UNIVERSITY OF ALBERTA

Elicitation and Speech Acts in the Maskwacîs Spoken Cree Dictionary Project

BY

Tanzi Reule

A THESIS
SUBMITTED TO THE FACULTY OF ARTS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF ARTS

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

June, 2018

UNIVERSITY OF ALBERTA

FACULTY OF ARTS

The undersigned certify that they have read and recommend to the Faculty of Arts for acceptance, a thesis entitled “Elicitation and Speech Acts in the Maskwacîs Spoken Cree Dictionary Project”, submitted by Tanzi Reule in partial fulfillment of the requirements for the degree of Bachelor of Arts.

….....

Honours Thesis Supervisor

….....

For the Department

ABSTRACT

The Maskwacîs Spoken Cree Dictionary Project is a collaboration between Miyo Wahkotowin Education and the University of Alberta's ALTLab. I describe how the project adapted the Rapid Word Collection Method to supplement the existing Maskwacîs Cree Dictionary and create recordings of lexical items and sentences. Treating the elicitation genre as its own genre of communication, I attempt to capture the elements of interaction and negotiation by creating an elicitation speech act taxonomy from a detailed annotation of a single elicitation session. I identify typical and non-typical task sequences and verify the sequences with samples from additional elicitation sessions. The report illustrates the regularity and uniqueness of the elicitation analysis from a discourse perspective and serves as a practical case study for researchers undertaking related projects.

ACKNOWLEDGEMENTS

My most sincere thanks to my supervisor Antti Arppe, M.Sc., Ph.D., at the University of Alberta for his patience, guidance, and detailed insight into the Maskwacîs Spoken Cree Dictionary Project.

My gratitude to Jordan Lachler, Ph.D., Atticus Harrigan, B.A. (Hons), and Katherine Schmirler, M.A., for their helpful advice and information.

Finally, many thanks to the participants of the Maskwacîs Spoken Cree Dictionary Project and Miyo Wahkotowin Education for allowing me to sit in on some recording sessions.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER I – INTRODUCTION

This report has two parts. The first is a description of the implementation of a semantically-organized elicitation methodology, the Rapid Word Collection Method, in the Maskwacîs Spoken Cree Dictionary Project, conducted by the University of Alberta in conjunction with Miyo Wahkotowin Education of the Ermineskin Cree Nation. Ordinarily used as an accelerated means of collecting lexical items, the Rapid Word Collection Method was used to supplement an existing Plains Cree dictionary in conjunction with an audio recording project. Secondly, in answer to the question "what exactly happens during an elicitation session?", I examine the elicitation activity as a unique genre of communicative event in which participants collaborate and negotiate as they produce lexical items and example sentences. I present a taxonomy of elicitation speech acts and identify a prototypical discourse sequence from elicitation request to the production of the item to orthographical clarification.

CHAPTER II – LITERATURE REIVEW - ELICITATION

2.1 Introduction.

This literature review will highlight relevant literature regarding the typological distinctions between communicative events gathered as data by documentary linguists, their merits, and how common elicitation methodologies relate to these data types. Although there is ongoing discussion concerning the distinction or interrelation between documentation and description in the context of endangered languages, for the purposes of this review they will be treated as discrete disciplines. This review will focus on data-gathering for documentation and will not explore ancillary issues; i.e., the role of theory in language document, collaboration, technology, or ethics.

2.2 Data types in language documentation.

As it aims to produce a "comprehensive record of the linguistic practices characteristic of a given speech community", language documentation is concerned with *primary* or *raw* data consisting of two types, observable linguistic behaviour and metalinguistic knowledge (Himmelmann, 1998, 2002; Austin & Grenoble, 2007).

In attempting to define the range of potential linguistic data, language documentarians frequently refer to Himmelmann's typologies of spontaneity (see Figure 1) and naturalness (see Figure 2). Some have noted, however, that these two continua may not be sufficient to describe all types of communicative events. Austin and Grenoble (2007) illustrate this problem with the example of a dialogical narrative, where the storyteller is frequently interrupted by the audience.

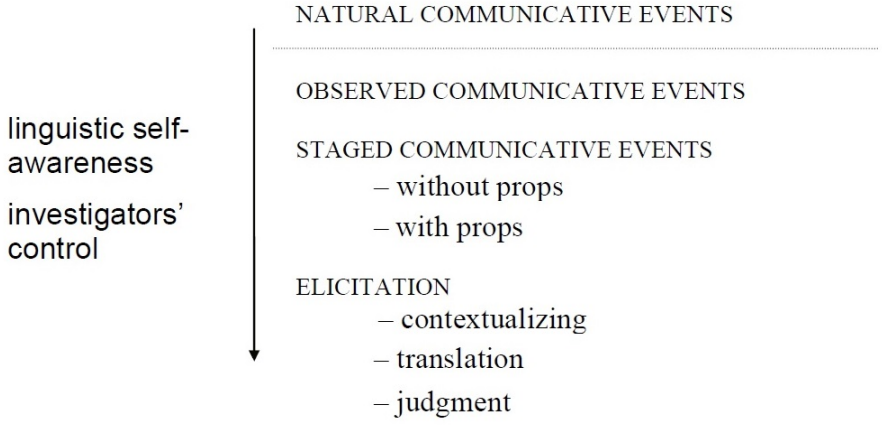
Indeed, the project that is the subject of this paper is difficult to classify according to these criteria alone, as will be discussed further below.

Looking outside the field of language documentation to language data collection in general, Gilquin and Gries (2008) proposed a more granular typology that distinguishes levels of naturalness between elicited fieldwork data, experimentation using usual language behaviour and language units, and experimentation using unusual language behaviour and units. Arppe et al. (2010) point out that, while linguistic judgments tasks are always described as unnatural, making language judgments is as necessary and natural as language production.

Figure 1: *Typology of spontaneity in communicative events (Austin & Grenoble, 2007).*

Parameter	Major Types	Examples
unplanned	exclamative	'ouch!', 'fire!'
⇕	directive	'scalpel!'
	conversational	greetings, small talk, chat, discussion, interview
	monological	narrative, description, speech, formal address
planned	ritual	litany

Figure 2: *Typology of naturalness in communicative events (Himmelmann, 2002).*



As data from all across the naturalness continuum began to be seen as a necessary part of language documentation, there has been some misconception that (observed) natural communicative events should be the only source of primary data (Himmelmann, 2012). This may have been a reaction to a too-heavy reliance on planned and elicited primary data in the early days of documentation. In spite of this misconception, most current sources agree that the greatest possible diversity of data types, methods, and participants is most valuable (Woodbury, 2003). Simply put, if both observable linguistic behaviour and speakers' metalinguistic knowledge are required for complete documentation (Berge 2010; Woodbury 2003), then both recordings of natural data and direct elicitation are crucial tools for researchers.

2.3 Elicitation methods and data types.

Due to the Observer's Paradox (Labov, 1972), the most natural communicative event (the unobserved event) is outside the reach of researchers. Opportunistically recorded unprompted communicative events are the most natural language data source available and will provide the best picture of current actual usage. A corpus of natural language data is for some documentarians replacing the traditional collection of narrative texts as the preferred source of primary language data. Creating a corpus is time-consuming, and non-frequent structures may be difficult to capture. It may also be difficult to ensure that all the data collected in these circumstances is data the participants will be comfortable publishing, particularly if the researcher is not fluent in the language.

Staged communication events, which often make use of prompts like map tasks (Brown et al., 1984) or the Frog Story (Mayer, 1969) or Pear Story (Chafe, 1980) have been used as a

means of generating discourse data. The added investigator control may be particularly useful in accessing specific higher-level features that are of interest (Caldecott & Koch, 2014).

As language documentation has grown as a discipline, the trio of a grammar, dictionary, and collection of narrative texts has been recognized as insufficient for complete documentation, but they nonetheless remain important as a foundation for further documentation and description (Berge 2010). Oftentimes, these resources cannot be compiled without the use of direct elicitation (Berge 2010; Mosel 2006). Despite the potential drawbacks, including the difficulty of controlling for bias; a participant's perception of pressure to "invent" suitable examples; the effects of repetition and fatigue (Himmelmann, 2012), direct elicitation is acknowledged as one of the most effective ways of collecting information about metalinguistic knowledge (Himmelmann, 2006). Again, staged communicative events and direct elicitation may be the only way of collecting data about infrequent structures (see Arppe & Järvik, 2007, for a discussion of the lack of universal correlation between frequency and grammatical acceptability.)

If there is any use of direct elicitation about which there is a lack of consensus, it may be using direct elicitation to collect lexical items. Ultimately, this may be less of an issue of conflicting theories than a conflict between theory and practice. "Corpus observation," or the recording of natural communicative events, is frequently implied to be the preferred method of collecting lexical items (Himmelman, 2006; Woodbury, 2003), and yet many authors acknowledge the practical function of a simple direct elicitation tool like a wordlist (Berge 2010; Crowley, 2007; Mosel, 2006). This practical benefit is recognized even "when a researcher has an acknowledged belief that language should be studied in context" (Berge, 2010).

2.4 Conclusion.

Here we arrive at another point of widespread agreement: complete documentation is of course an ideal. When facing the "grim realities of constrained resources" (Himmelman, 2012), each documentation project is an inevitable balancing act of collaborative goals, theoretical questions, and means. Much has been done to systematize documentation methodologies and data typologies, but the uniqueness of language communities and documentation projects means that not every method will be reflected in these typologies.

CHAPTER III. MASKWACÎS SPOKEN CREE DICTIONARY PROJECT

3.1 Origin of the Project.

In 2014, the University of Alberta's Alberta Language and Technology Laboratory (ALTLab) developed a partnership with Miyo Wahkotowin Education of the Ermineskin First Nation, located in Maskwacîs, Alberta, to develop a web-based electronic dictionary of Plains Cree.

Community elders had previously produced the Maskwacîs Cree Dictionary, which contained 8,986 words. Originally, the goals of the Maskwacîs Spoken Cree Dictionary Project were to

i) record careful, isolated pronunciations of each word in the Maskwacîs Cree Dictionary,

ii) record an example sentence for each dictionary word,

iii) record the informal Cree and English discussions concerning the words,

iv) transcribe the recordings and create a comprehensive text collection of Plains Cree, and

v) make these spoken words and sentences publicly available as part of a web-based electronic dictionary (Arppe, 2016).

Miyo Wahkotowin Education was willing to contract and reimburse local native speakers for their participation in the project.

3.2 Expansion of the Project.

The scope of the Maskwacîs Spoken Cree Dictionary Project (MSCDP) was subsequently expanded to include the elicitation of new words in order to supplement the existing Maswacîs Cree Dictionary. The supplementation was determined to be a valuable addition to the project since it could utilize the technical set-up already contemplated for the recording project. Also, since one might assume that the most frequent words were included in the original dictionary, the researchers hoped a systematic elicitation process might capture less frequent or easily-accessible words.

In order to proceed with the elicitation, an ontology needed to be chosen to organize the existing data and systematically elicit new data. Recording and eliciting data in alphabetical order would not create coherent sets of words. A semantically organized ontology, having the benefit of easy association of ideas, was preferred. Using semantic domains in elicitation is also supported by documentary linguists such as Crowley (2007) and Grimes (2002). The project chose to use Ronald Moe's Rapid Word Collection Method, materials for which are available from rapidwords.net.

3.3 The Rapid Word Collection Method

Background. After decades of experience in lexicography and documentation of endangered languages, Ronald Moe of SIL International was familiar with the technique of organizing dictionaries by semantic domains or "areas of meaning" (Moe, 2003). In 2004, he identified the benefits of semantic organization as

i) allowing for user ease in locating a desired word or meaning,

ii) facilitating research of the semantics and pragmatics of lexical sets, and

iii) giving translators the ability to compare and contrast words in a domain.

As part of the Bantu Initiative, Moe was tasked with producing a list of semantic domains that could be used to classify Bantu language dictionaries. While investigating how native speakers of languages in the Bantu family self-categorized words in their languages, Moe was struck by similarities to English lists of semantic domains he had collected from around the world. Based on this finding, Moe began work on a universal list of semantic domains that he believed could be used as an elicitation tool, a dictionary classification tool, and an aide to semantic investigation (Moe, 2003; 2004).

The result is an ontology of 1,800 semantic domains organized into 9 broad domains. The 9 broad domains contain cascading series of subdomains. These domains and subdomains are populated with items connected to the head domain by a lexical relation (see Table 1).

Table 1: *Examples of lexical relations*

Head Domain	**Lexical Relation**	**Items**
'game'	generic-specific (hypernym-hyponym)	*chess, checkers, charades, monopoly*
'head'	whole-part (holonym-meronym)	*eye, nose, mouth*

To improve the functionality of the list of domains as an elicitation tool, Moe added: (1) a simple statement of the central idea of each domain, (2) elicitation questions to act as prompts, and (3) sample English words. For example,

"What words refer to singing? *sing, serenade, warble, yodel, burst into song*

What words refer to singing without using words? *hum, whistle"* (Moe, 2003, p. 220)

(See Appendix A for further samples of domain organization and elicitation questions).

In terms of universality, Moe theorized that this ontology corresponds to the organization of the mental lexicon; i.e., that "the words of a language are all linked together in the mind in a gigantic multidimensional web of relationships [that] cluster around a central nexus" (Moe, 2003). He noticed that many of his domains are based on semantic primitives or a combination of semantic primitives and suggested this might be evidence in favour of their existence, while acknowledging this theory is not yet widely accepted (Moe, 2003).

Suggested implementation. From his experience producing dictionaries for minority languages, Moe estimated that, on average, words are collected at a rate of 650 words per year, or 2.5 words per working day (Moe, 2003). Utilizing his list of universal semantic domains, Moe developed the Rapid Word Collection Method (RWCM) as a means of efficiently eliciting lexical data from native speakers.

The following is a summary of the suggested implementation for the RWCM. These notes on implementation are available at rapidwords.net.

The RWCM is a word-collection workshop designed to collect at least 10,000 over approximately four weeks (including one week of preparation and training, two weeks of data collection, and an additional week of data entry and error correction).

One important preparatory step is to translate the English list of semantic domains into the relevant language of wider communication (LWC) if it is not already available. A minimum of 35 participants are divided into up to six teams, composed of:

i) a team leader, who also interprets the semantic domain questions to team members;

ii) language experts, who provide the words;

iii) scribes, who write the words down as they are provided by team members;

iv) glossers, who add glosses in the LWC; and

v) typists, who enter the words into the computer.

Three of the participants take on coordinating, logistics, and record-keeping roles. By the end of the error correction week, draft copies may be printed for the community, and the database should be made available online within 30 days.

One workshop held in Senegal in 2017 reported a collection of 12,485 words in 11 days (rapidwords.net, 2018). Moe (2003) reports workshops of 15 to 30 speakers collecting between 10,000 and 17,000 words in 8 to 10 days.

Moe lists a number of prerequisites for using the RWCM:

i) a motivated language community,

ii) a functional orthography,

iii) a sufficient number of literate participants,

iv) a sufficient number of bilingual participants, and

v) a workshop coordinator.

Rapid Words and the MSCDP. Whereas the RWCM is designed for undocumented languages where a dictionary must be created from scratch, the elicitation target of the MSCDP was to supplement an existing Plains Cree dictionary. The project had several other goals beyond

the scope of the RWCM: (1) to record careful repetitions, (2) to elicit example sentences, and (3) to create a web-based, online dictionary. Instead of a two-week workshop with 30 speakers, weekly sessions were held over a period of 4 years with three speakers per session.

Technically, the MSCDP fulfilled all of Moe's prerequisites: the community was motivated enough to hire native speakers (all bilingual and literate) to participate, a functional orthography already existed, and ALTLab and Miyo Wahkotowin were prepared to act as project coordinators.

Compiling Existing Maskwacîs Data. The first step in adapting the RWCM was to populate the Rapid Words ontology with the entries from the Maskwacîs Cree Dictionary (see Table 2).

Table 2: *Rapid Word domains populated by original Maskwacîs dictionary items*

Rapid Word Domain	**Maskwacîs Dictionary Items**
1. Universe, creation	863
2. Person	1124
3. Language and thought	889
4. Social behaviour	664
5. Daily life	937
6. Work and occupation	744
7. Physical actions	2256
8. States	1323
9. Grammar	164

Final methodology. Once a week, elicitors from the University of Alberta would travel to Maskwacîs to conduct recording sessions. Two sessions of two hours would take place in a day. One to two researchers from the University of Alberta would act as the elicitor and monitor the recording. Three to four native speakers would participate in the language tasks. If another set of speakers was available, two simultaneous sessions would be conducted. Language tasks included:

i) confirming the orthographic representation and gloss of the original Maskwacîs entries,

ii) working through the RWCM questionnaire sheets and providing new words,

iii) providing example sentences, and

iv) providing a final clear double repetition of all items.

All native speakers were individually mic'd and the entire session recorded on individual tracks. After the recording session, the files were annotated in Elan for Cree word/sentence, its English gloss, and any spontaneous Cree speech.

Results. As of April 25, 2018, 15,130 new words and 2,267 example sentences have been collected (see Table 3). Given the size of the existing dictionary, it is possible that many of the most familiar or frequent words were collected in the original Maskwacîs dictionary.

Table 3: *New words and sentences by domain*

Rapid Word Domain	**New Words**	**New Sentences**
1. Universe, creation	1129	420
2. Person	1858	75
3. Language and thought	1785	287
4. Social behaviour	1396	263

5. Daily life	1232	84
6. Work and occupation	1339	91
7. Physical actions	2052	232
8. States	1506	253
9. Grammar	244	52

Perspectives from the elicitors. I conducted two written interviews with University of Alberta researchers, Atticus Harrigan and Katie Schmirler. While both researchers found the semantic organization to be helpful, repetition of terms and culturally inappropriate or irrelevant terms were a source of difficulty. They both recommended editing the elicitation questions and prompts in advance for a long-term project like the MSCDP (Harrigan, 2018; Schmirler, 2018).

CHAPTER IV. ELICITATION SESSION DISCOURSE ANALYSIS

4.1 Introduction.

In this section, I will discuss the proposition that lexical and grammatical elicitation creates a new genre of communicative event that exhibits a regularized pattern of speech acts. I will identify the reasons a genre-specific taxonomy is necessary to describe these speech acts and outline my MSCDP speech act taxonomy. Finally, I will present the regularized speech act sequence developed from an in-depth analysis of one Maskwacîs elicitation session and contrast that with examples from sessions led by a variety of elicitors.

4.2 Elicitation communicative events.

Whereas observable linguistic behaviour (high on the naturalness continuum) consists of communicative events between native speakers that might be expected to occur outside of the instance a linguist records them, elicitation creates a new genre of communicative event between the elicitor and the native speaker (Himmelmann, 2006). Indeed, some argue that each area of human activity includes a distinct genre of speech (Bakhtin, 1986).

The elicitation communicative event is interactive because there are multiple interactants and collaborative by virtue of the fact that all interactants contribute turns and words (Wilkes-Gibbes, 1986 in Davies 2006). Theoretically, an elicitation session could be used to test established dialogic principles, as in Davies' (2006) map task experiment, such as:

Table 4: *Select hypotheses from Davies (2006) based on dialogic principles.*

Principle	Hypotheses
Gricean Cooperation	i) speakers will avoid unnecessary effort ii) speakers will improve at tasks iii) speaker effort will decrease
Principle of Least Collaborative Effort	i) dialogues will get shorter the more times participants do the task ii) there will be a decrease in average effort for later dialogues iii) speakers with equal commitment should be associated with more task success
Principle of Least Individual Effort	i) risks (of a failure of communication because of too-little effort) would be taken ii) risks would decrease over time iii) task success would improve as speakers negotiate trade-off more successfully over time

While these hypotheses could in theory apply to the task of elicitation, there are some important impediments regarding evaluations of risk and success. A map task has a clear binary measure of success in that participants successfully reach the destination or do not. In an elicitation session, a requested structure may not be produced because none exists or because participants could not recall it. And while of course some structures may be categorically ungrammatical or inaccurate, in general, accuracy or appropriateness is more likely to be a scalable attribute. Elicitation is fatiguing work, which must be considered when trying to quantify any decrease in time or effort. Speaking of effort, the "effort" referred to above is simply the amount of effort a participant exerts in order to make themselves understood, but an elicitation task requires a whole separate category of effort: that of generating or identifying the requested structures before communicating them. Finally, each elicitation request is essentially a new task with new (if semantically-related) material, so participants may not have the opportunity to increase efficiency as described above.

Elicitation interactions also involve negotiation, the elements of which are defined in discourse analysis literature as i) the requirement of shared information for collaboration; ii) the incompleteness of each participant's knowledge; and (iii) the need to resolve differences of belief (Sidner, 1994). Negotiation, consisting of proposals, acceptances or rejections, counterproposals, and retractions of proposals (Sidner, 1994), occurs on two levels in an elicitation session: between the elicitor and the native speakers as the elicitation request is negotiated, and between the native speakers themselves as they negotiate the appropriate structure.

If our aim is to describe how these elements of negotiation and collaboration manifest in the elicitation genre, or, put simply, what exactly occurs in this communicative event, we must describe both its constituent parts and their sequence. Austin (1962) and Searle (1975) provided two classical speech act taxonomies as a way of categorizing and sequencing individual utterances. Using Searle's taxonomy, one might hypothesize that most elicitor speech acts would be "directives" (requests), and most participant speech acts would be "representatives" (assertions, statements, suggestions, descriptions). Because the elicitation interaction is a social one, "expressives" (apologies, greetings, complaints) would be expected from both elicitors and participants, as well as some "verdictives" (*that's right*, *you got it,* or in Cree, *ekosi ôma*). Searle's taxonomy therefore, while adequate to generally categorize each utterance, will not enable very complex sequences to be established, so I have created a taxonomy based on the utterances found in the Maskwacîs recording sessions themselves. I conducted an in-depth annotation of a Maskwacîs elicitation session, coding each individual utterance in the appropriate

speech act category. I then compared the speech act sequence of each elicitation request so that any regular pattern could be revealed.

CHAPTER V. ELICITATION SPEECH ACT TAXONOMY

5.1 Methodology.

The in-depth annotation was conducted on the recording of the June 1, 2017, afternoon session. The session was 103 minutes long. In addition to the elicitor, there were three Cree speakers present, one male (M1) and two female (F1 and F2).

For the purposes of the in-depth annotation, I created a typology of speech acts with accompanying shorthand codes to account for every utterance in the elicitation session.

5.2 Elicitor-specific speech acts.

Elicitation techniques. Two types of elicitation occur in the Maskwacîs recording sessions: elicitation of lexical gap items, and elicitation of sentences. Lexical gap items are words or phrases that meet the target prompted by the RWCM questions. For example, "What words indicate that there is a very small amount of something?" For ease of reference, I will refer to lexical gap items (words and phrases) as phrases.

In this session, the elicitor used three techniques to elicit lexical gap items: typing out the English target (non-verbal); a direct translation request (for example, asking participants "How would you say X in Cree?"); and simply stating the English target. The distinction between sentence and phrase elicitation was preserved in the prefix, as seen in Table 5.

Table 5: *Elicitation speech acts*

Sentence Code	**Speech Act**	**Phrase Code**
SE1	typing (non-verbal)	PE1

SE2	direct translation request	PE2
SE3	statement of English target	PE3

5.3 Speaker-specific speech acts.

Negotiation of elicitation. After receiving an elicitation request, the speakers would negotiate an appropriate Cree equivalent among themselves. One speaker would offer an initial suggestion. Negotiation would continue back and forth until a final version was arrived at. This took place most often in Cree, but where English was used it has been noted (see Table 6).

Table 6: *Negotiation of elicitation speech acts*

Code	**Speech Act**
SW1	initial suggestion
SW2	back-and-forth negotiation (Cree)
SW2E	back-and-forth negotiation (English)
SW3	agreed-upon item

Explanations. Explanations of two kinds occurred in this session: when a speaker provided metalinguistic information about usage context or constraints and when a speaker explained the meaning of a Cree construction (See Table 7).

Table 7: *Speaker explanations*

Code	**Speech Act**
ME	metalinguistic explanation
EME	explanation of meaning (English)

Repetition. In the recording section of the session, speakers would say each item twice. At times they would practice the pronunciation, suggesting any amendments they thought were necessary. Amendments were coded as SW2 (see Table 6) or SPC1 (see Table 7) as appropriate.

In the elicitation section of the session, speakers would sometimes simply repeat the elicitation request as they were thinking of a response.

Table 8: *Repetition*

Code	**Speech Act**
PP1	practice before recording
RR	repeat for recording
R1	repetition of the request

5.4 Elicitor-speaker speech acts.

Negotiation of elicitation request. Peripheral to the activity of elicitation itself were discussions about amending the elicitation request (see Table 9). Both the elicitor and speakers participated in these negotiations. This interaction only occurred during sentence elicitation.

Table 9: *Negotiation of elicitation request*

Sentence Code	**Speech Act**	**Phrase Code**
SE4	amend the request	PE4

Clarifications. Clarifications of meaning as well as spelling or pronunciation were highly frequent. I distinguished interactions clarifying the meaning of a Cree construction from interactions clarifying the meaning of the target phrase or sentence (see Table 10). The distinction between the two types of spelling and pronunciation clarifications could also be

described as “pre- and post-agreement”; that is, before and after a consensus had been arrived at as to the correct spelling or pronunciation.

Table 10: *Clarifications*

Code	Speech Act
MC	clarification of the meaning of a Cree construction
MC1	clarification of the target meaning
SPC	spelling or pronunciation clarification to determine correctness
SPC1	spelling or pronunciation clarification for the purposes of the written record

Other. Utterances which are not directly result from the elicitation or recording process nevertheless form an important subset of speech acts that relate largely to managing the interaction from an interpersonal perspective were coded as “other” (see Table 11). These include jokes (O3C and O3E) and complaints (serious or intended for humour – O2C, O2E, and O4), in Cree and English, as well as softening comments (O1), for example:

Elicitor: “So I find these words by themselves pretty hard to translate or describe in English.”

O5 was used to capture utterances that steered the conversation away from the task at hand, for example, if an interactant would start telling a personal story. The remainder of the utterances that could not be categorized as any of the identified elicitation speech acts but were not complete digressions I placed in a catch-all “other” category (O).

Finally, I created a category for both expressing and requesting agreement (coded respectively as agreement and agreement?)

Table 11: *Other*

Code	Speech Act
O	other discussion
O1	softening comments
O2C	humorous complaint (Cree)
O2E	humorous complaint (English)
O3C	joke (Cree)
O3E	joke (English)
O4	complaint
O5	off-topic conversation
agreement	expression of agreement
agreement?	request for agreement

5.5 Results

Speaking time. In this session, it is clear that one speaker took the lead. Out of a 103-minute session, the elicitor spoke for 7.7 minutes, Male Speaker 1 (MS1) for 29.8 minutes, Female Speaker 1 (FS1) for 12.3 minutes, and FS2 for 10.1 minutes for a total of 59.9 minutes (see Table 12).

Table 12: *Speaking time*

Participant	Total Speaking Time	Number of Speech Acts	Mean Duration
Elicitor	464 s	171	2.71 s
Male Speaker 1	1786 s	600	2.97 s
Female Speaker 1	738 s	242	0.57 s
Female Speaker 2	605 s	168	3.60 s

Elicitation speech acts. This particular session contained mainly of sentence elicitation, which is valuable in the sense that anecdotal information from elicitors and the complaints of the speakers in this session raises the inference that sentence elicitation is the most difficult. It does, however, make comparisons between sentence and phrase elicitation more difficult given the lack of phrase examples.

Table 13: *Number of elicitation speech acts*

Sentence Code	Occurrences	Speech Act	Phrase Code	Occurrences
SE1	7	typing (non-verbal)	PE1	0
SE2	2	direct translation request	PE2	4
SE3	15	statement of English target	PE3	3

Negotiation of elicitation speech-acts by speaker. There is evidence here again for MS1 taking a lead role in the negotiation (see Table 14). He makes by far the greatest number of initial suggestions and takes an active role in each negotiation. He also is most likely to restate the agreed-upon final item.

I have included both of the agreement speech acts in this table, since the negotiation context is where they are most likely to occur (along with spelling and pronunciation clarification and, to a lesser extent, meaning clarification). It is perhaps not wholly unexpected that MS1, who occupies the leading role, is the only speaker who explicitly requests agreement, while the instances of agreement by FS1 and FS2 are more directly comparable to MS1 in number than the other speech acts.

Table 14: *Negotiation of elicitation speech acts by speaker*

Code	Speech Act	MS1	FS1	FS2
SW1	initial suggestion	26	8	6
SW2	back-and-forth negotiation (Cree)	84	50	27
SW2E	back-and-forth negotiation (English)	5	8	0
SW3	agreed-upon item	19	4	2
agreement	expression of agreement	45	48	38
agreement?	request for agreement	6	0	0

Explanations. Most of the metalinguistic explanation is provided by MS1. FS1 takes a slightly more active role in providing English explanations of meaning, but MS1 again provides the bulk of that information.

Table 15: *Explanation speech acts by speaker*

Code	Speech Act	MS1	FS1	FS2
ME	metalinguistic explanation	26	1	0
EME	explanation of meaning (English)	27	6	1

Repetition. In this category, one would expect that the numbers would be relatively equal, since each speaker in theory produces the same number of repetitions for the recording. Due to the practical limitations of my annotating practice, mentioned below under "Limitations", there is not a one-to-one relationship between speech acts and annotations. In addition, speakers sometimes produce only one repetition for the recording, for example, if they are having a particularly difficult time pronouncing the item. In spite of these factors, the numbers for the recording-related speech acts have a much higher degree of parity than in the other categories.

Table 16: *Repetition by speaker*

Code	Speech Act	MS1	FS1	FS2
PP1	practice before recording	18	17	12
RR	repeat for recording	36	42	48
R1	repetition of the request	15	4	2

Because it is a pause that causes repetitions for the recording to be annotated separately instead of two together, we can infer that higher RR numbers means more pauses between repetitions. This may be due simply to speech rate, but RR numbers might be to some degree inversely proportional to confidence or ease of production of the repetition items. The fact that lower RR numbers correspond to higher PP1 numbers strengthens this hypothesis somewhat.

Negotiation of elicitation request. Table 17 lists the total per-person counts for the negotiation of elicitation request utterances. Since the negotiation of elicitation request speech act refers to the entire interaction, the number of annotations is not necessarily reflective of the number of negotiations that took place. With this in mind, it does not appear sentence negotiations were highly frequent in comparison to the number of total speech acts, and no phrase negotiations took place.

Table 17: *Negotiation of elicitation request by speaker*

Sentence Code	Elicitor	MS1	FS1	FS2	Speech Act	Phrase Code	All
SE4	8	5	1	0	amend the request	PE4	0

Clarifications. Table 18 lists the total per-person counts for the clarification utterances. The elicitor is highly involved in interactions surrounding meaning clarification.

Perhaps the most important observation from this section is that the bulk of spelling and pronunciation clarification is for the benefit of the written record. The vast majority of the annotations for MS1 are SPC1. In addition to the speaking time involved, there are often long pauses between clarifications as a written form of the item is being recorded. If a speaker took over the task of providing the written record, it is possible a great deal of time would be saved.

Table 18: *Clarifications by speaker*

Code	Speech Act	Elicitor	MS1	FS1	FS2
MC	clarification of the meaning of a Cree construction	20	8	0	0
MC1	clarification of the target meaning	24	15	1	0
SPC	spelling or pronunciation clarification to determine correctness	22	34	9	0
SPC1	spelling or pronunciation clarification for the purposes of the written record	12	146	12	13

Other. Table 19 below lists the total per-person counts for the "Other" utterances. Expressions of agreement are used very frequently by the speakers, and far more so than by the elicitor, likely because they engage in the bulk of the negotiation. Softening comments are used more by the elicitor than any other participant, usually in anticipation of a difficult request. I noted that the dominant speaker (MS1) was more likely to make humourous complaints in English, while the other two speakers (FS1 and FS2) more often made humourous complaints in Cree. After seeing these numbers, I was reminded that MS1 more often explicitly addressed the elicitor or referred to him in his complaints than FS1 and FS2. In terms of MS1's high number request for agreement utterances, I hypothesize it may be due to his number of initial suggestion (SW1) utterances.

Table 19: *Other*

Code	Speech Act	Elicitor	MS1	FS1	FS2
O	other discussion	23	41	16	5
O (inaudible)	inaudible	3	0	0	0
O1	softening comments	5	2	0	0
O2C	humorous complaint (Cree)	0	3	12	13
O2E	humorous complaint (English)	0	10	0	2
O3C	joke (Cree)	0	4	4	2
O3E	joke (English)	3	7	3	0
O4	complaint	0	0	0	1
O5	off-topic conversation	1	2	5	1
agreement	expression of agreement	6	48	48	38
agreement?	request for agreement	0	6	0	0

Jokes. While I have not conducted a fulsome evaluation of the purpose of jokes, from an impressionistic perspective, humour plays an important role in the elicitation sessions. Tension can arise because of fatigue, the difficulty of a target, or (one of the perils of bilingual elicitation) the difficulty of overcoming non-equivalent elements between two languages. Humour acts as a way to break the tension and re-establish the group's unity toward a common purpose.

5.6 Limitations.

The number of annotations can only be used as a general indicator of patterns. In annotating the data, my first priority was an accurate time measurement rather than ensuring a speech act was entirely represented in a single annotation. In the case of a long pause, a single speech act may be represented by two annotations. In the case of repetitions for the recording, two back-to-back repetitions may be included under a single annotation, while two repetitions for

the recording that are separated by a long pause or intervening element may be represented by two annotations.

CHAPTER VI – TASK SEQUENCES

6.1 Broad Patterns.

There are three main tasks within an elicitation session: verification of the words in the original Maskwacîs dictionary, elicitation of new items, and the clean recording.

Borrowing the sections of sonata form, the elicitation task can be sequenced into three parts: Exposition, Development, and Recapitulation (recap). Each section contains standard and optional speech acts from those identified above. In the majority of tasks, each section occurred only once. I have called these "typical." In a non-typical task, sections may reoccur out of order; for example, when a form has been agreed upon (development), spelling is being confirmed (recap), and an amendment is suggested (redevelopment). Examples of typical and non-typical sequences are provided below.

Exposition: Introduction of the target. The main actor in the Exposition section is the elicitor, who will introduce the target to the speakers. The target may be a lexical gap item or a sentence. Depending on the elicitor's perception of the difficulty of the target and the response from the speakers, he or she may use one or more strategies to introduce the target. This may include any of the elicitation techniques outlined in Table 5 or clarification of meaning. Negotiation around amending the target (SE4) can occur in this section. Agreement may occur in any interactional task. Speakers may also introduce metalinguistic explanations or explanations of meaning (see Table 15).

Development: Negotiation of the item. The main actors in the Development section are the speakers, who will negotiate an item they deem meets the target requested by the elicitor. The standard speech acts include an initial suggestion (SW1) and amendments and negotiation (SW2), and selection of the approved item (SW3). Optional speech acts include explanations of

meaning (EME), meaning clarifications (MC1), metalinguistic explanations (ME), agreement, and other discussion (O to O5). In some of these optional speech acts, the elicitor is included in the interaction.

It should be noted that SW3, while not included under the optional speech acts, is not always present. At times there is a clear verbal consensus: one speaker presents the approved item (SW3), and the other speakers express agreement. At times, a speaker (usually the dominant speaker) moves directly to clarification of spelling and pronunciation. This does not mean consensus has not been reached, just that it has not been directly expressed. It is also possible in some instances that I was not able to recognize the presentation of SW3 because it was couched in a longer Cree utterance.

Recap: Recording the approved item. This section consists mainly of spelling and pronunciation clarification of the approved item selected by the end of the Development section. The main speech act is spelling and pronunciation clarification of the approved item for the benefit of the elicitor as he or she records the specific form (SPC1). At times the speakers negotiate the spelling and pronunciation they wish to approve. Other optional speech acts include explanation of meaning (EME), clarification of meaning (MC), agreement, and other discussion (O to O5). There are instances where the item itself is renegotiated (a return to the development section).

Coda. In some instances, additional speech acts occur after these sections are complete. These mainly fall into the "other" category, and can include jokes, softening comments, off-topic discussions, and metalinguistic explanation. I have not included the coda in any calculation of time or total number of speech acts.

6.2 Typical task.

Below is an example of a typical task sequence.

Exposition	Elicitor: MC1	clarifies the target meaning
	Elicitor: SE3	states an English sentence containing the target
	Elicitor: PE2	direct translation request for the target
Development	FS1: SW1	offers initial suggestion
	FS2: agreement	agreement
	MS1: SW2	negotiation of target
	FS1: SW2, SW2E	negotiation of target (in Cree and English)
	FS2: agreement, SW2	agreement, negotiation of target
	FS1: agreement	agreement
	FS2: SW2	negotiation of target
	MS1: SW2	negotiation of target
	FS2: SW3	statement of approved item
Recap	Elicitor: SPC	spelling/pronunciation clarification
	MS1: SPC1	spelling/pronunciation clarification
	FS2: SPC1	spelling/pronunciation clarification
	MS1: SPC1	spelling/pronunciation clarification
Coda	FS1: O	other discussion

Total task time (not including coda): 83 seconds. Recap time: 27 seconds.

6.3 Non-typical task.

Below is an example of a non-typical task sequence.

Exposition	Elicitor: SE3	states an English target sentence
Development	MS1: SW1	offers initial suggestion
Recap	MS1: SPC1 (x2)	spelling/pronunciation clarification
Redevelopment	MS1: SW2	negotiation of target
Recap	MS1: SPC1 (x6)	spelling/pronunciation clarification
	FS2: SPC1	spelling/pronunciation clarification
	MS1: SPC1 (x3)	spelling/pronunciation clarification
Redevelopment	FS1: SW2	negotiation of target
	MS1: agreement	agreement
	FS1: SW2	negotiation of target
	MS1: SW2	negotiation of target
	FS2: SW2	negotiation of target

Total task time: 85 seconds. Recap time (including all redevelopment): 76 seconds.

6.4 Speech act type-token ratio.

As a measure of variation, I calculated the type token ratio for each section of the speech act over 17 tasks, approximately one-third of the session. Out of the 17 tasks, 4 were excluded because they were non-typical. The exposition section had the highest TTR at 0.74, while the development and recp sections' TTRs were very similar at 0.45 and 0.43 respectively.

Table 20: *Speech act type-token ratio*

Section	**Types**	**Tokens**	**TTR**
Exposition	45	61	0.74
Development	45	101	0.45
Recap	59	136	0.43

A greater variety of speech acts and more speech acts per task are used during recap. While exposition and development have the same variety of speech acts, more speech acts per task are used in development. For exposition, the average number of speech acts per task is 4.7; for development, the average number of speech acts per task is 7.8. The average number of speech act types per task is the same for both sections at 3.5. In the recap section, the average number of speech acts per task is 10.5, and the average number of speech act types per task is 4.5.

Table 21: *Average speech acts and types per task*

Section	**Average Speech Acts/Task**	**Average Speech Act Types/Task**
Exposition	4.7	3.5
Development	7.8	3.5
Recap	10.5	4.5

6.5 Additional examples.

I conducted the same analysis on six samples from two other elicitation sessions, each led by different elicitors. I hypothesized that, while the structure of the task sequence would likely remain relatively constant, different elicitors may use different elicitation strategies than those I identified in my original session. Samples were selected randomly apart from pre-selecting sessions from separate elicitors and selecting two instances of phrase elicitation and one of sentence elicitation from each session. All the tasks followed the prototypical sequence explained above; however, I encountered two utterances distinct enough to warrant the creation of a new code.

The first I described as "metalinguistic clarification" (coded as MEC). In one session, the elicitor, rather than asking for clarification of *meaning*, asked a question clarifying *usage*. This prompted a metalinguistic discussion among the speakers.

In the same task, a meaning clarification from the elicitor prompted one of the speakers to suggest a new item (coded as NSW1). Thus, the elicitor's meaning clarification in the coda of one task simultaneously acted as the exposition of a new task. The full sequence appeared as follows (the new items are in bold type):

Combined task.

Exposition	Elicitor: PE2	direct translation request
	MS1: MC1	clarification of the target meaning
	Elicitor: MC1	clarification of the target meaning
Development	FS1: SW1	offers initial suggestion
	MS1: SW2	negotiation of target
	MS1: MC	clarification of the meaning of the Cree construction
	FS1: agreement	agreement
	FS2: agreement	agreement
Recap	MS1: SPC1	spelling/pronunciation clarification
	FS2: SPC1	spelling/pronunciation clarification
Coda	**Elicitor: MEC**	**metalinguistic clarification**
	MS1: MEC	**metalinguistic clarification**
	FS2: MEC	**metalinguistic clarification**
	MS1: EME	explanation of meaning (English)

Exposition.2	Elicitor: MC	clarification of the meaning of the Cree construction
Development.2	**FS2: NSW1**	**initial suggestion of new item**
	FS2: SW2	negotiation of target
	FS1: SW2	negotiation of target
	FS2: agreement, SW3	agreement, agreed-upon item
Recap.2	Elicitor: MC	clarification of the meaning of the Cree construction
	MS1: agreement	agreement

Phrase vs. sentence elicitation. Four out of the six samples I analyzed were phrase elicitation tasks. While the sample size is too small to make any stable inferences, the Development (negotiation) stage of phrase elicitation tasks tended to be shorter. There were between 1 and 3 turns in the four phrase negotiations (initial suggestions (SW1) and negotiation (SW2). The two sentence negotiations each took 5 turns. I did not include repetition of the agreed-upon item (SW3) or any clarifications or explanations of meaning in the turn counts.

Conclusions after reviewing additional examples. As expected, a change in elicitor did not require a new task sequence. Given that the tasks in all elicitation sessions are constrained by the Rapid Word elicitation sheets and prompts, this is not surprising.

The new utterance types may be due to the style of the individual elicitor, or they could be the result of the particularities of phrase elicitation, very little of which formed part of the original session.

The fact that negotiation of the target was completed in fewer turns during phrase elicitations than sentence elicitations is some evidence to support anecdotal accounts of researchers that sentences are significantly harder to produce than words or phrases.

CHAPTER VII – CONCLUSION

After four years and hundreds of hours of work by native speakers and University of Alberta researchers, the Maskwacîs recording project is drawing to a close. Not only has the dictionary doubled in size, but the project has generated a rich data source in addition to the new lexical items. The data gathered has furnished evidence of morphological change, and some recordings are currently being used for phonological analysis (Schmirler, 2018). As the recordings are annotated so that segments can be used to generate audible examples for the web-based dictionary, spontaneous Cree speech is also marked so it can be more easily extracted for future analysis. Metalinguistic information and even some oral history are also being marked and preserved in the sound files. Miyo Wahkotowin Education and the University of Alberta ALTLab are discussing future collaboration.

Creating a speech act taxonomy was an arduous process, not only because of the time required for annotation, but because of the unavoidable subjectivity involved in classifying utterances. There is no doubt a different author would process different results. Also, because the elicitation sessions have an inherent repetitive structure and are guided by formulaic elicitation sheets, it is perhaps not surprising that a speech act analysis shows evidence of regular sequencing. In spite of these limitations, what ultimately became clear was that there was enough regularity in the task sequencing of an elicitation to identify when an elicitation request went "smoothly" (in a typical sequence), and when it did not. It also became apparent that the difference between a typical and non-typical task sequence did not depend on the elicitation request technique used by the elicitor. Identifying the speech acts also revealed the interpersonal

nature of the elicitation activity, particularly in the use of humour to mitigate the effects of fatigue and frustration.

Further discourse analysis of the data correlated with features of the requested target (for example, investigating whether the lexical relation of the target to the semantic domain correlates with a systematic result or non-result across semantic domains), might provide further information, both about the linguistic structure of Cree, and about the ease of production of certain kinds of targets. Requests that require longer exposition, that is, more meaning clarification or metalinguistic explanations, could be analyzed as a group to see if they have features in common that make them more difficult targets for a Cree speaker. Items generated in non-typical task sequences might be good candidates for acceptability judgment experiments to determine whether there is more than one acceptable way of representing the target. The recordings of native speakers negotiating items in Cree could be a source for analysis of higher-level Cree discourse features. The tendency for one dominant native speaker to emerge in a session could also be explored in order to investigate whether this is determined by personal attributes like proficiency or confidence, or if it is a feature of discourse patterns.

Finally, my documentation of the methodology employed by the Maskwacîs Spoken Cree Dictionary Project is hopefully an interesting case study, not only of the novel adaptation of an already little-documented methodology (the Rapid Word Collection Method), but in that it illustrates two of the requirements of any language documentation project: that it be both flexible and feasible under its own constraints.

REFERENCES

AMEKA, FELIX K.; DENCH, ALAN; AND EVANS, NICHOLAS. 2008. *Catching Language: The Standing Challenge of Grammar Writing.* Berlin: Mouton de Gruyter.

ARPPE, ANTTI AND JÄRVIKIVI, JUHANI. 2007. Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3-2.131-159.

ARPPE, ANTTI; GILQUIN, GAËTANELLE; GLYNN, DYLAN; HILPERT, MARTIN; AND ZESCHEL ARNE. 2010. Cognitive corpus linguistics: five points of debate on current theory and methodology. *Corpora* 5(2).1-27.

ARPPE, ANTTI. 2016. *Miyo Wahkohtowin/Maskwacîs dictionary recording project* [PowerPoint slides]. Retrieved from ALTLab.

AUSTIN, JOHN L. 1962. *How to Do Things with Words.* Oxford: Clarendon.

AUSTIN, PETER K. (ed.) 2007. *Language Documentation and Description, vol 4.* London: SOAS.

AUSTIN, PETER K., & GRENOBLE, LENORE. 2007. Current trends in language documentation. In Austin (2007). 12-25.

BAKHTIN, MIKHAIL M.. 1986. *Speech Genres and Other Late Essays.* Emerson, Caryl and Holoquist, Michael (eds). Translated by McGee, Vern W. Austin: University of Texas Press.

BERGE, ANNA. 2010. *Adequacy in documentation.* In Grenoble and Furbee. 51-66.

BROWN, GILLIAN; ANDERSON, ANNE; YULE, GEORGE; AND SHILLCOCK, RICHARD. 1984. *Teaching talk: Strategies for production and asssessment.* New York: Cambridge University Press.

CALDECOTT, MARION AND KOCH, KARSTEN. 2014. Using mixed media tools for eliciting discourse in Indigenous languages. *Language Documentation and Conservation* 8.209-240.

COLE, PETER AND MORGAN, JERRY L. 1975. *Syntax and Semantics* (3). New York: Academic Press

CROWLEY, TERRY. 2007. *Field Linguistics: A Beginner's Guide.* Oxford: Oxford University Press.

DAVIES, BETHAN L. 2006. Testing dialogue principles in task-oriented dialogues: An exploration of cooperation, collaboration, effort, and risk. *Leeds Working Papers in Linguistics and Phonetics* 11.

DREW, PAUL AND HERITAGE, JOHN. (eds.) 2006. *SAGE Benchmarks in Social Research Methods: Conversation analysis.* (1-4). London: SAGE Publications Ltd.

FRAWLEY, WILLIAM; HILL, KENNETH C.; AND MUNRO, PAMELA. (eds.) 2002. *Making dictionaries: Preserving Indigenous Languages of the Americas*. Berkley and Los Angeles: University of California Press.

GERNSBACHER, MORTON ANN AND GIVÓN. (eds.) 1995. *Coherence in Spontaneous Text.* Amsterdam: John Benjamins Publihsingn Co.

GILQUIN, GAËTANELLE and GRIES, STEFAN. TH. 2008. Corpora and experimental methods: a state-of-the-art review. *Corpus Linguistics and Linguistic Theory* 5(1).1-26.

GIPPERT, JOST; HIMMELMANN, NIKOLAUS P.; AND MOSEL, ULRIKE. (eds.) 2006. *Essential of language documentation.* Berlin: Mouton de Gruyter.

GRIMES, JOSEPH. E. 2002. Lexical functions as a heuristic for Huichol. In Frawley, Hill, and Munro. 70-85.

GRENOBLE, LENORE A. AND FURBEE, N. LOUANNA. (eds.) 2010. *Language Documentation: Practice and Values.* Amsterdam: John Benjamins Publishing Company.

HIMMELMANN, NIKOLAUS. P. 1998. D

HARRIGAN, ATTICUS. 2018. Personal communication. (See Appendix B). https://voices.uchicago.edu/lvc/files/2013/11/himmelmann-documentary-and-descriptive-linguistics.pdf

HIMMELMANN, NIKOLAUS P. 2002. Documentary and descriptive linguistics (full version). In Sakiyama and Endo. 37-83.

HIMMELMANN, NIKOLAUS P. 2006. Language documentation: what is it and what is it good for? In Gippert, Himmelman, and Mosel (eds.) 1-30.

HIMMELMANN, NIKOLAUS P. 2012. Linguistic data types and the interface between documentation and description. *Language Documentation and Conservation* 6.187-207.

LABOV, WILLIAM. 1972. *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania.

MAYER, MERCER. 1969. *Frog, where are you?* New York: Dial Press.

MILLER, GEORGE A.; BECKWITH, RICHARD; FELLBAUM, CHRISTIANE; GROSS, DEREK; AND MILLER, KATHERINE. 1993. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography* 3/4.235-44.

MOE, RONALD. 2003. Compiling dictionaries using semantic domains. *Lexikos* 13.215-223. Retrieved from http://lexikos.journals.ac.za/pub/article/view/731.

MOE, R. 2004. Producing dictionaries using semantic domains. In Sin-wa. 55-61.

MOSEL, ULRIKE. 2006. Grammaticography: The art and craft of writing grammars. In Ameka, Dench, and Evans (eds.) 41-67.

RapidWords.net. 2017, October. Retrieved from rapidwords.net

SAKIYAMA, OSAMU AND ENDO, FUBITO. (eds.) 2004. *Lectures on endangered languages* 5. Kyoto: Endangered Languages of the Pacific Rim.

SCHMIRLER, KATHERINE. 2018. Personal communication. (See Appendix C).

Semdom.org (2018, April). Retreived from semdom.org.

SEARLE, JOHN R. 1975. Indirect speech acts. In Cole and Morgan.

SIDNER, CANDACE L. 1994. Negotiation in collaborative activity: a discourse analysis. *Knowledge-Based Systems* 7(4).265-267.

SIL INTERNATIONAL. 2017. *RWC overview* [PDF document]. Retrieved from RapidWords.net. Website: http://rapidwords.net/sites/default/files/docs/RWC%20Overview.pdf.

SIN-WA, CHAN. (ed.) 2013. *Translation and bilingual dictionaries.* Tübingen: Niemeyer.

WILKES-GIBBS, DEANNA. 1995. Coherence in collaboration: Some examples from conversation. *Coherence in Spontaneous Text.* In Gernsbacher and Givón. 239-267.

WIERZBICKA, ANNA. 1996. *Semantics: Primes and universals.* Oxford & New York: Oxford University Press.

WOODBURY, ANTHONY C. 2003. Defining documentary linguistics. In Austin (2007). 35-51.

APPENDIX A

Sample Domain and Elicitation Sheet

1.1.1.1 Moon

Use this domain for words related to the moon. In your culture people may believe things about the moon. For instance in European culture people used to believe that the moon caused people to become crazy. So in English we have words like "moon-struck" and "lunatic." You should include such words in this domain.

(1) What words refer to the moon?
- *moon, lunar, satellite*

(2) What words refer to how the moon moves?
- *rise, set, sink*

(3) What words refer to the time when the moon rises?
- *moonrise, rising of the moon,*

(4) What words refer to the time when the moon sets?
- *moonset, setting of the moon,*

(5) What words refer to when the moon is shining?
- *moon is shining, by the light of the moon, moon is/comes out,*

(6) What words describe where the moon is shining?
- *moonlit, be in the moonlight, lit by the moon, shine on,*

(7) What words describe when or where the moon doesn't shine?
- *moonless night, eclipse of the moon, eclipse (v),*

(8) What words refer to the light of the moon?
- *moonlight, moonbeam, moonshine,*

(9) What words describe the brightness of the moon?
- *bright, pale,*

(10) What words describe the appearance of the moon?
- *man in the moon, harvest moon*

(11) What words refer to the spots on the moon?
- *lunar sea, crater,*

(12) What words refer to the phases of the moon?
- *phase, new moon, full moon, half moon, quarter moon, crescent, sliver, moon is waxing (getting bigger), moon is waning (getting smaller), wax and wane*

(13) What words refer to the time it takes for the moon to go through its phases?
- *(lunar) month*

APPENDIX B

Atticus Harrigan Interview

Maskwacis Spoken Cree Dictionary Project

Interviewee: Atticus Harrigan

1. How long have you been involved in the project, and what was your role(s)?

I've been involved for around 4 years now, since just after the inception of the project. I've helped coordinate and schedule the recordings, acted as an elicitor, managed and preformed annotations, and managed the inventory of equipment.Earlier on, I helped prepare the elicitation documents.

2. Please describe the evolution of the project as you are aware of it (methodology, goals).

The goals and methodology of the project have more or less stayed the same throughout the project. There have been slight tweaks, such as the recording of metadata, the marking of new sentences/words, and running parallel sessions with single elicitors. We also moved spaces from the Nehiyaskweyahk Cree department's Sohki House to the Education board's head office.

3. What aspect of the project did you expect to be most difficult for elicitors? for participants? Please be as specific or as general as you wish.

I imagined getting into grammatical terms, race terms, religion, and sexuality to be most difficult for participants. For elicitors I imagine (and still believe) the annotation to be the most arduous.

4. What was most difficult for elicitors? for participants? Please be as specific or as general as you wish.

Issues of sex, religion, and grammar were difficult for participants, and annotations were difficult for elicitors; however, by far the most difficult aspect for everyone (that I saw) was the moving of spaces.

While sound insulation was slightly better in the head office location, the Sohki house offered a cozy environment, more or less adequate recording conditions, staff available at predictable hours (so we were never locked out of recording rooms), and staff regularly provided lunch and conversation with the participants. Conversely, the head office was very sterile feeling. We were recording in a kitchen that felt more like something from an old abandoned house and various offices which were often cramped and extremely hot. We were fairly secluded from any lunch options, and participants reported general discomfort and displeasure with the location. There was also a fridge which made a loud humming noise in the kitchen (which was often unplugged for recordings). People regularly interrupted sessions to use the kitchen for their lunch (though they did their best to avoid using it when we were there) and we regularly arrived to find our recording areas locked with no one available to unlock them for half an hour. There were so many disadvantages to this recording space that I'd actually suggest doing recording on campus, which while inconvenient for participants and takes away from the linguistic ecology, is still less of a waste of time and significantly more comfortable.

5. What was the value of using Rapid Words or semantic domains in this project? Were there any drawbacks?

The advantage of RapidWords is a readymade ontology which provides a semantic structure to the day's recording. The difficulty in this process is that RapidWords is very clearly written from a white christian perspective. Many questions were inappropriate for the Cree culture, so there is some editing that needs to be done. Like all ontologies, it is not perfect and says more about the creator than the material. There was also many repeated words and prompts which should have been deleted.

6. Did you receive data that you did not expect? If so, what do you think is the best way to capitalize on its acquisition?

I was surprised at the amount of terms we received for different races.

7. What advice would you give to someone at the beginning of a similar project?

Carefully curate your documents to not be repetitive, as speakers don't like wasting their time. Make sure your questions are relevant and respectful. Be ready to change on the fly.

APPENDIX C

Katherine Schmirler Interview

Maskwacis Spoken Cree Dictionary Project

Interviewee: Katie Schmirler

1. How long have you been involved in the project, and what was your role(s)?

I have been involved in the project since September 2015. In that time, I have primarily been one of the primary elicitors, but I have also worked on other aspects for the eventual implementation of the online spoken dictionary. This has involved comparing the content of the Maskwacîs Dictionary to the lexicon of our morphological model, which is based on the nêhiyawêwin: itwêwina / Cree: Words dictionary, as well as assisting in the development of the morphological model and the paradigms that will underlie the online dictionary.

2. Please describe the evolution of the project as you are aware of it (methodology, goals).

When I came into the project, it had already been going on for a year, so most of the initial problems had been worked out. As time went on it became clear that the project was going to take longer than originally expected, so we began to make some changes, such as aiming for words rather than sentences, which were an initial goal, and also changing from two elicitors to one per session, to make better use of our time.

3. What aspect of the project did you expect to be most difficult for elicitors? for participants? Please be as specific or as general as you wish.

When I joined the project, I had only limited fieldwork experience from a few fieldwork courses, all of which involved a number of students talking with one speaker. Because of that, the larger group was daunting for me. I was also expecting to have a difficult time with the transcription, even with my

background studying Plains Cree, and other elicitors had even less experience. For the speakers, I had expected more problems to arise with the transcriptions being projected to the room

4. What was most difficult for elicitors? for participants? Please be as specific or as general as you wish.

One of the most difficult aspects for everyone was simply tiredness. By the second half of the afternoon, everyone was tired, mentally drained, and sometimes a bit grumpy. That often meant slowing down, or stopping early. It was important to read the room and not push anyone too far, but overall, as problems go, it could have been worse.

5. What was the value of using Rapid Words or semantic domains in this project? Were there any drawbacks?

I found the semantic domains of Rapid Words to be a useful tool for the most part, especially with the Maskwacîs Dictionary also coded for the domains; it got everyone into the right headspace. A downside was that we were working with this outline designed for a wide variety of languages, and we didn't take the time to tailor it more for Plains Cree. This sometimes led to detours into unrelated words, spending too much time trying to come up with words for not entirely relevant concepts, or just losing the flow when we encountered weird or sometimes almost offensive terms.

6. Did you receive data that you did not expect? If so, what do you think is the best way to capitalize on its acquisition?

In terms of lexical items I can't recall anything overly surprising, but some interesting aspects of morphology and phonology have arisen. Some of the morphological patterns seen in Maskwacîs suggest that some things have changed in the 40+ years since the last detailed grammatical description was published, and in my own research I have made use of the Maskwacîs recordings to investigate

phonetic/phonological questions, which has led to some interesting results I wasn't expecting from the descriptions I've read.

7. What advice would you give to someone at the beginning of a similar project?

First, definitely plan on it taking more time than you expect. If it's possible, start the project and get a feel for how long a section takes before setting any firm end dates. Also, think about the size of group you want to work with. Three ended up working out pretty well for us. With three speakers, there was usually a good chance of getting all words we looked at, while two speakers couldn't always think of them. Four or more speakers would have a lot of words, but also came with a lot of time spent on discussion that was less relevant. For just creating a dictionary, more speakers might be a good thing, but seeing as we were also recording each word, four or more speakers could really slow down the process.

APPENDIX D

Speech Act Codes Reference List

Table 5: *Elicitation speech acts*

Sentence Code	**Speech Act**	**Phrase Code**
SE1	typing (non-verbal)	PE1
SE2	direct translation request	PE2
SE3	statement of English target	PE3

Table 6: *Negotiation of elicitation speech acts*

Code	**Speech Act**
SW1	initial suggestion
SW2	back-and-forth negotiation (Cree)
SW2E	back-and-forth negotiation (English)
SW3	agreed-upon item

Table 7: *Speaker explanations*

Code	**Speech Act**
ME	metalinguistic explanation
EME	explanation of meaning (English)

Table 8: *Repetition*

Code	**Speech Act**
PP1	practice before recording
RR	repeat for recording
R1	repetition of the request

Table 9: *Negotiation of elicitation request*

Sentence Code	**Speech Act**	**Phrase Code**
SE4	amend the request	PE4

Table 10: *Clarifications*

Code	**Speech Act**
MC	clarification of the meaning of a Cree construction
MC1	clarification of the target meaning
SPC	spelling or pronunciation clarification to determine correctness
SPC1	spelling or pronunciation clarification for the purposes of the written record

Table 11: *Other*

Code	**Speech Act**
O	other discussion
O1	softening comments
O2C	humorous complaint (Cree)
O2E	humorous complaint (English)
O3C	joke (Cree)
O3E	joke (English)
O4	complaint
O5	off-topic conversation
agreement	expression of agreement
agreement?	request for agreement