

Language Shift, Language Technology, and Language Revitalization: Challenges and Possibilities for St. Lawrence Island Yupik

Lane Schwartz

University of Illinois at Urbana-Champaign

Department of Linguistics

lanes@illinois.edu

Abstract

St. Lawrence Island Yupik is a polysynthetic language indigenous to St. Lawrence Island, Alaska, and the Chukotka Peninsula of Russia. While the vast majority of St. Lawrence Islanders over the age of 40 are fluent L1 Yupik speakers, rapid language shift is underway among younger generations; language shift in Chukotka is even further advanced. This work presents a holistic proposal for language revitalization that takes into account numerous serious challenges, including the remote location of St. Lawrence Island and Chukotka, the high turnover rate among local teachers, socioeconomic challenges, and the lack of existing language learning materials.

Keywords: St. Lawrence Island Yupik, language technology, language revitalization

Piyuwhaaq

Sivuqam akuzipiga akuzitngi qerngughquteghllagluteng ayuqut, Sivuqametutlu pamanillu quteghllagmillu. Akuzipikayuget kiyang 40 year-eneng nuyekliigut, taawangiinaq sukallunteng allanun ulunun Iliighaqut nuteghatlu taghnughhaatlu; wataqaaghqaq pamanu quteghllagmi. Una qepghaqaghqaq aaptaquq ulum uutghutelleghqaaneng piyaqnaghngaan uglalghii ilalluku uyavantulanga Sivuqaamillu Quteghllagemllu, apeghutghistetlu mulungigatulangitnengllu, kiyaghtaallghemllu allangughnenganengllu, enkaam apeghutusipagitellghanengllu.

1. The Inuit-Yupik language family

The Inuit-Yupik-Unganam Tunuu languages constitute the northernmost language family in the Western Hemisphere. Unganam Tunuu is indigenous to far southwest Alaska and the Aleutian Islands; the remaining languages constitute the Inuit-Yupik language family. The eastern branch of the Inuit-Yupik family, the Inuit languages, represent a dialect continuum indigenous to the Arctic and near-Arctic coast of North America, encompassing Greenland and the north coast of Canada and Alaska. The western branch of the family, the Yupik languages, are indigenous to western Alaska and the Bering Strait region, including St. Lawrence Island, Alaska and the Chukotka Peninsula of Russia. The Inuit-Yupik-Unganam Tunuu language family is notable as the only language family in the world indigenous to both North America and Asia (see Figure 1). This work will focus on St. Lawrence Island Yupik, a variety indigenous to St. Lawrence Island, Alaska and parts of the Bering Sea coast of Chukotka, Russia. St. Lawrence Island Yupik is the only language within the Inuit-Yupik family spoken natively on both continents.

1.1. Terminology

The Inuit-Yupik language family has historically been called the Eskimo language family. Within Alaska, the term Eskimo has been used to encompass the Inuit peoples of northern Alaska and the Yupik peoples of western and southwestern Alaska. Outside of Alaska, especially in Canada, the term Eskimo is now considered derogatory. In any case, the term is an exonym, and the Inuit Circumpolar Council has requested that its use be discontinued (ICCR, 2010). However, the use of the term Inuit to refer to Yupik peoples and languages obscures the fact that there is a historical and linguistic distinction between Inuit and Yupik.

We therefore forgo the further use of the term Eskimo and instead use the term Inuit-Yupik to refer to the language family that encompasses the Inuit and Yupik languages, and the term Inuit-Yupik-Unganam (Fortescue et al., 2010) to refer to the broader family that also encompasses Unganam Tunuu (exonym: Aleut).¹

The term Central Siberian Yupik was proposed in the 20th century to refer to the language spoken on St. Lawrence Island and across the Bering Strait in villages including Ungaziq, Chukotka (Russian name: Chaplino). The modifier term “Central” was chosen based on the idea that this language was centrally located amongst the three Yupik languages spoken in Chukotka both geographically and in terms of language relatedness (Michael Krauss, P.C. 2017). Subsequent research has made a strong case that Sirenik (one of the three) should be considered to form its own branch of the Inuit-Yupik family, rather than a branch of Yupik (see Figure 1). The modifier term “Siberian” was chosen to contrast with the Alaskan varieties of Yupik, with the term Siberia historically used in Alaska to refer to Chukotka. However, within Russia the term Siberia generally refers to a much broader region that mostly includes territory far west of Chukotka. In the interest of accuracy and clarity, we therefore follow Schwartz et al. (2020) in arguing that the term Central Siberian Yupik should be replaced by the term St. Lawrence Island Yupik in English and the term Chaplinski Yupik in Russian (refer-

¹If a more concise name for the Inuit-Yupik-Unganam language family is at some point needed within the academic literature, the term *Iyut* could perhaps be proposed to the broader Inuit, Yupik, and Unganam communities. The term is an acronym formed from the language names (Inuit-Yupik-Unganam Tunuu) and has a surface form ending in *-t* which is consistent with how plural nouns are inflected in many of the languages in this family.

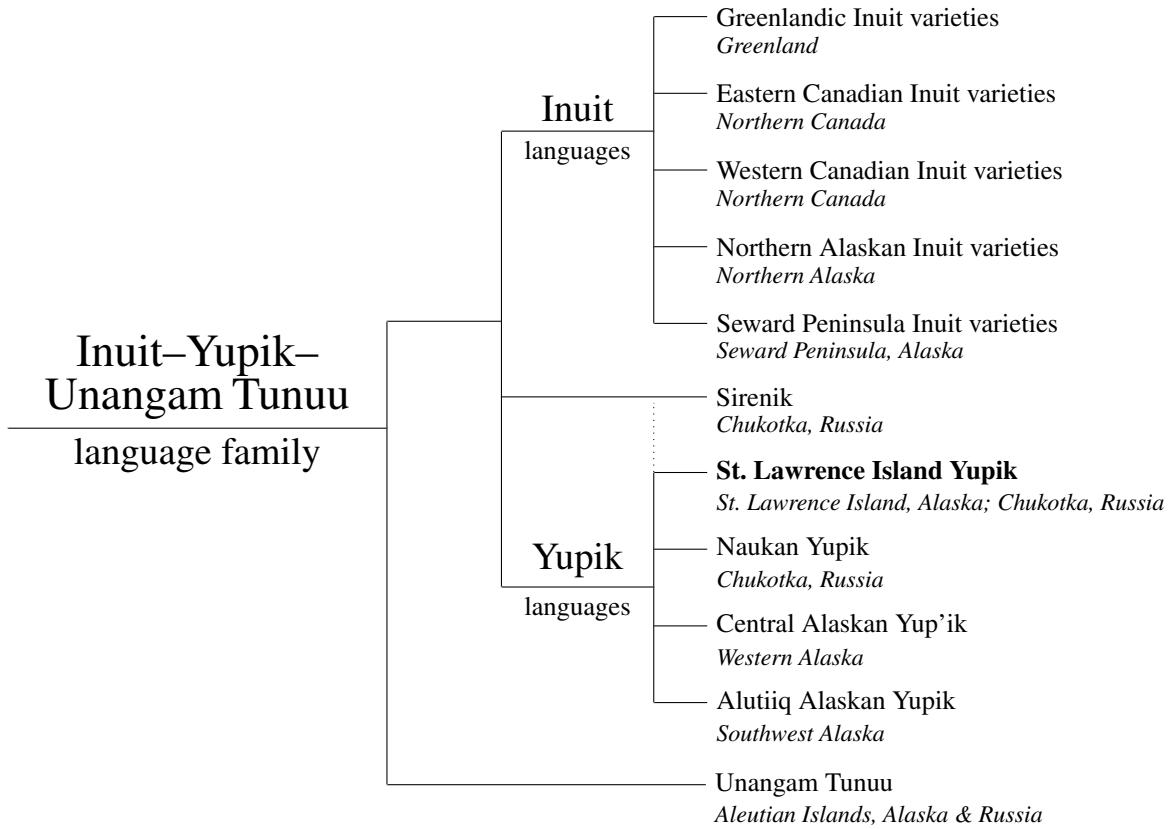


Figure 1: Inuit-Yupik-Unangam Tunuu language family (Fortescue et al., 2010; Krauss et al., 2011)

ring to Chaplino and New Chaplino, the respective Russian names of a prominent historical village and a contemporary Yupik village in Chukotka). The Yupik terms *Yupigestun* and *Akuzupik* are endonyms for this language, with specific terms such as *Sivuqaghdmistun* and *Ungazighdmistun* used to refer to the specific varieties spoken on St. Lawrence Island and Chukotka, respectively (Jacobson, 1990).

2. Status of St. Lawrence Island Yupik

St. Lawrence Island Yupik (ISO 639-3: *ess*) is a polysynthetic language indigenous to St. Lawrence Island, Alaska, and the Chukotka Peninsula of Russia. We use the term *Yupik* to refer to the language and to any individual Yupik person; we use the Yupik plural word *Yupiget* to refer to multiple Yupik persons. Schwartz et al. (2020) estimate 800–900 fully fluent L1 Yupik speakers out of an ethnic population of approximately 2400–2500 Yupiget. The majority of fluent Yupik speakers live in the villages of Gambell and Savoonga on St. Lawrence Island or in Chukotka, with smaller numbers residing in larger settlements, primarily Nome and Anchorage. While the vast majority of St. Lawrence Island Yupiget born in or prior to 1980 (individuals who as of late 2019 were 40 years old or older) are fluent L1 Yupik speakers (Krauss, 1980), rapid language shift is underway among younger generations; language shift among Yupiget in Chukotka is even further advanced (Morgounova, 2007). During linguistic fieldwork conducted over the period of 2016–2019, we observed very little Yupik use among Yupiget born prior to 1980, and widespread use of English among this age group.

2.1. Yupik Language in Education

Yupik language pedagogical materials were developed in the Soviet Union in the early 20th century (Krupnik and Chlenov, 2013) and in Alaska in the late 20th century (Koonooka, 2005). These materials were used in Russia into perhaps the 1950s (Krupnik and Chlenov, 2013) and on St. Lawrence Island until the early 2000s (Koonooka, 2005). The schools on St. Lawrence Island today have some basic Yupik language instruction, but very little of the pedagogical materials developed in the 1980s and 1990s are in use. The existing Yupik language books are archived at the Alaska Native Language Archive at the University of Alaska Fairbanks and the Materials Development Center in the Gambell School; these books are not easily accessible by members of the St. Lawrence Island community.

2.2. Challenges

The Yupik communities on St. Lawrence Island (and in Chukotka) are geographically very isolated. Like many indigenous communities, these communities have high rates of poverty, and many individuals within these communities struggle with various associated socioeconomic challenges. Within the school system, there is a chronic struggle to recruit and retain highly qualified teachers willing to serve these isolated communities. The certified teachers recruited to the schools are not Yupik speakers. A Yupik instructional curriculum was developed in the early 1990s, but it is in need of major pedagogical revision, and generally assumes a much higher level of Yupik proficiency among both instructors and students than is the case today. No pedagogi-

cal material designed to teach Yupik from scratch exists.

3. Yupik Language Technology

The polysynthetic nature of Yupik means that Yupik words are commonly composed of multiple morphemes. This means that the development of even simple language technologies such as spell-checkers and electronic dictionaries is substantially more complicated than for isolating or analytic languages. The first known publicly released Yupik language technology was released by Schwartz and Chen (2017) and included a simple web-based rule-based spell-checker that ensured that each word conformed to Yupik phonotactic and orthotactic requirements, as well as transliteration utilities allowing user-entered Yupik text to be transliterated into the Cyrillic orthography used in Chukotka, a Latin orthography used on St. Lawrence Island, a fully transparent alternative Latin orthography designed for pedagogical use, the Americanist phonetic notation used by several existing Yupik linguistic works, and the International Phonetic Alphabet notation. Chen and Schwartz (2018) implemented a prototype finite-state morphological analyzer for Yupik capable of decomposing a Yupik word into its constituent morphemes. Hunt et al. (2019) implemented a prototype web-based electronic dictionary for Yupik that incorporated a Javascript version of the Chen and Schwartz (2018) finite-state morphological analyzer. Schwartz et al. (2019a) explored the viability of using recurrent neural network methods to learn a more generalized morphological analyzer from the finite-state analyzer of Chen and Schwartz (2018). During the summer of 2019, a six-week summer research workshop (Schwartz et al., 2019c) explored this issue in substantially more depth, including the development of a language model prototype for use in text completion on mobile devices.

4. Goals and Proposals for Language Yupik Language Education and Revitalization

Over the course of our linguistic fieldwork conducted in Gambell between 2016 and 2019, a consistent theme voiced by many members of the Yupik community has been a desire for substantially strengthened Yupik instruction in the schools, ideally in the form of a Yupik language immersion program. This desire was communicated directly by the tribal council of the Native Village of Gambell, the elected representatives of the Yupik people in Gambell, during multiple meetings with the author of this work over the course of three years, as well as in conversations between this author and numerous members of the community. In addition to a strong Yupik-language educational program in the St. Lawrence Island schools, an additional goal stated by some members of the community is the establishment of self-study materials for use by younger Yupiget who are partially fluent, passively fluent, or not fluent in Yupik. Such materials would be especially helpful for younger adults of child-rearing age who are not themselves fluent Yupik speakers, but who have a desire to speak and transmit Yupik to their own children.

Current efforts in support of Yupik language education and revitalization efforts include the recent creation of a community-organized language committee on

St. Lawrence Island (the Kaalguq Committee), language technology development led by the author of this work (see §3.), Yupik language documentation efforts led by Dr. Sylvia Schreiner of George Mason University and the author of this work, and ongoing efforts (begun in 2016) to digitize all printed Yupik-language materials into accessible PDF and plain-text formats.

Continued leadership from members of the St. Lawrence Island Yupik community is critical for the development and success of formal language revitalization efforts, as is consultation with stakeholders in the St. Lawrence Island community, including tribal and local governments, the Bering Strait School District, and the Alaska Native Language Center. The proposals in the following subsections (§4.1.–§4.3.) represent one possible forward path toward the long-term goal of a robust Yupik-language immersion program in the St. Lawrence Island schools.

4.1. Mobile-friendly digital Yupik library

During the 1970s through the early 1990s, a substantial number of Yupik-language materials were developed for educational use by the Nome Schools, the Bureau of Indian Affairs, the Alaska Native Language Center, and the Bering Strait School District. Since 2016, over ninety such books have been identified and scanned. This includes pre-primers for use in pre-school and early elementary settings, a set of three mid-elementary readers, and a three-volume set of stories by Yupik elders appropriate for use by advanced Yupik-language students.

Current efforts are underway to digitize all of these books into plain-text formats, and to obtain high-quality audio recordings of as many as possible. Once in plain-text format, each book (with associated audio if available) can be converted into an interactive e-book in e-pub format (Schwartz et al., 2019b). These e-books could then be gathered into a mobile-friendly collection, which community members could download for offline access when mobile internet services are active. Mobile phones are in widespread use on St. Lawrence Island, although mobile data access is relatively unreliable. By prioritizing books with the lowest reading level, this collection of audio e-books could be utilized by instructors in the St. Lawrence Island schools, as well as by adult language learners and their children at home.

4.2. Development of pedagogical materials

There are currently no Yupik self-study materials, no materials designed to teach Yupik from scratch in an educational setting, and no Yupik-language pedagogical materials for teaching other subjects (such as mathematics) using Yupik as the language of instruction. These facts represent major challenges to the eventual establishment of a Yupik language immersion school.

A medium-term goal, then, is the development of materials for use in learning Yupik. Ideally, such materials should be designed to be dual-use whenever possible, in order to support use in the St. Lawrence Island schools as well as use by adults in a self-study scenario. Materials should be designed keeping in mind that Yupik instruction in the

schools could be led by Yupik-dominant or English dominant teaching aides who may lack formal training, or potentially by (English-speaking) certified teachers from outside St. Lawrence Island who have formal training but not in language instruction.

Language technology could potentially be used to assist in the development of such pedagogical materials. Machine translation and translation memory technologies trained for translation from English (or from a different Inuit-Yupik language) into Yupik could potentially be used in a computer-aided translation scenario in which existing subject-matter textbooks are translated into Yupik. Yupik-language spell-checking and grammar modelling technology could also play a role.

4.3. Immersion programs

The academic literature on language revitalization (Hinton and Hale, 2001; Hinton, 2013) regarding best practices includes three related techniques that have been shown to be successful in creating new speakers of endangered languages: language nests (Hinton, 2013), language immersion programs (Hale, 2001), and master-apprentice programs (Hinton, 2001). In language nests, fluent adult speakers work with very young children in child-care or preschool-like settings. In school immersion programs, the immersion language is spoken by fluent instructors as the language of instruction in schools. In master-apprentice programs, an adult language learner or learners interact with a (usually elder) fluent speaker of the language in an immersion setting over an intense 2-3 year period, typically for 20–40 hours per week.

The establishment of a Yupik language immersion program in the St. Lawrence Island schools is a stated goal of many in the St. Lawrence Island community. Achieving this long-term goal will require coordination and buy-in from the Bering Strait School District, in addition to the development of Yupik-language subject-matter textbooks. An equally important requirement is the availability of instructors who can teach using Yupik as the language of instruction.

While it is theoretically possible that a sufficiently well-motivated student could learn Yupik through self-study (assuming that such self-study materials are first developed), the most promising path to the development of Yupik-fluent instructors is the establishment of a master-apprentice program. Full-time master-apprentice programs in which the master and the apprentices are paid a living wage for their time have been shown to be effective in developing new language speakers in a relatively short period of time (2–3 years). Community-led efforts to seek grant funding could provide one possible source for the establishment of such a master-apprentice program. Innovative partnerships between the Bering Strait School District, the St. Lawrence Island tribal governments, and the Alaska Native Language Center at the University of Alaska could also be explored, in order to create a novel program which could provide Yupik proficiency through a master-apprentice program while simultaneously providing students with some form of (possibly distance-based) teacher training.

5. Acknowledgements

I want to offer my immense thanks to the people of St. Lawrence Island for graciously sharing their language with me. The St. Lawrence Island Yupik language is a critical part of the cultural heritage of the St. Lawrence Island Yupik people. It has been and continues to be an immense honor to be welcomed into the St. Lawrence Island community, and to be trusted to engage in this work. I want to offer my great thanks to Petuwaq and all of the Yupik speakers who have worked with my colleagues and I since 2016. I want to offer my deep gratitude to the people of St. Lawrence Island, who first welcomed my family in 1982, and have generously continued to do so in the decades since. I want to thank the board members and staff of the Native Village of Gambell, the City of Gambell, and Sivuqaq, Inc, as well as the faculty and staff of Gambell Schools and the staff of Gambell Lodge.

I want to acknowledge the National Science Foundation and its Documenting Endangered Languages program, which has generously supported this research under NSF Award 1761680.

Finally, I want to offer my profound thanks to Willem de Reuse, Michael Krauss, and especially to Steve Jacobson for their decades-long work documenting the Yupik language. None of my work would have been possible without the outstanding grammar of Yupik that Dr. Jacobson developed over the course of many years working with the St. Lawrence Island community, as well as the incredible Yupik dictionary that he edited.

Igamsiqanaghhalek!

6. Bibliographical References

Chen, E. and Schwartz, L. (2018). A morphological analyzer for St. Lawrence Island / Central Siberian Yupik. In *Proceedings of the 11th Language Resources and Evaluation Conference (LREC'18)*, Miyazaki, Japan, May.

Fortescue, M., Jacobson, S., and Kaplan, L. (2010). *Comparative Eskimo Dictionary with Aleut Cognates*. Alaska Native Language Center, Fairbanks, Alaska, 2nd edition.

Hale, K. (2001). Linguistic aspects of language teaching and learning in immersion contexts. In Leanne Hinton et al., editors, *The Green Book of Language Revitalization in Practice*, page 227–235. Brill.

Leanne Hinton et al., editors. (2001). *The Green Book of Language Revitalization in Practice*. Brill.

Hinton, L. (2001). The Master-Apprentice language learning program. In Leanne Hinton et al., editors, *The Green Book of Language Revitalization in Practice*, page 217–226. Brill.

Leanne Hinton, editor. (2013). *Bringing our Languages Home: Language Revitalization for Families*. Heyday, Berkeley, California.

Hunt, B., Chen, E., Schreiner, S. L., and Schwartz, L. (2019). Community lexical access for an endangered polysynthetic language: An electronic dictionary for St. Lawrence Island Yupik. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*,

pages 122–126, Minneapolis, Minnesota, June. Association for Computational Linguistics.

ICCR. (2010). Inuit circumpolar council resolution 2010–01 on the use of the term inuit in scientific and other circles.

Jacobson, S. A. (1990). *A Practical Grammar of the St. Lawrence Island/Siberian Yupik Eskimo Language, preliminary edition*. Alaska Native Language Center, Fairbanks, Alaska.

Koonooka, (Petuwaq), C. (2005). Yupik language instruction in Gambell (St. Lawrence Island, Alaska). *Études/Inuit/Studies*, 29(1/2):251–266.

Krauss, M., Holton, G., Kerr, J., and West, C. T. (2011). Indigenous peoples and languages of Alaska. ANLC Identifier G961K2010.

Krauss, M. (1980). Alaska Native languages: Past, present and future. *ANLC Research Papers*, 4.

Krupnik, I. and Chlenov, M. (2013). *Yupik Transitions — Change and Survival at Bering Strait, 1900-1960*. University of Alaska Press, Fairbanks, Alaska.

Morgounova, D. (2007). Language, identities and ideologies of the past and present Chukotka. *Études/Inuit/Studies*, 31(1-2):183–200.

Schwartz, L. and Chen, E. (2017). Liinnaqumalghiit: A web-based tool for addressing orthographic transparency in St. Lawrence Island/Central Siberian Yupik. *Language Documentation and Conservation*, 11:275–288, September.

Schwartz, L., Chen, E., Hunt, B., and Schreiner, S. L. (2019a). Bootstrapping a neural morphological analyzer for St. Lawrence Island Yupik from a finite-state transducer. In *Proceedings of the 3rd Workshop on the Use of Computational Methods in the Study of Endangered Languages Volume 1 (Papers)*, pages 87–96, Honolulu, February. Association for Computational Linguistics.

Schwartz, L., Schreiner, S. L., Zukerman, P., Soldati, G. M., Chen, E., and Hunt, B. (2019b). Initiating a tool-building infrastructure for the use of the St. Lawrence Island Yupik language community. International Year of Indigenous Languages 2019: Perspectives Conference, October.

Schwartz, L., Tyers, F., Park, H., Steimel, K., Strunk, L., Haley, C., Zhang, K., Kirov, C., Knowles, R., Levin, L., Littell, P., Lo, J., Prud'hommeaux, E., and Jimmerson, R. (2019c). 2019 JSALT Workshop on Neural Polysynthetic Language Modelling, June-August.

Schwartz, L., Schreiner, S., and Chen, E. (2020). Community-focused language documentation in support of language education and revitalization for St. Lawrence Island Yupik. *Études/Inuit/Studies*, In press.

Schwartz, L. (2018). NNA: Collaborative research: Integrating language documentation and computational tools for Yupik, an Alaska Native language. U.S. National Science Foundation Award 1761680.