigms (which would be facilitated by hopping from one lexeme to another in a less structured approach), nor do any mistranscriptions remain of individual word-forms in the paradigms. In this manner, we will be able to make available from the very onset an intelligent online dictionary for Tsuut'ina, with full features and functionality, e.g., the ability to (a) generate full and correct inflectional paradigms, (b) include full audio linked to all word-forms in these paradigms, as well as (c) recognize and linguistically analyzed each and every word-form in these paradigms—even if we can only implement this for a small set of verb lexemes, at least in the very beginning. This will allow for the informed examination of, and accurate feedback from various stakeholders for, a fully Tsuut'ina version of an intelligent online dictionary in terms of its linguistic content, rather than having to somehow explain (away) and go back to filling in missing sub-paradigms for some tense/aspect/mode, or correcting some incorrect word-forms in the paradigms. This resource is anticipated to gradually grow as more paradigms for verbal lexemes are individually created and reviewed. To date, this process has resulted in over 1,000 pages of completed paradigm review sheets for verb phrases requested for the Tsuut'ina language curriculum, with 45h18m of corresponding audio recordings of inflected word-forms. This work is still underway, with more material expected to be created again over the coming few months and committee review ongoing.

5 Lessons learned

In our experiences in this multi-year collaboration, we have noted positive outcomes from several practices that we have increasingly come to favour over time:

1. *Showing vs. telling:* In our initial conversations about this project, we found it valuable to be able to show what the tools we were discussing might look like, rather than simply talk about them in general terms. For new technologies with few precedents, or where the only precedents are currently available for unrelated languages, it can sometimes be difficult to picture what a particular tool or resource may look like in the target language and imagine how it might be useful. In this project, being able to share and discuss mock-ups of these tools, and later develop those into limited-but-working prototypes for preliminary evaluation, provided a valuable way forward for us in developing a common understanding of what we hoped to work towards.

2. *Change, communication, and responsiveness:* Since this partnership officially began in 2019, several project team members have transitioned into and out of key roles, new processes for the review of Tsuut'ina language materials have come into effect, and priorities have continued to shift for local language programs (here, in the direction of curriculum and teaching). Changes such as these have, at times, required significant deliberation to determine how best to proceed, including through extended pauses when members of the team needed to assess how these shifts might affect their planned contributions. Maintaining communication between project partners and deliberately expanding the circle of those involved—from an initially small group of community language leaders and university-based collaborators to a wider group of first-language speakers serving as reviewers, Tsuut'ina language teachers, and curriculum developers—has been crucial to seeing this project continue to develop, helping to ensure that it remains relevant in the context of local language work.

3. *Slow but steady wins the race:* In current work in developing collections of lexical material, computational morphological models, and related language technologies, it is not uncommon for breadth of lexical coverage and the rapid gathering of information to be presented as important features of useful, real-world resources and approaches (cf. Boerger

& Stutzman, 2018 on the motivations behind Rapid Word Collection methods). While this emphasis on broad-coverage and efficient lexicography is understandable, we would note here that initial attempts to draw on existing, relatively extensive language resources as a quick starting point were ultimately less successful than focusing on a much more restricted set of materials that were identified as the immediate needs of local language programs and processed in a slower, more deliberate manner. This approach has brought to light notable gaps between the vocabulary needed by current language education and revitalization programs and the outputs of previous generations of language documentation (cf. Mithun, 2007; Amery, 2009). We also find value in review processes that serve to build both consensus and community around such work, as is arguably the case here.

In this project, what began as a series of preliminary discussions among a small group of community and university-based language workers to see what could be possible to support Tsuut'ina language initiatives with new, digital tools has grown into a considerably broader partnership—one that now involves a much larger community of Tsuut'ina language teachers and curriculum developers, first-language speakers, and university-based researchers and students as key contributors. We look forward to seeing how this partnership continues to develop from here as these resources continue to grow—one verb phrase at a time.

Acknowledgments

We gratefully acknowledge the contributions made by many individuals who have been involved in this partnership, including Karoline Antonsen, Steven Crowchild, Josh Holden, and Ruben Mögel. This work has been funded by a Partnership Grant (895-2019-1012) from the Social Sciences and Humanities Research Council (SSHRC) of Canada.

References

Rob Amery. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation and Conservation* 3(2):138–148. http://hdl.handle.net/10125/4436.

Kenneth R. Beesley and Lauri Karttunen. 2003. *Finite State Morphology*. CSLI Publications, Stanford, CA.

Brenda H. Boerger & Verna Stutzman. 2018. Single-event Rapid Word Collection workshops: Efficient, effective, empowering. *Language Documentation & Conservation* 12:235–273. http://hdl.handle.net/10125/24766.

Calgary Roman Catholic Separate School Division. 1996. *Nanagusja: A Tsuut'ina (Sarcee) Language Development Program. Teacher's Guide*. Learning Resources Distributing Centre, Alberta Education. https://eric.ed.gov/?id=ED436084.

Eung-Do Cook. 1984. *A Sarcee grammar*. Vancouver, BC: University of British Columbia Press.

Pliny Earle Goddard. 1915. *Sarsi texts*. University of California Publications in American Archaeology and Ethnology 11, No. 3, pages 189–277. Berkeley, CA: University of California Press. http://digitalassets.lib.berkeley.edu/anthpubs/ucb/text/ucp011-004.pdf.

Joshua Holden, Christopher Cox & Antti Arppe. 2022. An expanded finite-state transducer for Tsuut'ina verbs. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, et al. (eds.), *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 5143–5152. Marseille, France: European Language Resources Association. https://aclanthology.org/2022.lrec-1.551.

Mans Hulden. 2009. Foma: a finite-state compiler and library. In *Proceedings of EACL*, pages 29—32, Athens: Association for Computational Linguistics.

Ryan Johnson, Lene Antonsen & Trond Trosterud. (2013). Using finite state transducers for making efficient reading comprehension dictionaries. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA) 2013*: 59–71.

Joyce McDonough, Jared O'Loughlin & Christopher Cox. 2013. An investigation of the three tone system in Tsuut'ina (Dene). In *Proceedings of Meetings on Acoustics* 19(1):060219. https://doi.org/10.1121/1.4800661.

Marianne Mithun. 2007. What is a language? Documentation for diverse and evolving audiences. *STUF – Sprachtypologie und*

Universalienforschung 60(1):42–55. https://doi.org/10.1524/stuf.2007.60.1.42.

John Onespot & Edward Sapir. 1922. *Sarsi (Tsuut'ina) notebooks #1–7*. Unpublished field notes, ms. American Council of Learned Societies Committee on Native American Languages. American Philosophical Society, Philadelphia, PA.

Keren Rice. 2000. *Morpheme order and semantic scope: Word formation in the Athapaskan verb* (Cambridge Studies in Linguistics 90). Cambridge: Cambridge University Press.

Keren Rice. 2020. *Morphology in Dene languages*. Oxford Research Encyclopedia of Linguistics. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.629.

Edward Sapir. 1925. Pitch accent in Sarcee, an Athabaskan language. *Journal de la Société des Américanistes de Paris* 17: 185–205.

Bruce Starlight, Gary Donovan & Christopher Cox. 2016. From archival sources to revitalization resources: Revisiting the Tsuut'ina notebooks of John Onespot and Edward Sapir. Paper given at the American Philosophical Society symposium 'Translating Across Time and Space: Endangered Languages, Cultural Revitalization, and the Work of History', Philadelphia, PA, October 13–15, 2016.

Bruce Starlight, Patrick Moore & Christopher Cox. 2018. *Documenting conversations in Tsuut'ina*. Endangered Languages Archive: http://hdl.handle.net/2196/00-0000-0000-0013-2FA3-9.

Bruce Starlight and Gary Donovan. 2019. *Tsuut'ina Verb Phrase Dictionary: Book One*. Calgary, AB: n.p.

Bruce Starlight & Christopher Cox. 2023. On vowel length contrasts in Tsuut'ina language work. Paper presented at the 2023 Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), online, January 20–22, 2023.

Bruce Starlight & Christopher Cox (eds.). 2024. *Isúh Áníi: Dátł'ìshí Ts'ìká áa Guunijà / As Grandmother Said: The Narratives of Bessie Meguinis*. Regina, SK: University of Regina Press.

Trond Trosterud. (2006). Grammatically based language technology for minority languages. In *Lesser-known languages of South Asia*, pages 293–316. De Gruyter Mouton, The Hague.

A Appendix A: Example verb paradigm summary

Table 1 presents a partial verb paradigm summary for the Tsuut'ina lemma *ànàyidiʔò* "he/she/it will lose it (solid obj.)", showing inflected Tsuut'ina word-forms associated with the Non-Past tense/aspect/mode category and their English free translations. In a complete paradigm summary for this lemma, similar tables would be included not only for the Non-Past, but also for the Past, Progressive, Repetitive, and Potential categories. The 13 verb forms shown in this table represent all possible subject person-number

Non-Past	**Ànàdisʔò.**	"I will lose it."
	Ànàdíʔò.	"You will lose it."
	Ànàyidiʔò.	"He/she/it will lose it."
	Ànàdaàʔò.	"We both will lose it (solid obj.)."
	Ànàdasʔò.	"You both will lose it (solid obj.)."
	Ànàgiyidiʔò.	"They both will lose it (solid obj.)."
	Ànàts'idiʔò.	"Someone will lose it (solid obj.)."
	Nominalized Verb Phrase	
	Ànàyidiʔò-hí	"the one who will lose it (solid obj.)"
	Ànàyidiʔò-hà	"the one that will lose it (solid obj.)"
	Distributive Plural	
	Ànàdàdaàʔò.	"Each and every one of us will lose it (solid obj.)."
	Ànàdàdasʔò.	"Each and every one of you will lose it (solid obj.)."
	Ànàdàgiyidiʔò.	"Each and every one of them will lose it (solid obj.)."
	Ànàdàts'idiʔò.	"Each and every one will lose it (solid obj.)."

Table 1: Partial verb paradigm summary for the Tsuut'ina lemma *ànàyidiʔò* "he/she/it will lose it (solid object)."

combinations in Tsuut'ina (including forms with distributive plural marking and two distinct forms of deverbal nominalization) when appearing with a third-person singular direct object. This limited set of forms is sufficient to determine both the inflectional paradigm to which this lexeme belongs as well as its constituent morphemes. Moreover, by holding the person and number of any object marking constant across all subject forms, summary charts such as this are able to concisely represent verbs that mark one or more objects morphologically, which may otherwise have several thousand distinct inflected forms.