Nominal Classification in Michif

by

Olivia Nathene Sammons

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics
University of Alberta

Abstract

The maintenance of multiple systems of nominal classification is typologically uncommon, as is the transfer of noun class systems in language contact situations (Corbett 1991; Good 2012). Michif (ISO 639-3: crg), a critically endangered language spoken by members of the Métis Nation on the northern Great Plains, presents an exception to both of these generalizations, having inherited two systems of nominal classification from its source languages—French-derived gender (masculine/feminine), and Algonquian-derived animacy (animate/inanimate) (Bakker 1997; Papen 2003a). This study investigates Michif nominal classification in detail, considering both the relationship between the animacy and gender values observed in Michif and their equivalents in Cree and French, and the assignment of animacy and gender values to loanwords from English. Corbett (1991) questions whether or not any clear-cut examples of languages with "two independent gender systems" (188) can be identified cross-linguistically, and others have claimed that masculine-feminine gender in Michif is either weakening (Gillon & Rosen 2018) or fossilized (Stoltzfus & Boissard 2016). However, through quantitative investigation of animacy and gender assignment patterns, this study finds that Michif has two independent and productive grammatical categories of noun classification, each inherited from a different source language.

The data analyzed in this study are drawn from a subset of a 60-hour multimodal corpus of contemporary spoken Michif, developed by the author in collaboration with 42 members of Métis communities in Manitoba, Saskatchewan, and Alberta, Canada primarily during the period of 2011–2016. Statistical analysis of a dataset consisting of 261 lemmas from this corpus finds that the animacy and gender values of Michif nouns

align with those of its Cree and French source languages in the overwhelming majority of cases. This large-scale alignment, even in cases of semantic irregularity, indicates that these systems have largely been inherited in their full complexity in Michif. With more than one source language introducing syntactic and semantic categories into Michif, this finding underscores the importance of (a)symmetry in speakers' linguistic competence in the development of models of language genesis in contact situations.

Additional statistical tests find no signs of interaction between animacy and gender systems, motivating a treatment of Michif as having two separate, co-existing systems of nominal classification, rather than a single, merged gender system—a cross-linguistically uncommon result which has relevance to current typologies of nominal classification (e.g., Corbett 1991). This is further substantiated by the observation of a statistically significant difference between gender assignment patterns in French-origin lemmas as compared to English-origin lemmas, while no such difference is found in animacy assignment patterns. In addition, the synchronic results of this study partially corroborate the hypothesis that French-derived gender will be less stable than animacy over time (cf. Gillon & Rosen 2018), although it is found that animacy is also not immune to regularization to a default grammatical value.

This study concludes that animacy and gender remain productive categories in Michif, rather than appearing only as fossilized elements in nominal constructions. This is supported by the observations that a) every lemma in the language must have values for animacy and gender, as indicated by the mandatory nature of grammatical agreement for these categories; b) with few exceptions, these values are stable and shared by speakers; and c) these values are always assigned to new lexical items brought into the

language, even when the resulting classifications cannot be easily attributed to inheritance, as in the case of English borrowings.

Preface

This dissertation is an original work by Olivia N. Sammons. The research project, of which this dissertation is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name “Michif Language Documentation”, No. 10819, February 12, 2010 (renewed March 22, 2011; February 15, 2012; March 13, 2013; March 13, 2014; March 10, 2015; and March 22, 2016). No part of this dissertation has been previously published.

Acknowledgements

This work has been the result of several influences and sources of support over the last several years. First and foremost, I would like to thank Verna DeMontigny, my primary Michif teacher and collaborator. Thank you for welcoming me into your home and family, and for being so generous with your time, experiences, stories, language, and humour. I am forever grateful.

I would also like to acknowledge the many other Michif speakers who participated in this project, without whom this work would not be possible: Eliza Aubichon, Cecile Burroughs, †Daniel Daigneault, †Victoria Daigneault, †Louis Ducharme, Louise Dufour, Liz Durocher, Tony Durocher, James Favel, Marie Favel, †Rita Flamand, Edna Fleury, George Fleury, Harvey Fleury, Irene Fleury, Lawrance Fleury, Mary Fleury, †Mervin Fleury, Norman Fleury, Angus Gardiner, †Victoria Genaille, Louise Gregory, Thérèse Laliberté, Tony Laliberté, Shirley LaRocque, George Lavallée, †Clifford Ledoux, Grace Ledoux-Zoldy, Yvonne Longworth, Lawrence Morin, Max Morin, R.J. Morin, George Pelletier, Harvey Pelletier, Louis Roy, Thomas ("T.J.") Roy, Stanley Smith, Edwin St. Pierre, Harriet St. Pierre, Marie Tanner, and Gail Welburn. Of these, I am particularly indebted to †Mervin Fleury and T.J. Roy for introducing me to other members of their communities, and for their overall enthusiasm, support, and facilitation of our sessions. *Kihchi-marsii!*

This dissertation would not be what it is today without the help of my supervisory committee members—David Beck, Terry Nadasdi, and Nicole Rosen. David Beck, my committee chair, has been supportive at every stage of this research, from the conceptualization (and subsequent reconceptualizations) of this project, to providing advice on fieldwork and other professional matters. He has also spent countless hours providing detailed and thoughtful feedback throughout the writing process, substantially improving the overall quality of this work. Terry Nadasdi introduced me to Canadian French, guided me through my first attempts at sociolinguistic research, and has generally offered support whenever needed. Nicole Rosen provided me with the opportunity to work on the Michif Dictionary Project, helped introduce me to Verna DeMontigny, and also provided helpful comments on earlier drafts of this dissertation. I am also grateful to the external members of my committee, Peter Bakker and Evangelia Daskalaki, for their time and input.

This research was primarily funded by a grant from the Hans Rausing Endangered Languages Project, and was supplemented by grants from the Jacobs Research Fund and the Phillips Fund for Native American Research. Thanks are due to the linguistics department at the University of Alberta for hosting these grants, and particularly to Elizabeth French for her help in navigating the administrative aspects of this funding.

Outside of my committee, I would like to thank Conor Quinn for advice on grant writing and for stimulating discussions at various stages of writing and analysis. I am also grateful to Christopher Cox and Yvonne Lam for their help with statistics. Thank you to Tracy Hirata-Edds for checking in from time to time and for providing feedback on very early drafts of several of the chapters in this dissertation. You are a saint. Thanks also go to my fellow graduate students in the linguistics department at the University of Alberta, particularly Simon Fung, Michelle Garcia-Vega, Sullay Kanu, Kevin Penner, Conor Snoek, Tamara Sorenson-Duncan, and Corey Telfer. I would also like to thank the Sauk Language Department team for their understanding and willingness to temporarily put other projects on hold while I completed this dissertation.

There are several other influences which have guided me over the years and which continue to do so. Heartfelt thanks are due to Marcellino Berardo for providing me with my first introduction to Algonquian linguistics; to Heather Souter for introducing and encouraging me to work on Michif; and especially to Mary Linn, for modelling what it is to be a good researcher, mentor, colleague, and friend.

With the last several years having been somewhat transient, finding good places in which to pursue this work has been challenging at times. I would like to thank CoSpace North in Whitehorse for providing me with a workspace at a discounted rate so that I could have a dedicated place to work in the company of others. This was a lifesaver for me, especially in the dark Yukon winter months. I would also like to thank Jeff Mühlbauer for generously letting me use his office at Brandon University during the summer months to meet with speakers and to have a place to work while "in the field."

To my in-laws, Doug and Adeline Cox, thanks for letting me store my fieldwork supplies in your basement, and for allowing me to keep my car parked in your driveway for months at a time while I was living in Whitehorse so that I would be able to get around for fieldwork when "down South". This made everything so much easier.

To my parents, Oliver and Paula Sammons, thank you for your constant love and support throughout these many, many years of schooling. I'm finally done!

To Charlotte, my sunshine girl, and to the little one on the way. I'm so lucky to be your mama.

Finally, to Chris – my rock, my love, my constant source of encouragement and support. You've been there every step of the way, whenever I needed a sounding board, technical advice, a diversion, or just a good dose of perspective. You never seemed to doubt that I would finish even when I doubted myself. Your love and support mean the world to me. Thank you for everything.

Table of contents

List of Tables

List of Figures

Abbreviations

1	first person
2	second person
3	third person
AN	animate
ART	article
CONJ	conjunct marker
DEM	demonstrative
EXCL	exclusive
FEM	feminine
FUT	future
GEN.OBJ	general object
HES	hesitation
IN	inanimate
INCL	inclusive
INDEF.ACTOR	indefinite/unspecified actor
INTERJ	interjection
MASC	masculine
NEG	negation
NA	noun, animate
NI	noun, inanimate
NON3	non–third person
OBV	obviative
PL	plural
POSS	possessive
PREP	preposition
PROX	proximate
PST	past
PTV	partitive
SG	singular
VII	verb, inanimate intransitive
VAI	verb, animate intransitive
VAIT	verb, pseudo-transitive
VTI	verb, transitive inanimate
VTA	verb, transitive animate
VOL	volitional

[inc] = incomprehensible

Chapter 1: Introduction

Contact languages have been underdocumented relative to other kinds of natural languages, and nominal classification systems in contact languages are particularly understudied. The tacit assumption has been that intricate systems of nominal classification are one of the features most likely to be lost in the processes of simplification that are often associated with the formation of contact varieties such as pidgins and creoles. This study brings new evidence from Michif (ISO 639-3: crg) into this discussion which might contribute to existing typologies of nominal classification systems, particularly where contact languages are concerned.

Michif, one of the traditional languages of the Métis, emerged in the early 19th century as a result of intermarriage between First Nations women and European fur traders in the Red River settlements of Manitoba (Bakker 1997; Rosen & Souter 2009a). Their descendants began speaking a new language that combined elements from Algonquian languages—Cree and Saulteaux—with Canadian French. Michif is considered to be a contact language in the sense of Thomason (2001) in that it is a new language which arose in a contact situation and cannot be traced back to a single source (158).[1] As societal multilingualism was common among Métis communities historically (Rosen 2016a; Rosen & Souter 2009b), many scholars have concluded that Michif originated within a context of stable bilingualism involving Plains Cree and French,

[1] Note that this differs from another usage of the term, in which "contact language" refers more generally to languages of wider communication between groups with no common language. This latter sense essentially treats contact languages as lingua francas, and commonly presumes that these are simplified varieties, driven by the expressive needs of speakers, which developed in contact situations. As Thomason (2001) notes, "Under that definition, any language that is used for intergroup communication is a contact language—including not only pidgins and creoles, but also nonpidgin/noncreole languages like English, which is certainly the most widely used lingua franca in the modern world" (158).

which contributed to the observed mixture of French and Cree we see today (cf. Bakker & Papen 1997: 353; Thomason & Kaufman 1988: 228–33). Michif is thus also commonly identified as a mixed language, i.e., "a type of contact language that arises as the result of the fusion of two languages, normally in situations of bilingualism" (Meakins 2017).

In addition to identifying features of nominal classification in Michif which might be of typological relevance, this study aims to make a contribution to the descriptive literature on Michif, which has noted the occurrence of both animacy and gender largely in passing in previous works, without much further elaboration (but see Hogmen 1981; Papen 2003a for exceptions). Features of the Michif nominal system have only recently begun to receive more dedicated attention in linguistic research than in the past, largely concentrating on developing and testing generative-syntactic models of Michif nominal structure (see, e.g., Gillon & Rosen 2018; Mathieu & Strader 2015, among others). This work represents the first full, monograph-length study devoted to nominal classification in Michif, and contributes to this ongoing discussion by offering a quantitative, corpus-based perspective on nominal classification as it appears in Michif discourse. While previous studies in this area have drawn primarily on elicitation and published lexical sources for their data, the data analyzed in this study are drawn from a subset of a 60-hour multimodal corpus of contemporary spoken Michif, developed by the author in collaboration with 42 members of Métis communities in Manitoba, Saskatchewan, and Alberta, Canada primarily during the period of 2011–2016.

1.1 Sociolinguistic overview

Michif is currently spoken in both western Canada and the northern United States. In Canada, it is reportedly spoken by 725 individuals, located primarily in Saskatchewan (45.5%), Manitoba (18.6%), and Alberta (15.9%) (Government of Canada 2016). However, these numbers are likely significantly inflated due to the use of this label to also refer to distinct Métis varieties of Cree and French (Papen 2005a; Rosen 2007). In the United States, there are approximately 75 speakers on the Turtle Mountain Reservation in North Dakota (Lewis, Simons & Fennig 2013). Even given the highest estimates of speakers, Michif is a critically endangered language, and would likely rank at 8b ("Nearly Extinct") of 10 on the Expanded Graded Intergenerational Disruption Scale (EGIDS, Simons & Fennig 2018), with 1 reflecting use as an international language, and 10 indicating that a language is dormant. At stage 8b, all remaining speakers are of the grandparent generation or older, with few opportunities to use the language. In addition, speakers are geographically scattered, as are potential adult language learners.

Multilingualism has traditionally been an integral aspect of Métis identity. In the late 19th and early 20th centuries, it was not uncommon for most Métis to be conversant in Cree, French, and English, and in some cases even Saulteaux, Gaelic, and/or Blackfoot. With this linguistic knowledge, Métis people were in a unique position to act as guides and interpreters for European explorers and traders in exchanges with First Nations people (Bakker 1997: 164; Rhodes 1982; Rosen & Souter 2009b; Sealey & Lussier 1975). It was this intense community multilingualism that set the stage for the emergence of Michif as a distinct language. Today, all Michif speakers are fluent in English, and

some may also speak French, Cree, and/or Saulteaux. In both Canada and the United States, intergenerational transmission of the language has ceased, and the majority of Michif speakers are sixty years of age or above.

1.2 Michif as a mixed language

Mixed languages are typically categorized on the basis of a combination of structural and sociohistorical criteria, as detailed in Meakins (2013; 2017) and Thomason (2001), among other sources. From this perspective, Michif qualifies as a mixed language because it (a) was created by bi-/multilinguals (and thus, unlike pidgins and creoles, is not driven by expressive needs on the part of its creators); (b) serves as an identity symbol for a new ethnic group; (c) consists of structural components drawn from two or more source languages; and (d) was formed through a particular set of linguistic processes (e.g., code-switching and/or borrowing, relexification, or metatypy).[2]

There are several key differences between the social factors typically encountered in the development of pidgin and creole languages and those observed with mixed languages. Creoles generally arise in situations in which one or more substrate languages and a superstrate language are in contact. Speakers of these languages find themselves in situations of prolonged inter-group contact with a need to communicate with members of the other speech communities, yet have no language in common. Creoles thus arise as a means of facilitating communication between speakers of multiple languages who cannot

[2] It is worth noting here that some scholars contest the classification of Michif as a mixed language (e.g., Rhodes 1977), the usefulness of that classification (e.g., Gillon & Rosen 2018), and even the validity of the category of mixed languages in general (see, e.g., Matras & Bakker 2003 for discussion). While there is no single, generally agreed-upon set of criteria used to determine whether or not a language falls into this category, the characterization of mixed languages most prevalent in the literature is framed around socio-historical conditions at the time of genesis, with a range of variation in potential structural outcomes (Meakins 2017).

otherwise understand one another. Critically, creole development is characterized on the whole by situations in which creole "creators" have limited access to the language(s) of their interlocutors.

The contact situation in which Michif arose would have been markedly different from those of most creole languages, where there is a politically and socially dominant European group to which speakers of other languages are made to adapt. Rather, Michif arose from intermarriage between two groups that, at the time, were likely not in a strongly differentiated power relation. There is reason to believe that the forebears of Michif who were in contact with the French fur traders would have been Ojibwe people who were fluent in Ojibwe, but also spoke Plains Cree as a second language since it was the lingua franca in the area at the time (Bakker & Papen 1996: 9). In addition, it would seem necessary for there to have been widespread community bilingualism during the formative period of a language like Michif, which maintains complex structures from both major source languages, to stabilize (Crawford 1985a: 49). Crawford elaborates:

> Another way of restating some of the above is that it must have required some sort of sympathetic co-existence or a balance of prestige between Cree and French groups to produce Michif. Whereas in most contacts between languages one language dominates, to the eventual extinction of the other, in this case [sic] at least during the period giving rise to Michif, the direction of such dominance is not clear. At least it must he [sic] accepted that the French component is strong and not typical of a language being absorbed into another… The most common pattern in language contact is for one language, often that of the intruder or colonist, to dominate over time. Thus creoles tend to become more like the dominant (often European) languages involved in them. In the Michif case that would be expected to be French. The fact that this was not the [c]ase with Michif indicates that there were factors which made the Cree language and culture, if not dominant, at least strong enough to resist assimilation to French. (1985a: 49–50).

A clear-cut dichotomy between superstrate and substrate languages is thus problematic in the Michif context, as this presumes that one source language is more

dominant than the other, which is not what is reflected in the socio-historical situation at the time of genesis.

As with many creole languages, mixed languages also emerge in situations of intense and sustained contact between multiple linguistic groups. Unlike the originators of creole languages, however, the creators of mixed languages are typically members of multi-ethnic communities and fluently speak all of the major languages present in the contact scenario. It is also claimed that mixed languages often arise as a marker of in-group identity through which speakers can differentiate themselves from neighboring groups. Unlike pidgins and creoles, mixed languages typically arise through situations of bilingualism:

> [Mixed languages]…show considerable diversity in structure, social function, and historical origins; nonetheless, they all emerged in situations of bilingualism where a common language is already present. In this respect, they do not serve a communicative function, but rather are markers of an in-group identity. (Meakins 2017: 1)

Mixed languages thus differ significantly from creoles in at least two ways: first, for mixed languages, generally only two languages are involved in the contact scenario; and, second, speakers have full access to both the so-called superstrate *and* substrate languages.

Indeed, in the historical case of the Red River Métis, symmetric bilingualism (and/or multilingualism involving other Indigenous languages and possibly English) among the members of a pervasively bilingual speech community is considered to have been the norm. Speakers of Michif at the time of genesis were fully bilingual, if not multilingual (Bakker 1997: 163–7; Bakker & Papen 1997: 352; Rosen & Souter 2009b). It would follow, then, that this may have had consequences for the resulting form of

Michif, as these speakers would have been able to bring their full knowledge of two linguistic systems to bear on this new language. Contexts of societal bilingualism or multilingualism such as that found in Michif entail radically different speaker knowledge of the syntactic and semantic properties of items in source languages from that of pidgins and creoles. In situations of societal bilingualism, speakers can be assumed to have immediate access to the forms, meanings, and combinatorics of *all* of the languages in their repertoire in which they are proficient, and not only the one or two that comprise a substrate. In turn, this implies that new languages that emerge under these conditions might incorporate syntactic and semantic information from more than one source language in a consistent way, rather than drawing on only one language (the superstrate) for phonological forms and another (the substrate) for meanings and combinatorics. The analysis of nominal classification presented in this study demonstrates that Michif has done just this, incorporating both Cree-derived animacy and French-derived gender into the language, thus maintaining two distinct systems of nominal classification. The Michif case suggests, therefore, that (a)symmetry of speakers' knowledge of the languages involved in contact situations can be of crucial importance to the results of language contact.

1.3 Language contact and the lexicon

A number of theories have been proposed within the field of contact linguistics to account for the ways in which lexicons develop in contact situations. Perhaps the most common theory is that of *relexification,* which has been described as a mechanism of language change essentially involving the replacement of the lexicon of one language

with that of another (see, e.g., Muysken 1981, among others). Bakker (1994; 1997) proposes another term, *language intertwining,* as an alternative to relexification to refer to a process of language genesis which involves the "combination of the grammatical system (phonology, morphology, syntax) of one language with the lexicon of another (1997: 203). Unlike relexification, the term "language intertwining" focuses on the combined nature of the resulting language in which both source languages have the same weight, rather than replacement of elements from one language with those from another.

Van Coetsem (2000) discusses ways in which languages in contact influence one another, drawing a distinction between borrowing and what he refers to as *imposition*. He claims that in situations of linguistic transfer, there is a source language and a recipient language. In situations of linguistic borrowing, the recipient language acts as the agent, and is the language in which the speaker is more proficient. The recipient language grammar is more stable and resistant to change. On the other hand, in situations where imposition occurs, the source language is the agent. It is also the language in which the speaker is more proficient and is more stable and resistant to change than the recipient language (Winford 2008: 126–127)

Finally, Ansaldo (2011) and Ross (2007) propose a model of *metatypy*, which is characterized as a contact phenomenon involving a "model language" and a "replica language" within a bilingual speech community. In this process, the morphosyntactic constructions of the so-called replica language are restructured to match those of the "model language", both in terms of semantics and in morphosyntactic properties (Ross 2007: 116).

1.4 Research objectives

One question central to this study is the extent to which the original nominal classification systems of both of Michif's primary source languages have been maintained. In the case of both Cree-origin animacy and French-origin gender, this entails considering the relationship of Michif animacy and gender values relative to those of their Cree and French translation equivalents, as well as whether or not Michif appears to apply the same criteria for nominal classification as is found in these source languages. Such an assessment would reveal whether animacy and gender assignment in Michif can be ascribed wholly to patterns inherited from its source languages, or whether there is evidence of divergence from these models. Understanding the role of inheritance in the development of nominal classification systems like Michif may provide insights into how systems of nominal classification transfer or develop in bilingual contact situations more generally, which may in turn be relevant to models of language contact.

This study also investigates whether or not Michif's animacy and gender systems are productive, have been lost, or have become fossilized. From a cross-linguistic perspective, the existence of a language such as Michif, which potentially has two productive systems of nominal classification, is rare. If, on closer investigation, it turns out that one or more of these systems is no longer being applied consistently to lexical items or has disappeared entirely, then the overall profile of Michif is much more in line with what has been reported for other languages around the world, thus striking a much less unusual typological profile. Conversely, if it is found that Michif does indeed have multiple productive classification systems, then this result is potentially relevant for typologies of nominal classification, where multiple co-existing systems within a

language are extrememly rare (see Chapter Three for further discussion), as well as for ongoing documentary and descriptive work involving Michif.

Similarly, this study aims to determine whether or not the nominal classification systems commonly reported for Michif have remained independent of one another, or if they have potentially been merged or lost. In theory, Michif could either be analyzed as having two independent systems of classification (i.e., sex-based gender on the one hand, and animacy on the other), or as one combined system involving all logically possible combinations of animacy and gender (i.e., animate-feminine, animate-masculine, inanimate-feminine, inanimate-masculine). In his study of nominal classification, Corbett (1991) raises the question of whether or not any clear-cut examples of languages with "two independent [nominal classification] systems" (188) can be identified cross-linguistically. In the cases that have been analyzed previously in the literature, there has not been much of any evidence to decide unequivocally between these competing analyses. Determining whether or not languages with two independent nominal classification systems exist requires some way of demonstrating characteristics that are more plausibly attributed to separate systems rather than a joint one. This study finds that Michif does indeed maintain two independent systems of nominal classification, thereby responding to some of these open questions in the literature regarding whether or not multiple systems with different classificatory bases can exist in a single language.

Finally, this study investigates how animacy and gender values are assigned to English borrowings in Michif. This particular question falls at the intersection of the above-mentioned research questions related to inheritance, productivity, and independence. English-origin forms present a particularly valuable point of evidence

concerning the dynamics of synchronic gender assignment in Michif, as these animacy and gender values cannot be attributed to inheritance from Cree or French. Borrowings such as these allow us to investigate whether or not these forms default to a particular gender and/or animacy value, or whether they may follow natural animacy and gender or some other pattern, perhaps based on the values of equivalent forms in their source languages.

Overall, the results of this study provide an example of increased grammatical complexity as an outcome of language contact. Through the corpus-based analysis described above, this study finds that both animacy and gender have been maintained as active categories in contemporary spoken Michif, applying consistently to new lexical items in ways that neither follow from natural animacy and gender nor default to a single animacy or gender value. While both of these nominal classification systems show a high degree of correlation with the animacy and gender assignment patterns of their source languages, we also find evidence of divergence from these patterns that set Michif apart from those source languages. This includes instances of semantic analogy that extend exceptional classifications to a wider subset of lexical items than is found in a given source language, as well as instances of apparent regularization, where the animacy and/or gender values of some nouns are brought more closely into alignment with the natural animacy and gender attributes of their referents.

1.5 Dissertation outline

This dissertation is organized as follows. Chapter Two provides background on Michif, including linguistic varieties and an overview of previous work on Michif. In Chapter

Three, I define and discuss the concept of nominal classification and provide an overview of relevant literature. Chapter Four provides a structural overview of animacy and gender in Michif and its source languages, while Chapter Five discusses animacy and gender assignment in Michif and its source langauges. Chapter Six outlines the methods used to assemble the corpus of spoken Michif and construct the dataset used in this study. In Chapter Seven, I use corpus data to show that both animacy and gender categories came into Michif historically and remain productive and independent. I also show that both animacy and gender categories systematically follow their source languages rather than notional animacy and gender in the vast majority of cases. Finally, in Chapter Eight, I consider the implications of these results for Michif language description, typologies of nominal classification, and Michif language documentation and lexicography, as well as directions for future research.

Chapter 2: Background on Michif

Although the language of the Métis people is commonly referred to as "Michif," this term actually encompasses a range of linguistic varieties. In this chapter, I present an overview of Michif varieties and previous research and publications concerning Michif. Section 2.1 describes the major Michif varieties, including Northern Michif (2.1.1), Michif French (2.1.2), and Southern Michif (2.1.3). Section 2.2 outlines both academic and community-level publications on Michif, beginning in the early 1970s and proceeding to the present day, with a particular focus on the Southern Michif variety.

2.1 Michif varieties

Métis people have been engaged in constant migration throughout the prairie provinces in western Canada for generations and, in fact, pride themselves on their mobility, versatility, and multilingualism (Rosen & Souter 2009b). The Métis Nation comprises a non-contiguous group of speech communities spread throughout western Canada and the northern plains of the United States. Many of these communities have experienced generations of separation, leading to the development of several varieties which can be roughly divided into three groups: Southern Michif, Northern Michif, and Michif French (Souter 2018a: 5–6).[3] While all three varieties share the same source languages, they

[3] These varieties have also been classified as "Michif-Cree" (Southern Michif), "Île-à-la-Crosse Michif" (Northern Michif), and "Métis French" (Burnouf, Fleury & Lavallée 2007). While the terms used here are not in widespread use among speakers, I follow (Souter 2018a) in adopting the terms "Southern Michif," "Northern Michif," and "Michif French" as a way of avoiding potential confusion over the varieties intended (e.g., since "Michif-Cree" has been used to refer both to the variety spoken in Île-à-la-Crosse, SK and to varieties spoken in southern Manitoba and Saskatchewan, despite notable linguistic differences between the two). I also do this to (a) recognize that what is referred to as "Île-à-la-Crosse Michif" is spoken in other communities in northern Saskatchewan as well, and to (b) capture the fact that all three varieties are spoken by Métis people and are often referred to as some form of Michif. As Burnouf *et al.* (2007) state, "…if Métis people speak [these differing varieties] and call them Michif, then by sociological reality, they are Michif languages…" (iii).

differ in the extent to which features of the source languages are represented in lexical, grammatical, phonological, and semantic domains. Some varieties are more influenced by French, while others are more influenced by Cree. What follows is a brief survey of these languages, including the extent to which contact with Cree and French is represented in each.

2.1.1 Northern Michif (Métis Cree)

Northern Michif is spoken in northern Saskatchewan (Buffalo Narrows, Ile-à-la-Crosse) and possibly in areas of northern Alberta (Bakker 1997: 144). Its speakers often refer to this variety as "Michif", but may also use the terms "Métis Cree", "French-Cree", "Île-à-la-Crosse Michif", "Michif-Cree", or "Cree-Michif", or simply "Cree". It has also been referred to as "Church-Cree" due to the introduction of a small number of French nouns to the language by Francophone priests in the nineteenth century (Burnouf, Fleury & Lavallée 2007: iii). The structure of Northern Michif differs considerably from the Southern Michif variety spoken in Manitoba, southeastern Saskatchewan, and North Dakota. First, some consider it to be influenced by Woods Cree rather than, or in addition to, Plains Cree (Burnouf, Fleury & Lavallée 2007: iii), though other sources claim a solely Northern Plains Cree origin (e.g., Bakker 1997: 148). In addition, Northern Michif has extremely limited French influence in its nominal lexicon compared to Southern Michif (cf. Ahenakew 2009; Fleury 2000), where virtually all nouns have French-origin phonological forms. Bakker (1997) also notes a lack of French-origin adverbs or prepositions. The Northern Michif lexicon thus more closely resembles Cree than French, with some instances of borrowed French nouns. This leads to limited mutual

comprehension between speakers who do not share another language in common (e.g., for Southern Michif speakers who do not also speak or understand Cree or Ojibwe, or for Northern Michif speakers who do not also speak or understand French).

Bakker (1997) observes additional phonological and morphological differences between Northern and Southern Michif. Morphologically, Northern Michif differs from Southern Michif in that it inflects for the conjunct order in main clauses, while Southern Michif does not. Northern Michif also uses the first person inclusive pronoun *kiiyanaaw*, as opposed to Southern Michif, which uses *kiiyanaan*. Additionally, Northern Michif does not verbalize French nouns in the same manner that Southern Michif does. For example, Southern Michif may verbalize French-origin nouns such as *laboo-iwan* 'it is muddy' (from French *la boue* 'mud') or, *kaa-liselibreetiichik* 'when they celebrate' (from French *célébrer* 'to celebrate'), whereas forms such as these are not used or understood in Northern Michif (Bakker 1997: 146).

Phonological differences exist between Northern and Southern Michif as well. For instance, /k/ is voiced intervocalically in Northern Michif, but is voiceless in Southern Michif. This pronunciation is closer to Cree than to Southern Michif (Bakker 1997: 146–7). Northern Michif speakers are typically also fluent in Plains Cree, while today, Southern Michif speakers rarely speak additional languages other than English. Southern Michif distinguishes between /e:/ and /i:/ (as in southern Plains Cree dialects), while these two sounds have merged into /i:/ in Northern Michif (as in Northern Plains Cree dialects). Sibilants are also pronounced as /s/ and /c/ in Northern Michif, versus /ʃ/ and /tʃ/ in Southern Michif. Based on these differences, Bakker claims that Northern and Southern Michif are most likely to have developed independently, and that Northern

Michif is therefore best considered a dialect of Cree rather than a mixed language (Bakker 1997: 146–7).

2.1.2 Michif French (Métis French)

Michif French is a dialect of French developed and spoken by the Métis people in Manitoba and Saskatchewan. Speakers may also refer to this variety as simply "Michif" or "French". According to Bakker, it is this variety of French that served as one of the source languages for Michif, as opposed to standard French or Canadian French (1997: 81). Michif French is a dialect of French with some Algonquian influence, rather than a dialect of Cree (Rosen & Souter 2009b). It exhibits several phonological, lexical, and syntactic differences from Canadian French that have been claimed to be the result of Cree influence (Bakker & Papen 1997: 301; see also Papen & Bigot 2010; Rosen & Lacasse 2014 for discussion of other aspects of Michif French).

Relatively little is known about this variety of Michif. A few basic descriptions exist (e.g, portions of Bakker 1997; Papen 1984a; Papen 1984b, among others), as well as a short introduction to the language self-published by a group of Michif French speakers (Millar et al. 2016), although there have been no in-depth studies to date. As is the case with Métis Cree, further documentation and analysis are needed to determine ways in which Michif French may differ from Southern Michif, Northern Michif, and Canadian French.

2.1.3 Southern Michif (Michif)

Typically, the term "Michif" refers to the contact language described as having Cree-origin verbs and French-origin nouns, rather than the variety of Cree with French influences, as described in 2.1.1, or the variety of French with Cree influences, as described in 2.1.2. It is what linguists tend to refer to when discussing "Michif." Further, in studies of mixed languages, it tends to be a reference point against which other mixed languages are measured. Most documentation and linguistic description have focused on this variety (e.g., Bakker & Papen 1997; Crawford 1976; Laverdure & Allard 1983; Rhodes 1977; Rosen 2007), while those spoken elsewhere have received comparatively little attention. The literature on Michif mentions communities of speakers in San Clara and Boggy Creek, Manitoba; Camperville and Duck Bay, Manitoba; and St. Lazare, Manitoba in Canada, as well as on the Turtle Mountain Reservation in Belcourt, North Dakota in the United States (Crawford 1985a). I have additionally worked with Southern Michif speakers from Ste. Madeleine and The Corner (near Binscarth, Manitoba) and Crooked Lake, Saskatchewan. This variety was also once reportedly spoken in Métis communities in Montana, Minnesota, and Wisconsin, although it appears that there are no longer any first-language speakers in these areas (Bakker & Papen 1997: 357–358; Peterson 1978). Speakers may also refer to this variety as "Michif-Michif", "Cree", or "Michif-Cree". Broadly speaking, nouns and noun phrases of this variety are derived from French, while verbs and verb phrases are derived from Cree, making Southern Michif unlike any other contact language (Meakins 2013). From this point forward,

unless otherwise specified, when I refer simply to "Michif", it is the Southern Michif variety to which I am referring.

2.2 Previous linguistic work

Linguistic research on Michif began in the early 1970s with John Crawford at the University of North Dakota.[4] He worked primarily with speakers of Southern Michif from the Turtle Mountain Reservation near Belcourt, North Dakota. Crawford's earliest publications focus mainly on describing the general nature of this previously unknown language to determine how it should be classified. Crawford (1973) provides a brief introduction to Turtle Mountain Michif, addresses dialect variation, and discusses Michif's relationship to its Algonquian source languages. Crawford (1976) outlines major linguistic influences on Michif and also touches briefly on the issue of language endangerment. Crawford (1979) discusses challenges in developing a standardized Michif orthography, while Crawford (1985b) provides a very brief look at dialect variation. Working with Patline Laverdure and Ida Rose Allard, both fluent speakers of Turtle Mountain Michif, Crawford also edited the first Michif dictionary (Laverdure & Allard 1983), an important source of information for many subsequent studies (e.g., Bakker 1997; Wolfart 2010). Crawford (1985a) reports on Michif language use in four Métis communities (Belcourt, ND; San Clara and Boggy Creek, MB; Camperville and Duck Bay, MB; and St. Lazare, MB). Finally, Crawford (1985c) explores whether Michif would best be classified as "nothing worthy of note", a creole language, a dialect of Cree,

[4] The focus of this section is on Southern Michif, though there is a larger literature associated with other varieties of Michif. See, for example, Ahenakew (1997) and Hogmen (1981) on Northern Michif, and Douaud (1980; 1989) and Papen (1984a; 1984b; 1993) on Michif French.

or a mixed language. While eliminating the first possibility, Crawford concludes that more research would be necessary to arrive at a definitive classification. Crawford's research on Michif also resulted in a substantial collection of unpublished audio recordings which are now housed at the University of North Dakota (Pasch 2013). In addition to his own work, Crawford also supervised several graduate students who conducted linguistic research on Michif. Their work is discussed below.

Richard Rhodes also began work on Michif in the 1970s on the Turtle Mountain Reservation in Belcourt, North Dakota. In his first publication on Michif (1977), Rhodes provides a 20-page morphosyntactic sketch of the language, which includes inflection charts for all four Algonquian-derived verb classes (Inanimate Intransitive, Animate Intransitive, Transitive Inanimate, Transitive Animate) in the independent, conjunct, subjunctive, and imperative modes. Based on this analysis, Rhodes suggests that Michif is a dialect of Plains Cree which borrows heavily from French, rather than a language in its own right. However, Rhodes (1986) retracts this view, and instead argues that Michif is indeed a mixed language. Other publications include Rhodes (1987), which discusses Métis myths, and Rhodes (2001), which addresses differences between narrative strategies found in Michif and Plains Cree texts. Finally, Rhodes (2008) discusses Ojibwe influence on the Cree component of Michif, while Rhodes (2009) describes the historical development of Michif from a phonological perspective.

Peter Bakker has published the most extensively on Michif. His earliest publications on the language appeared in the late 1980s and early 1990s, and primarily concern the genesis and classification of the language (e.g., Bakker 1989a; Bakker 1989b; Bakker 1990). Bakker (1991) discusses Ojibwe influence on Michif, while a brief

overview of the language in the context of mixed languages is found in Bakker (1994). A handful of short texts (representing both Northern and Southern Michif varieties) with English translations can be found in Bakker (1996; 1997), while Bakker et al. (1998) provides a Michif story with an accompanying translation into Dutch. A number of grammatical features are also outlined in Bakker's chapter on Michif in the *Atlas of Pidgin and Creole Language Structures* (Michaelis et al. 2013). Bakker (2011) provides an overview of literary works published in the Michif language and discusses the choice of languages and orthographies used therein.

Bakker has also contributed to a number of Métis community publications. Bakker (2004a) provides a very brief introduction to the language, with a particular focus on Michif's typological uniqueness, while Bakker (2004b) provides an overview of the various Michif orthographies that have been used. Bakker (2004c) provides a basic introduction to the Michif verb, highlighting differences between Cree and Michif inflection. Also included in this chapter are sample verb paradigms and a list of verbal affixes and their meanings. Bakker (2004d) discusses the ambiguity of the term "Michif", and briefly outlines the different linguistic varieties with which the label may be associated. Finally, in Bakker & Barkwell (2004), the authors provide three Michif texts with English translations—"Why Bears Have No Tail," "The Trapper and the Wolf," and "The Three Bears".

The most detailed account of Michif thus far is found in Bakker (1997). This work, based on Bakker's dissertation research (1992) and fieldwork in the late 1980s and early 1990s, represents the first major publication on Michif. The volume presents a lengthy discussion on the historical emergence of the Métis and the genesis and

development of Michif. It also profiles Michif-speaking communities throughout Canada and the United States, and touches on linguistic differences between them. Also included is a brief structural overview of the language, which characterizes Michif in relation to its source languages. As the first monograph-length study of Michif, Bakker (1997) represents a significant advancement in linguistic and historical research on the language. Nevertheless, the linguistic description that Bakker provides is brief, constituting only 40 of the volume's 316 pages, and is based primarily on data from a translation-based questionnaire and example sentences taken from Laverdure & Allard (1983). A handful of short Michif texts (ca. three pages) are also included in a specialized orthography developed by Bakker, but do not appear to have been drawn on significantly in his linguistic analysis. As a result, many grammatical topics are mentioned only briefly (e.g., verbal paradigms, for which no full examples are included), while others (e.g., derivation) are not addressed at all. While its grammatical treatment of the language is thus somewhat limited, Bakker (1997) nonetheless represents a landmark in studies of Michif, and provides a solid foundation for continued research.

As with Peter Bakker, Robert Papen also began linguistic research on Michif in the late 1980s, focusing largely on issues in Michif phonology, the genetic classification of Michif, and the development of Michif orthographies. Both Papen (1986) and (2003b) contribute to the debate over whether Michif is comprised of one or two phonological systems. Further publications include Papen (1987a), which addresses variation among Michif speakers; Papen (1987b), which describes selected lexical and phonological processes in Michif; and Papen (2003a), which discusses gender and the classification of Michif among language types. Papen (2004) proposes an orthography for Michif and

discusses its motivations. Papen (2005a) also addresses Michif orthography development, but includes a roughly two-page overview of Michif grammar as well, and Papen (2005b) provides a general overview of Michif for French-language audiences. Using a corpus of Michif French developed by Guy Lavallée, Papen & Bigot (2010) analyze the use of three unusual verb forms in the third person plural—*sontaient* (*être*), *ontvaient* (*avoir*), and *fontsaient* (*faire*). Finally, Papen (2011) describes the phenomenon of liaison in Michif, asserting that it is still productive in Michif today, while Papen (2017) revisits the split phonology hypothesis in Michif.

Together with Peter Bakker, Robert Papen has authored two other publications on Michif. The first, Bakker & Papen (1996), provides information on the geographical distribution of Métis languages. The second, Bakker & Papen (1997), discusses the historical origins of Michif, but also offers a four-page structural sketch of the language, in which grammatical processes of French and Cree origin are treated separately. While this sketch provides a useful point of departure for future linguistic investigations, the analysis is influenced by the authors' assumption that the Cree and French components of Michif operate under two separate sets of rules, rather than viewing and describing the language as a single coherent system.

Significant work on Michif has also been produced by Nicole Rosen, beginning in the early 2000s. Rosen (2003) describes the patterning of demonstratives in Michif, and argues that the language can be analyzed as a single system, rather than positing two grammars to account for the data. Rosen (2006) provides the first analysis of Michif word stress, and compares it against that of its source languages, showing that Michif's system is distinct from both. In her doctoral dissertation (Rosen 2007), Rosen provides the first

systematic phonological description of Michif. One chapter also includes a 68-page sketch of Michif morphology. An appendix providing a list of Michif verbal affixes and their meanings is also included. Based on her findings, Rosen argues that Michif is comprised of a single merged phonological system, rather than two separate phonological systems which are stratified according to the source languages. Rosen (2008) uses Michif as a case study to investigate potential outcomes of language contact, with a particular focus on conflict sites between Cree and French grammars in phonological, morphological, and syntactic domains. Rosen & Souter (2009b) discuss the challenges of Michif language revitalization due to multilingualism, linguistic variation, and geographical distance, while Rosen & Lacasse (2014) provides a comparison of back vowels between Michif French and Manitoba French.

More recently, Rosen has conducted lexical research in Manitoba Métis communities towards a digital dictionary of Michif (Rosen 2016b). Rosen has also contributed Michif lexical items and short phrases to the Algonquian Linguistic Atlas, a multimedia atlas of Algonquian languages available online (2018a), and has recently produced work on the Michif noun phrase, which is discussed further below.

Rosen has also contributed to a number of community-level publications. For example, Rosen (2004a) discusses challenges in adult Michif language instruction, while Rosen (2004b) provides a learner-oriented description of the Michif stress system. Rosen is also a contributor to Barkwell et al. (2004), which provides a partially annotated bibliography on a wide range of Métis linguistic and cultural topics. Finally, Rosen & Souter (2009a; Rosen & Souter 2015) provide a teaching guide for a 12-week Michif language course for adult learners which covers a number of cultural and grammatical

topics and includes vocabulary lists, explanations of key grammatical points, communicative tasks, and practice exercises.

Notable contributions have also been made by graduate students of John Crawford at the University of North Dakota. Boteler (1971) examines syncretic medical practices among the Métis of Turtle Mountain, and includes the names of herbs used in the treatment of illness in Michif, Cree, and Ojibwe. Peske (1981) discusses the historical origins of the French component of Michif. Both Weaver (1982) and (1983) examine obviative marking in Michif, with the former study investigating how the general absence of Cree nouns in Michif has affected the proximate-obviative distinction, while the latter examines the effects of language attrition on this process. Evans (1982) investigates the claim that Michif is comprised of two separate, co-existing phonological systems, and concludes that the two systems appear to be in the process of converging, while Andrella (1983) comes to the opposite conclusion. Speers (1983) provides a Michif narrative with English translation and analyzes its structure, with a particular focus on the introduction of new information. Lovell (1984) examines Michif reflexive clauses from the perspective of Relational Grammar, while Wildeman (1989) describes code-switching in Michif. A collection of unpublished notes containing basic grammatical information was also produced through a field methods course on Michif by students at the Summer Institute of Linguistics (University of North Dakota Session) (Bitterman et al. 1976).

A number of scholars outside of North Dakota have published on Michif as well. Weston (1982) investigates benefactive and dubitative constructions in Michif, while Orser (1984) discusses its genetic classification. In addition, Wolfart (2010) discusses Michif negation strategies, and both Prichard & Schwayder (2014) and Fitzsimmons et al.

(2015) return to the question of whether Michif's phonology is stratified according to historical source. Sammons (2013a) provides a Michif historical narrative about one family's forced departure from their community with both free and literal translations, while Sammons (in revision) provides an analysis of applicative constructions in Northern Michif. Kitaoka and Strader (2016a; 2016b) examine the structure of Michif relative clauses. Other work by Strader and Kitaoka has focused on discontinuous elements in Michif, including Strader (2013) and Kitaoka & Strader (2016a; 2016b; 2017a; 2017b; 2017c). There have also been a number of presentations and publications on the phonetics of Michif in recent years. These include Rosen, Muehlbauer, and Lacasse (2010a; 2010b), Rosen & Brodner (2012), and Rosen, Stewart, and Sammons (2015; 2016a; 2016b; in prep), among others.

Several recent works in the area of Michif language documentation and revitalization have also begun to emerge. While McCreery (2013) focuses on challenges in the acquisition of Cree by adult learners, parallels are drawn to Michif, and potential means of addressing these challenges are explored. Sammons (2013b) and (2013c) both discuss the role of distance in Michif language documentation and revitalization, while Sammons (in revision) uses examples from Michif to discuss the role of conversation in language documentation. Mazzoli (in press) uses Michif as a case study to explore some of the challenges related to collaborative language research and language revitalization. Finally, Souter (2018a) describes the process of developing a set of Michif language video resources to support independent adult language learning which are available online (Souter 2018b).

There has also been an upsurge recently in work on the noun phrase in Michif in general, and on gender in particular. Browne (2005), a Master's thesis, discusses the relexification process in Michif. Sammons (2013d) examines Michif possessive constructions. Gillon (2015; 2016) investigates plurality in Michif and Innu-aimun, an Algonquian language, while Gillon & Rosen (2016) examines the mass-count distinction in Michif. Strader (2014) is a Master's thesis analyzing the structure of the Michif determiner phrase, while Strader (2016; 2017a) investigate adjectival agreement in Michif. Most relevant for the purposes of this study are several recent presentations and publications specifically investigating gender in Michif. These include, for example, Stoltzfus & Boissard (2016), which argues that the French morphosyntactic components of Michif, including masculine/feminine gender, are fossilized. In addition, Mathieu & Strader (2015) and Strader (2017b; 2017c) propose that Michif nominal syntax is distinct from that of its source languages, while Gillon & Rosen (2015a; 2015b; 2018) and Rosen & Gillon (2015; 2017) use evidence from gender to argue that Michif is an Algonquian language with extensive French borrowing, rather than a mixed language. Sammons (2017) presents preliminary results from this study about nominal classification in Michif. Previous research on grammatical gender in Michif is discussed further in Chapter Five.

Outside of the academic context, many community-level publications have also emerged in recent years. The majority of these materials have been produced through regional Métis organizations in Canada and their associated research institutions (e.g., Gabriel Dumont Institute in Saskatchewan, Louis Riel Institute in Manitoba), although individual authors have occasionally published through other venues, as well. These can

be grouped into pedagogical materials, children's books, resource guides, and other miscellaneous publications:

1. *Pedagogical materials:* In addition to those mentioned above, current pedagogical materials for Michif include introductory language lessons (Bakker & Fleury 2004; Flamand 2002; Fleury 1999a; Gordey & Fleury 2011) and basic primers (Fleury 1999b; Fleury 2000).
2. *Children's books:* A number of bilingual children's books translated into Michif from English have also appeared within the last decade (e.g., Burton & Patton 2011; Panas & Whitford 2004; Patton & Burton 2007; Pelletier 2007, among others).
3. *Resource guides:* Additional resource materials in the form of annotated bibliographies (Barkwell, Dorion & Préfontaine 2001; Barkwell, Fleury & Morin 2017) and general resource guides are also available. In particular, recent edited volumes by Barkwell (2004a; 2004b) include contributions from both Michif speakers and academic scholars (see above for contributions by Bakker, Papen, and Rosen) on topics such as conversational phrases, vocabulary items, basic grammar, language teaching methods, writing systems, songs, storytelling, and more. Another important resource is Burnouf et al. (2007), which provides the same sets of words and phrases translated into each of the three Michif varieties, allowing the reader to systematically compare them and gain a sense of their similarities and differences.
4. *Miscellaneous publications:* Other resources of note include a book of Michif prayers, written by speaker Grace Ledoux-Zoldy and transcribed by Arthur J.C.

Schmidt (Ledoux-Zoldy 2010); a book of personal reminiscences with partial translations into Michif (St. Pierre 2012); a card game (Souter 2008); and even a graphic novel in Michif (Fleury et al. 2008).

5. *Electronic resources:* An increasing number of Michif language resources are becoming available online. The Virtual Museum of Métis History and Culture (Gabriel Dumont Institute of Métis Studies and Applied Research 2018) provides access to a wealth of historical materials related to Métis language and culture, including recordings of oral histories, archival documents, and photographs. The virtual museum also links to a number of new learning resources that have been developed through the Gabriel Dumont Institute of Métis Studies and Applied Research, including a Michif dictionary, language lessons, mobile apps, and more. In addition, the Michif Internet Resource Center (Pyle 2018) assembles links to a number of Michif language resources available online, including a Michif language Facebook group, a podcast providing Michif language lessons (McCreery 2012), a YouTube channel showcasing videos of conversations between Elders in Michif, and links to a number of websites. This site also has a blog component providing Michif language lessons and discussing other topics related to Michif language and culture.

While the above description is by no means exhaustive, it nevertheless gives a sense of the degree of engagement with issues of language endangerment and revitalization in communities throughout the Métis diaspora.

Michif has also attracted considerable interest from the wider linguistic world, and is often cited in the literature on language contact and introductory linguistics textbooks as a classic example of a mixed language. However, these references are typically limited to a cursory introduction and rarely go beyond a brief description of Michif's typologically unusual mixture of French nouns and Cree verbs (see, e.g., Matras 2009; Thomason 2001; Thomason & Kaufman 1988). Michif has also been at the center of considerable debate over the validity of mixed languages as a linguistic classification distinct from pidgins and creoles. An entire volume, *The Mixed Language Debate: Theoretical and Empirical Advances*, is devoted to this discussion, and includes numerous references to Michif (Matras & Bakker 2003). While thoroughly engaging with these debates over classification is beyond the scope of this dissertation, by providing more concrete data on Michif, this project may further inform these debates in the future.

Chapter 3: Nominal classification

Michif has been claimed to have two classificatory systems for nouns, a gender-based system inherited from French, which distinguishes between masculine and feminine nouns, and an animacy-based system inherited from Cree, which distinguishes between animate and inanimate nouns. These distinctions can be seen in the examples below:

(1) awa la fii

awa	la	fii
DEM:AN:SG	the:FEM:SG	girl

'this girl$_{\text{FEM:AN}}$'

(Grace Zoldy; 2012–07–13)

(2) li kaab ooma

li	kaab	ooma
the:MASC:SG	rope	DEM:IN:SG

'this rope$_{\text{MASC:IN}}$'

(Mary Fleury; 2013–09–13)

In (1), the demonstrative *awa* indicates that the referent noun *fii* 'girl' is animate, while the article *la* indicates that the referent is feminine. Similarly, in (2), the use of the inanimate demonstrative *ooma* and the masculine article *li* indicates that the referent is both inanimate and masculine. The French-origin system corresponds to what is commonly known as "gender", whereas the animacy system may or may not be analyzed as such, depending on the scholar.

In this study, I refer to both of these classificatory systems as types of "nominal classification". Matthews (2014) defines nominal classification, also know as noun class or grammatical gender, as a system "in which a class to which a noun is assigned is reflected in the forms that are taken by other elements syntactically related to it" (Comrie 1999: 458). That is, nouns belonging to a particular noun class trigger grammatical

agreement with neighboring elements within the clause. Both animacy and gender in Michif can be considered to be types of nominal classification, as all nouns must belong to both an animacy class and a gender class, and agreement with other clausal elements is required for both categories.

There is some debate in the literature as to whether or not grammatical gender and noun class can be considered to be the same phenomenon. The terms "nominal classification" or "noun class" and "gender" or "grammatical gender" have often been used interchangeably, depending on the linguistic tradition (Aikhenvald 2000: 19). There is also a lack of consensus among scholars as to whether non-sex-based systems involving animate and inanimate values should be considered to be grammatical gender systems in the same right as sex-based, masculine/feminine systems (e.g., Mel'čuk 2008: 306). For instance, Ibrahim (1973) claims that masculine and feminine are the only true types of gender values, and that animate and inanimate values should be considered "subgenders". This position has been contested by more recent scholarship, however, which questions the need for drawing any such distinction between sex-based noun categories and animate/inanimate categories (Corbett 1991: 5). Aikhenvald (2000) further states:

> Since gender systems show some correlation with sex, many non-linguists (and a few linguists) erroneously confuse 'linguistic' gender and sex. However, sex represents biological categorization, and gender represents grammatical categorization. (19)

In the case of Michif, whether or not both animacy and gender systems are considered gender, noun class, or a combination of the two depends largely on how these terms are defined. In this chapter, we examine three key perspectives on nominal

classification and their implications for Michif (3.1). We also examine the types of values found in languages with nominal classification systems (3.2) and discuss common types of assignment systems found cross-linguistically (3.3). Finally, Section 3.4 outlines three cases of languages which can be analyzed as having combined nominal classification systems, including Michif.

3.1 Perspectives on nominal classification

This section provides an overview of some recent scholarship on nominal classification, focusing in particular on that of Greville Corbett, Igor Mel'čuk, and Alexandra Aikhenvald. This section also identifies points of commonality and difference among these positions, and evaluates how they apply to Michif. Before proceeding, it is important to note that in many of the sources discussed below, the term "gender" is used both to refer to systems of nominal classification in general, and to certain specific kinds of nominal classification systems, protypically those involving masculine/feminine, and possibly neuter values. While I retain these authors' individual uses of the term "gender" when discussing their work in the sections below or directly citing them, for the sake of clarity, in the remainder of this work I make a distinction between "nominal classification", which is used to refer to the grammatical category as a whole, and "gender" as a specific type of nominal classification system involving masculine and feminine values.

Though "gender" and "noun class" have been used as distinct terms in the past, Corbett (2013a) argues that there is no substantive difference between the two and that the different terms are merely a result of different linguistic traditions. He uses the terms

"gender" or "grammatical gender" to refer to any type of nominal agreement class, whether it is based on sex-based distinctions or non-sex-based distinctions. Rather than a particular type of semantic basis, Corbett maintains that agreement is the main criterion by which grammatical gender is determined:

> The defining characteristic of gender is **agreement:** a language has a gender system only if we find different agreements ultimately dependent on nouns of different types. In other words, there must be evidence for gender outside the nouns themselves. (2013a)

This agreement with the noun can be realized through adjectives, determiners, verbs, demonstratives, anaphoric pronouns, adverbs, and numerals. Adopting agreement as the defining characteristic of gender allows for languages that show no overt marking of gender on the noun to nevertheless be treated as having gender. Two nouns can be considered to have the same gender value if they take the same agreement marking in different constructional contexts (Corbett 2013a). Likewise, two nouns can be considered to belong to different gender classes if they trigger different forms of agreement, even while other properties of the noun, such as case and number, remain the same (Corbett 2006: 750; Kibort & Corbett 2008). Consider the following examples from Russian:

(3)

a.	žurnal-Ø	leža–l–Ø	na	stole.
	magazine$_{MASC}$–NOM:SG	lay–PAST–MASC	on	table
	'the magazine lay on the table'			
b.	knig–a	leža–l–a	na	stole.
	book$_{FEM}$–NOM:SG	lay–PAST–FEM	on	table
	'the book lay on the table'			

c. pis'm–o leža–l–o na stole.
letter$_{NEU}$–NOM:SG lay–PAST–NEU on table
'the letter lay on the table'

(Corbett 2013a)

In the examples above, we see that the three sentences are virtually identical except for the nouns and their corresponding verbal agreements, which change based on the gender of the noun being used, as do the case and number endings on the noun. In (3)a, the lack of an overt feminine subject on the verb indicates that the subject *žurnal* 'magazine' is masculine. In (3)b, the verb takes the agreement suffix *-a*, indicating that the subject *kniga* 'book' is feminine. In (3)c, the verb takes the neuter agreement suffix *-o*, indicating that the subject *pis'mo* 'letter' is neuter. These instances of verbal agreement with the noun thus provide evidence of a gender system in Russian (Corbett 2013a).

In Michif, both animacy and gender classes trigger agreement with the noun, as shown in (4)–(7):

(4) la pchit fii awa
the:FEM:SG little$_{FEM}$ girl DEM:AN:SG
'this little girl'

(Mervin Fleury; 2013–09–13)

(5) li pchi gaa awa
the:MASC:SG little$_{MASC}$ guy DEM:AN:SG
'this little guy'

(Mervin Fleury; 2013–09–13)

(6) ta meezoñ ooma
2SG.POSS:FEM:SG house DEM:IN:SG
'your house here'

(Verna DeMontigny; 2013–09–13)

(7) li kaab ooma
the:MASC:SG rope DEM:IN:SG
'this rope'

(Mary Fleury; 2013–09–13)

In (4), the noun *fii* ‘girl’ occurs with the feminine article *la*, showing agreement for gender, as well as well with the animate demonstrative *awa*, showing agreement for animacy. This is in contrast to (5), in which the noun *gaa* ‘guy’ occurs with the masculine article *li* as well as the animate demonstrative *awa*, indicating that it is both masculine and animate. In (6), the noun *meezoñ* ‘house’ occurs with the feminine possessive *ta* ‘your’, as well as the inanimate demonstrative *ooma*, showing that it is both feminine and inanimate. Meanwhile, in (7), the noun *kaab* ‘rope’ is both masculine and inanimate, as evidenced by its occurrence with the masculine article *li* and the inanimate demonstrative *ooma*. Thus, we find that in Michif, both animacy and gender trigger agreement phenomena with the noun. From Corbett’s perspective then, which considers gender or nominal classification to be a type of nominal agreement class, both animacy and gender qualify as a type of nominal classification and can each be treated as such in their own right.

Mel’čuk (2008) is in agreement with Corbett in grouping gender and noun class together, stating that “gender and noun class are not conceptually identical, but they are intuitively close enough.” In other words, both gender and noun class are conceptually and formally similar enough to each other to be treated together for some purposes. While there is no clear-cut division between gender and noun class, Mel’čuk identifies eight features which can generally be used to distinguish the two. The former option in each of these criteria is generally associated with gender, and the latter with noun class:

1. The small/large number of classes.
2. Relevance/irrelevance of biological sex as the basis of classification.
3. Absence/presence of semantic motivation for classification.
4. Absence/presence of an autonomous and non-cumulative marker (of the agreement class) in the noun.
5. Absence/presence of an autonomous and non-cumulative marker reflecting the agreement class of the controller noun in the target (= agreeing) word-form.
6. Relevance/irrelevance of classification to the formal aspect of the noun's inflection.
7. Autonomy/non-autonomy of classes with respect to inflectional meanings (first of all, with respect to grammatical number).
8. Autonomy/non-autonomy of classes with respect to derivational meanings (for instance, with respect to diminutivity). (Mel'čuk 2008: 323)

More specifically, Mel'čuk asserts that the agreement classes in a language can be considered to be gender if the following conditions are met:

1. The number of these classes is small: 2 to 4.
2. They manifest a direct link with the biological sex of the being denoted by the noun: a noun referring to a male belongs to one class and that referring to a female to another class.
3. Beside the sexual division, these classes do not show a sufficiently visible semantic motivation: in most cases, there is no direct link between the meaning of a noun and its gender**1**.
4. These classes do not have an autonomous and non-cumulative marker in the noun: gender**1** is not expressed in a nominal wordform by a special morphological means, for instance, by an affix which exclusively expresses gender**1**. (Gender**1** is rather a ***covert category***.)
5. The corresponding markers in wordforms that agree with the noun are cumulative: gender**2**, which reflects gender**1**, is expressed, as a rule, in combination with other grammemes, such as number or case.
6. These classes are relevant for the formal aspects of the noun's inflection: the choice of particular number/case affixes depends on the agreement class of the noun.
7. A change in the agreement class of a noun is not used in **L** to express an inflectional meaning characterizing this noun – for instance, grammatical number — nor does it systematically accompany the expression of an inflectional meaning by a separate linguistic sign.
8. A change in the agreement class of a noun is not used in **L** to express a derivational meaning characterizing this noun – for instance, diminutivity. (Mel'čuk 2008: 324–25)

Algonquian languages, in which nouns are divided into animate and inanimate agreement classes, do not meet Condition 2 above, since the division is not based on biological sex. Nevertheless, Mel'čuk claims that the agreement classes in Algonquian languages satisfy the remaining seven conditions, and thus can be considered to have grammatical gender (Mel'čuk 2008: 331). Below is a summary of Mel'čuk's criteria as they apply to Michif:

1. There are two values for animacy (animate, inanimate), and two values for gender (masculine, feminine).
2. Masculine/feminine gender manifests a link with the biological sex of animate referents. Animacy shows no link with the biological sex of the referent.
3. For inanimate referents, there is no direct link with biological sex.
4. Both animacy and gender are covert categories in Michif, with no overt marking on the noun.
5. Agreement for both animacy and gender is reflected in other wordforms, in combination with other markers, such as number and person.
6. Nouns in Michif must occur with either articles or possessive markers, which indicate both masculine/feminine gender and number. They may optionally occur with demonstrative pronouns, which indicate both animacy and number.
7. A change in either the animacy or gender class of a noun is not used in Michif to express an inflectional meaning characterizing the noun. Both animacy and gender marking are expressed through linguistic signs which also convey features such as number and person, and thus do not require separate linguistic signs.
8. A change in either the animacy or gender class of a noun is not used in Michif to express a derivational meaning characterizing this noun.

Thus, as shown here, both animacy and gender in Michif sufficiently satisfy the criteria outlined by Mel'čuk to be considered types of gender or noun class.

Aikhenvald (2017) also claims that "gender" and "noun class" are simply different terms that arose through different linguistic traditions, but which refer to the same linguistic phenomenon. Specifically, Aikhenvald notes that the term "gender" was often used in descriptions of Indo-European and Semitic languages which typically have masculine and feminine categories, while the term "noun class" eventually came into use for other linguistic systems with agreement classes involving semantic properties other

than masculine or feminine gender (e.g., humanness, animacy, shape) (362–3). However, Aikhenvald differs from the scholars mentioned above in claiming that both genders and noun classes together can be characterized as one type of *noun categorization device*, a morpheme "which occur[s] in surface structures under specifiable conditions, denoting some salient semantic characteristics of the entity to which an associated noun refers" (Aikhenvald 2017: 361). Other types of noun categorization devices include noun classifiers, numeral classifiers, classifiers in possessive constructions, verbal classifiers, locative classifiers, and deictic classifiers (2017: 361–2), though for our purposes here we focus solely on genders and noun classes. Aikhenvald (2017) outlines several additional cross-linguistic properties of genders/noun classes. These are summarized below:

1. The number of classes in a language is always restricted and countable.
2. If a language has a gender system, every noun in that language must belong to one or more noun classes. This is not the case for other types of noun categorization devices.
3. While other types of noun categorization devices are assigned to nouns purely on a semantic basis, nouns may be assigned to noun classes based on semantic, morphological, or phonological grounds, or some combination of these.
4. Gender is manifested through obligatory agreement with either a predicate or modifier that co-occurs with the noun (adapted from Aikhenvald 2017: 363).

These criteria are generally in line with the definitions and conditions for establishing gender advanced by Corbett and Mel'čuk, though some points merit further discussion. In particular, point 3 specifies that gender values are assigned on semantic, morphological, and/or phonological grounds. However, this criterion does not leave open the possibility that the gender values for some nouns may be completely arbitrary and/or a result of diachronic processes which are not recoverable in the synchronic grammar. Aikhenvald goes on to state that "[n]oun class or gender assignment is always linked to the meaning

of a noun: it will include humanness, animacy or sex" (366). This is in line with Corbett, who maintains that there is always at least some degree of semantic basis within a gender system, though it may not apply equally well to all nouns, particularly in cases where gender assignment appears to be arbitrary. In genral, however, these criteria apply fairly well to both animacy and gender in Michif:

1. The number of animacy classes is restricted to two (animate, inanimate), as is the number of gender classes (masculine, feminine).
2. Every noun in Michif must belong to both an animacy class and a gender class.
3. Michif nouns appear to be assigned their animacy and gender values at least partially on semantic grounds, though this does not hold in all cases.
4. Both animacy and gender agreement is obligatorily marked with other elements with which the noun occurs, including articles and possessives (for gender) verb stem selection, verbal inflection, and demonstrative and interrogative pronouns (for animacy).

In sum, the criteria for gender/noun class laid out by Aikhenvald applies to both animacy and gender categories in Michif.

In Michif, both Cree-origin animacy, which has animate/inanimate values, and French-origin gender, which has masculine/feminine values, can be treated as forms of nominal classification according to the criteria outlined by the authors above. Both systems trigger agreement of the noun with other forms related to it, have a limited number of values, and apply to all nouns. While Mel'čuk makes an exception to his criterion that sex-based distinctions are required in order for such a phenomenon to be considered gender/noun class, both Corbett's agreement-focused definition and Aikhenvald's emphasis on exhaustive categorization unproblematically allow for a treatment of animacy as a form of grammatical gender and/or nominal classification.

Before proceeding, a further note about terminology is in order. Some preceding studies of Michif have adopted terms to describe natural animacy and gender and grammatical animacy and gender that differ from the ones employed in this dissertation. In the following discussion, we adopt the label "arbitrary" applied to Michif by Gillon & Rosen (2018) to refer to animacy and gender values that are not in alignment with natural animacy or gender. At the same time, where this study uses the terms "natural gender" and "natural animacy" to refer to the real-world attributes of nominal referents, Gillon & Rosen (2018) instead refer to these same classes as "semantic gender" and "semantic animacy". These differences in labels bear noting here to facilitate comparison between the results of this study and those reported in previous research.

3.2 Nominal classification values

In languages with nominal classification systems, nominal agreement classes are based to at least some degree on semantic grounds concerning the physical properties of the noun such as biological sex, animacy, and humanness, but also occasionally shape, size, and extent (Aikhenvald 2017: 363). Nouns in a given class are associated with a particular value, and can be broadly grouped into sex-based and non-sex-based gender systems.[5]

Sex-based systems are by far the most common type of nominal classification system found cross-linguistically (Corbett 2013b; Ibrahim 1973: 70). In such systems, nouns are divided into classes whose values are generally associated with the biological sex of their animate referents, although inanimate nouns in such languages may also be assigned values on an arbitrary basis. In languages such as French, for example, all nouns

[5] In sex-based systems, nominal classification values are generally found to coincide with the biological gender of the referent, though this does not apply in all cases.

are divided into masculine and feminine gender classes. Nouns in the masculine class are prototypically used to denote male human or animal referents (e.g., French *le garçon* 'the boy' *le père* 'the father', *le frère* 'the brother', *le coq* 'rooster'), whereas nouns in the feminine class prototypically refer to female human or animal referents (e.g., French *la fille* 'the girl', *la mère* 'the mother', *la soeur* 'the sister', *la poule* 'chicken'). While these semantic criteria form the basis of such categories, other inanimate and abstract nouns may be assigned to either category based on their formal properties, or arbitrarily (Kibort & Corbett 2008). Thus, French assigns nouns such as *le gouvernement* 'government' and *le fromage* 'cheese' to the masculine class, based on their morpho-phonological properties (i.e, nouns ending in *-ment* and *-age* tend to be masculine), even though neither noun has an inherent association with masculine or feminine sex (Corbett 1991: 57–62).

While a binary distinction between masculine and feminine is common among sex-based systems, other values and forms of organization are also attested cross-linguistically. In addition to having masculine and feminine gender classes that are prototypically associated with male and female referents, some languages have an additional neuter class that comprises nouns with no obvious sex value. As in binary sex-based systems, nouns can also be assigned to the neuter class arbitrarily or on the basis of their formal properties. Examples of nouns with neuter gender include German *das Kind* 'child', Latin *tempus* 'time', and Slovene *mésto* 'town'. As with other gender categories, exceptions in which grammatical and natural gender are in conflict can be found. For example, in German, the term *das Mädchen* 'girl' is grammatically neuter, even though the natural gender would be feminine. This is due to the diminutive suffix *-chen,* which

requires neuter gender (Corbett 1991: 66). Generally, however, neuter gender is used for referents without sexual characteristics.

While these kinds of sex-based gender systems are prevalent cross-linguistically, other divisions between sex-based gender categories are possible. For example, languages such as Danish, Swedish, and Dutch merge the historical masculine and feminine gender classes into one common gender, with neuter gender being applied to all other nouns (Corbett 1991: 247). Likewise, some languages that include sex-based gender categories may have other gender classes that are not centered on the biological sex of their referents. These types of gender systems may include classes for vegetables or plant products alongside sex-based categories of masculine and feminine. One such example is Dyirbal, a Pama-Nyungan language which has masculine and feminine gender classes, a third class for edible fruit and vegetables, and a fourth class for residue (i.e., items that do not fit into any of the other three classes; Corbett 1991: 15–19).

By contrast, in non-sex-based nominal classification systems, nouns are divided into classes based on other physical properties. For example, some languages make a distinction between animate and inanimate. In such languages, animate values are used to denote biologically animate beings, as well as certain inanimate objects which are nevertheless classified as being grammatically animate (Mithun 1999: 98; Kibort & Corbett 2008). Likewise, languages with this distinction use inanimate values to refer to non-living entities (Kibort & Corbett 2008). In Plains Cree, for example, nouns referring to living beings such as *atim* ‘dog’, *iskwêw* ‘woman’, and *môswa* ‘moose’ are animate, while those referring to inanimate objects and abstract concepts such as *maskisin* ‘shoe’, *mitêhimin* ‘strawberry’, and *kisêwâtisiwin* ‘kindness’ are grammatically inanimate.

However, there are some inanimate objects which are nevertheless grammatically animate in Cree, such as *asikan* 'sock', *ayôskan* 'raspberry', and *ospwâkan* 'pipe'. As a source language of Michif, the intricacies of the Plains Cree system of non-sex-based classification are particularly relevant to this study, and are further discussed in Chapter Five.

Other examples of non-sex-based nominal classification values include human vs. non-human, as found in some Dravidian languages spoken in India; and rational vs. non-rational, as found in Tamil and some other Dravidian languages. These divisions may be used on their own, or in combination with masculine/feminine and animate/inanimate values (Aikhenvald 2017: 363–4). For our purposes here, we focus on the categories relevant to Michif—the distinction between masculine and feminine and the distinction between animate and inanimate. In the following section, we discuss the different types of nominal classification systems found cross-linguistically.

3.3 Assignment systems

The question of how nouns are assigned to particular noun classes and how speakers know which value to assign to any given noun has long been a subject of linguistic investigation. Several hypotheses have been put forward in this regard. Corbett refers to models that attempt to capture the ways in which nominal classification values are assigned to nouns in a language as "assignment systems" (1991: 7), and claims that the types of assignment systems found throughout the world can be grouped into three categories—*strictly semantic*, *predominantly semantic,* and *formal* (Corbett 2013c). These are each discussed below.

In *strictly semantic* systems, values for virtually the entire noun inventory can be drawn based purely on semantics (Corbett 2013c). One example of a language with a strictly semantic assignment system is Kannada, a Dravidian language spoken in southern India. In this language, all nouns referring to male humans, deities, demons, or heavenly bodies have masculine values, while those denoting female humans, deities, demons, or heavenly bodies have feminine values. All other nouns are neuter (Corbett 2013c).

In *predominantly semantic* systems, values for most nouns are based on semantic features, though there are a number of exceptions (Corbett 2014: 112). Corbett identifies Bininj Gun-Wok, a member of the Gunwinyguan language family of northern Australia, as an example of such a system. In this language, noun class assignment is loosely based on a masculine/feminine/neuter/vegetable distinction, though there are many unpredictable exceptions (Corbett 2013c).

Finally, in a *formal* assignment system, some nouns are assigned values based on their semantic properties, and the residue (i.e., nouns that do not have a semantic designation) are then assigned noun class values based on formal rules. These rules may be based on either phonological or morphological properties, or a combination of the two. Corbett notes that semantics can always override formal factors and that no language ever assigns noun class values based on formal rules alone (2014: 114). He also claims that, if semantic and formal rules are ever in conflict, the semantic rules typically take precedence (Corbett 2013c).

An example of a language with a formal assignment system based on phonological rules is Qafar, an Eastern Cushitic language spoken in northeastern Ethiopia and Djibouti. In Qafar, sex-differentiable nouns are assigned either masculine or feminine

values, while all remaining nouns are assigned values based on a set of phonological rules. For instance, nouns ending in an accented vowel are feminine (e.g., *karm**à*** 'autumn'), while all others are masculine (e.g., *gil**à**l* 'winter' and *t**à**mu* 'taste') (Corbett 2013c).

An example of a language with a formal assignment conditioned by morphological properties is Russian. There are four primary morphological classes of nouns in Russian, each of which has its own characteristic inflectional patterns. These inflectional classes are highly correlated with noun class assignment. In most cases, Class I nouns have masculine values, Class II and III nouns have feminine values, and Class IV nouns are neuter. According to Corbett (2013c), membership in one of these inflectional classes generally determines a non-sex-differentiable noun's value. For example, a noun like *kost'* 'bone' that appears with Class III inflectional endings will most likely be feminine, and one like *khleb* 'bread' that takes Class I endings can be expected to be masculine. Although this is generally the case, sex-differentiable nouns may present exceptions, with some nouns belonging to an inflectional class that is more often associated with another noun class. For instance, a noun like *djadja* 'uncle' receives Class II inflectional endings, but is treated as being grammatically masculine due to its semantics. Thus, while morphological criteria are important for noun class assignment in Russian, they may be overridden by semantics in the case of sex-differentiable nouns.

In sum, we have seen that languages with noun classification systems assign noun class values based on a combination of semantic and formal features. Of particular interest to this study are the noun class assignment systems for Michif's source languages, Plains Cree and Michif. Corbett classifies Plains Cree as having a

predominantly semantic system of noun class assignment, while French has a formal assignment system (Corbett 1991; Corbett 2013c). In Chapter Seven, we examine the assignment of noun class values to Michif nouns, comparing the animacy and gender values of nouns found in spontaneous Michif discourse against those of equivalent nouns in Michif's source languages.

3.3.1 Assignment of nominal classification values in loanwords

A related question for languages with nominal classification systems is how noun class values are assigned to new words that are borrowed from other languages. This question is of particular relevance to Michif, which can be seen as having "borrowed" most of its noun inventory from French, and which is also experiencing increased influence from English.

According to Corbett, generally speaking, loanwords undergo the same process of noun class assignment as other nouns in a given language. Thus, if a language has a predominantly semantic assignment system, this same process will apply to any loanwords that are brought into that language (1991: 74). For example, in Northern Cheyenne (Algonquian), which reportedly has a semantic assignment system, loanwords can be assigned their animacy values solely on the meaning of the referent, with no regard to form (Straus & Brightman 1982: 100, cited in Corbett 1991: 71).

Scholars have identified other possibilities for loanword noun class assignment as well, such as *semantic analogy*, a process of concept association in which a loanword assumes the same noun class value of an existing noun in the language with a similar meaning (Poplack, Pousada & Sankoff 1982, cited in Corbett 1991: 75). Another hypothesis is that loanwords are automatically assigned to an unmarked class,

irrespective of meaning or form, although this idea has generally been met with skepticism (Poplack, Pousada & Sankoff 1982, cited in Corbett 1991: 77–8). Another possibility is that, if the borrowed word is replacing a native word, it simply retains the noun class value that it originally had in the donor language (Ibrahim 1973: 61–2). Corbett claims, however, that this is unlikely, except when a conscious effort is made by very educated speakers (1991: 81). Ibrahim also suggests that phonetic factors, such as homophony and rhyme, may be at play (1973: 61–2). For example, if a borrowed word is homophonous or rhymes with a native word, it might take that same noun class value. In general, however, Corbett claims that no factors outside of the normal noun class assignment rules of the borrowing language are necessary:

> [T]he assignment of loanwords depends on the same types of factor[s] as the assignment of native words…Only if the rules for native words do not cover all the cases is there any justification for postulating additional factors. The normal situation is one in which borrowings are assigned in essentially the same way as are native words. (1991: 81)

Having now introduced several criteria for defining and identifying nominal classification, as well types of noun class assignment systems found cross-linguistically and ways in which loanwords may be assigned to different noun classes in a language, we now turn to examples of languages which have been claimed to have multiple and/or combined systems of nominal classification.

3.4 Combined nominal classification systems

Also of interest to our discussion of the cross-linguistic properties and distribution of languages with nominal classification systems is the existence of languages which appear to have multiple coexisting systems of nominal classification, which Corbett refers to as

combined gender systems (1991: 184; 2014: 101). In the following sections, we discuss three examples of such languages—Burmeso (3.4.1), Mba (3.4.2), and Michif (3.4.3).

3.4.1 Burmeso

Burmeso, a West Papuan language spoken in Indonesia, is one language that is reported to have a combined nominal classification system. In Burmeso, we find two intersecting systems of noun classification which are not based on entirely the same set of formal or semantic properties. The first system has six noun agreement suffixes which trigger different agreement markers on the verb, which has two possible inflectional classes. This system is based on gender and animacy, but also has other, more restricted categories, including one category for mass nouns, two categories for specific lexical items, and a catch-all category for residue, as shown in Table 1:

Table 1: Noun class markers on Burmeso verbs (after Donohue 2001: 100, 102, cited in Corbett 2014: 102)

	Assignment	**Inflection class 1**		**Inflection class 2**	
		e.g., *-ihi-* 'see'		e.g., *-akwa-* 'bite'	
		SG	**PL**	**SG**	**PL**
I	male	j-	s-	b-	t-
II	female animate	g-	s-	n-	t-
III	miscellaneous	g-	j-	n-	b-
IV	mass nouns	j-	j-	b-	b-
V	banana, sago tree	j-	g-	b-	n-
VI	arrows, coconuts	g-	g-	n-	n-

This noun class marking on verbs can be seen in the following examples:

(8)

a. da nawak g–ihi–maru
1SG woman.SG II.SG–see–TODAY'S.PAST
'I saw a woman'

b. da mibo j–ihi–maru
1SG banana.SG V.SG–see–TODAY'S.PAST
'I saw a banana'

c. jamo nawak n–akwa–ru
dog.SG woman.SG II.SG–bite–TODAY'S.PAST
'the dog bit a woman'

(Donohue 2001: 99–101, cited in Corbett 2014: 102)

In (8)a, the singular noun *nawak* 'woman', which belongs to Class II, triggers the use of the singular Class II prefix *g-* on the verb, which belongs to inflection class 1. This is in contrast to (8)b, in which the singular noun *mibo* 'banana', which belongs to Class V, triggers a different prefix, *j-*, on the verb. Meanwhile, in (8)c, the Class II noun *nawak* 'woman' triggers the singular Class II prefix *n-* on the verb, rather than *g-*, as in (8)a, since the verb *-akwa-* 'bite' belongs to inflection Class 2.

Adjectives, on the other hand, have a completely different set of noun class agreement suffixes which show sensitivity to animacy (animate, inanimate) and gender (masculine, feminine, neuter), without the other distinctions noted in Table 1 (Donohue 2001, cited in Corbett 2014: 101–4). These adjective agreement suffixes are shown in Table 2:

Table 2: Noun class agreement suffixes on adjectives in Burmeso (Donohue 2001: 106, cited in Corbett 2014: 103)

Gender	**Singular**	**Plural**
Masculine	-ab	-od(o)
Feminine	-an	-od(o)
Neuter	-ora	-or(o)
Masculine inanimate	-ab	-or
Feminine inanimate	-an	-or
Neuter animate	-ora	-od

This noun class marking on adjectives is exemplified in (9):

(9)

a. da de koya bek–abo
1SG 1SG.POSS grandfather.SG good–M.SG
'my grandfather is well'

b. da d–asia bek–an
1SG 1SG.POSS–grandmother.SG good–F.SG
'my grandmother is well'

c. da de koysorad bek–odo
1SG 1SG.POSS grandson.PL good–ANIM.PL
'my grandsons are well'

(Donohue 2001: 99–101, cited in Corbett 2014: 102)

In (9)a, the noun *koya* 'grandfather' is both masculine and singular, triggering the use of the masculine singular suffix *-abo* on the adjective *bek* 'good'. In (9)b, the subject *asia* 'grandmother' is instead feminine, triggering a feminine agreement suffix, *-an*, on the adjective. In (9)c, the plural subject *koysorad* 'grandsons' triggers the animate plural agreement suffix *-odo* on the adjective. Thus, the examples presented in (8) and (9) show that Burmeso has a combined nominal classification system, with two different sets of

noun class agreement markers based on different semantic criteria, one appearing on verbs and one on adjectives.

3.4.2 Mba

Another possible instance of a language with a combined nominal classification system is Mba, a Ubangian language spoken in the Democratic Republic of Congo (Corbett 1991: 184–8; Corbett 2011: 465–6). Mba can be analyzed as having two separate systems of nominal classification: one Bantu-like system which triggers attributive agreement marking according to the noun used, and another system used optionally for animate pronouns (Corbett 2011: 465). Consider the following examples:

(10)

a. kíá (ɓi) k–ímá
snake ANIM 5–one
'one snake'

b. kásá k–ímá
leaf 5–one
'one leaf'

(Corbett 1991: 186)

Both of the nouns in (10) take class 5 agreement, as indicated by the class 5 prefix *k-* attached to the attributive *ímá* 'one'. However, since *kíá* 'snake' is animate in (10)a, it may optionally be used with the animate pronoun *ɓi*. This is not possible for *kásá* 'leaf' in (10)b, which is inanimate.

The Bantu-like system is comprised of six classes (1/2, 3/4, 5/6, 7/2, 9/6, 11/2) which are assigned based on morphological criteria, while the system involving pronouns

is based on semantic criteria related to animacy. The personal pronouns are provided in Table 3:

Table 3: Mba personal pronouns (Corbett 1991: 185)

	Singular	**Plural**
Male human	ndé	ɓé
Other animate	ɓi	

Based on these two sets of agreement markers, it is a question then as to how many noun agreement classes Mba has. Given that there are six classes for attributive agreement and three for pronominal agreement, there would be 18 possible combinations of these classes. However, only 11 of these combinations are attested. These are presented in Table 4:

Table 4: Consistent agreement patterns in Mba (adapted from Corbett 2011: 466)

Noun class	**Attributive agreement**		**Pronoun / optional agreement**	**Combined noun class**
	Singular agreement	**Plural agreement**		
1	w	y	ndé	1/2 male personal
2	w	y	ɓi	1/2 animate
3	w	y	Ø	1/2 inanimate
4	l	s	Ø	3/4 inanimate
5	k	z	ɓi	5/6 animate
6	k	z	Ø	5/6 inanimate
7	g	y	ndé	7/2 male personal
8	g	y	ɓi	7/2 animate
9	g	y	Ø	7/2 inanimate
10	ny	z	Ø	9/6 inanimate
11	m	y	Ø	11/2 inanimate

One way of analyzing noun classification in Mba is that each noun has two separate noun class values which stem from each system, as described thus far. However, Corbett notes that alternative analyses are also possible. Instead of analyzing each noun as having two separate noun class values, each from a separate system, it is also possible to treat each attested intersection of noun class values from both classes as forming classes of their own. For instance, rather than posit that a given noun is classified twice, once as a male human and once as belonging to a particular Bantu noun class, Corbett notes that it is possible to treat this noun as belonging to a single, complex noun class, as shown in the rightmost column. Corbett does not advocate for one of these analyses over the other, but observes that each approach entails a distinct set of consequences for further linguistic analysis. A combined system analysis requires two separate sets of agreement rules, thus

complicating the syntax. On the other hand, an analysis which treats Mba as having a single nominal classification system requires a more complex noun class specification. Neither approach thus necessarily reduces the inherent complexity of the semantic criteria that appear to underlie nominal agreement patterns in this language.

3.4.3 Michif

Finally, Corbett identifies Michif as another example of a language with a combined nominal classification system, stating that "given the right circumstances, truly weird systems can arise" (2006: 269). Corbett asserts that Michif has inherited nominal classification systems from both of its major source languages—Canadian French (Indo-European; Romance) and Plains Cree (Algonquian). As described in Chapter Two, French makes a distinction between masculine and feminine gender, while Cree distinguishes between animate and inanimate nouns. Each noun in Michif is associated with two values, one from each of these categories (Bakker 1997; Corbett 2006: 270). Both of these categories trigger different agreement with different associated elements. For example, definite articles and adjectives (when preceding the noun) have been reported to agree in gender, while verbs and demonstratives agree in animate/inanimate animacy. This is illustrated in examples (11)–(14), reproduced from Section 3.1 above:

(11)	la	pchit	fii	awa
	the:FEM:SG	$\text{little}_{\text{FEM}}$	girl	DEM:AN:SG
	'this little girl'			

(Mervin Fleury; 2013–09–13)

(12) li pchi gaa awa
the:MASC:SG little$_{MASC}$ guy DEM:AN:SG
'this little guy'

(Mervin Fleury; 2013–09–13)

(13) ta meezoñ ooma
2SG.POSS:FEM:SG house DEM:IN:SG
'your house here'

(Verna DeMontigny; 2013–09–13)

(14) li kaab ooma
the:MASC:SG rope DEM:IN:SG
'this rope'

(Mary Fleury; 2013–09–13)

In (11), we see that *fii* 'girl' is both feminine and animate. The French-origin definite article *la* 'the' indicates feminine gender, while the Cree-origin demonstrate *awa* 'this one' indicates animacy. This differs from *gaa* 'guy' in (12) only in the use of masculine definite article *li* 'the'.[6] In (13), we see that *meezoñ* 'house' is both feminine and animate through the agreement that it triggers with surrounding elements. The French-origin possessive pronoun *ta* 'your' indicates feminine gender, while the Cree-origin demonstrative *ooma* 'this one' indicates that the noun is inanimate. Likewise, in (14), *kaab* 'rope' is also inanimate since the same demonstrative *ooma* is used, but it is masculine because it occurs with the masculine definite article *li* 'the'. The details of this system are discussed further in Chapter Four.

Given these facts, it would be possible to analyze Michif as either having two separate systems of nominal classification, one based on masculine/feminine values and another based on animate/inanimate values, or as having one nominal classification

[6] As noted in Chapter Two, adjectives in French can be inflected for gender (e.g., *petit* 'little (m.)', *petite* 'little (f.)'. However, a great deal of variability is found in these and similar forms in Michif. Adjectives such as these are therefore not are not treated as being reliable indicators of gender in Michif in this study.

system with four possible values (masculine-animate, masculine-inanimate, feminine-animate, feminine-inanimate). Either of these analyses places Michif in a typologically uncommon space, since, as we saw above, the majority of the world's languages either have no noun classification at all, or have noun class systems with less than three values. Combined nominal classification systems are even less common cross-linguistically. In Chapter Eight, I argue that Michif is better analyzed as having a combined nominal classification system, based on the results of this study presented in Chapter Seven.

Chapter 4: Nominal classification in Michif and its source languages

Several features of Michif have been claimed to distinguish it from other mixed languages. In addition to its apparent grammatical complexity, Michif is commonly reported to exhibit an unusual split in its source languages, having Algonquian-derived verb phrases and French-derived noun phrases. This differs from the profiles of other mixed languages, where typically the substantial parts of the grammar are inherited from one source language, while the lexicon is inherited from another source language (Meakins 2013). This division of the lexicon between Algonquian and French can be seen in the following examples, in which French-origin elements are bolded and Cree-origin elements are non-bolded:[7]

(15) **lii kuuvart** ee-kii-akotaachik akota

lii	**kuuvart**	ee–kii–akot–aa–t–ik	akota
the:PL	blanket	CONJ–PST–hang$_{VTA}$–3>3'.CONJ–3SG.CONJ–3PL:AN	there
ART	N	V	PRT

'they hung blankets there'

(Victoria Genaille, 2012–10–15)[8]

[7] There are several Michif orthographies in use, though none of them has emerged as the standard, and different orthographies are used in the sources cited here. To facilitate comparison, the orthography adopted here strives to follow that used in the Michif Dictionary Project, as found in Rosen (2016b). In some cases, the spellings differ slightly, reflecting potential dialectal variation. Vowels in the orthography used here are represented as follows: a = /a/; aa = /a:/; ae = /æ/; e = /ɛ/; ee = /e/; i = /ɪ/; ii = /i/; o = /ɔ/; oo = /o:/; u = /ʊ/; uu = /u/. The symbol "ñ" indicates that the preceding vowel is nasalized. Consonants follow conventions similar to English, with the following exceptions: sh = /ʃ/; zh = /ʒ/; ch = /tʃ/. The spellings for any cited examples (e.g., from Bakker 1997; Bakker & Papen 1997; Rosen & Souter 2009a) have been adjusted to this orthography and I have added my own interlinear glosses. Unless otherwise noted, the data presented here are taken from my corpus or field notes. I follow Chelliah and de Reuse (2011: 213) in noting how each data point has been acquired, whether through elicitation or corpus searches. Examples which have been elicited are noted as such. Otherwise, all data are from spontaneous, unprompted speech. Note also that false starts have been removed from all examples and glosses.

(16) eekotee gii-wiikinaan **zhuus-a kat añ** ee–ayaayaan

eekotee	ni–kii–wiiki–naan	**zhuus–a**	**kat**	**añ**
there	1–PST–live$_{VAI}$–1PL	until	four	year
PRT	V	PREP	NUM	N

ee–ayaa–yaan
CONJ–have$_{VAIT}$–1SG.CONJ
V

'we lived there until I was four years old'
(Victoria Genaille, 2012–10–15)

(17) **not meezoñ** kaa-waapahtamaahk yaeñk **dañ li feu**

not	**meezoñ**	kaa–waapaht–am–aahkw
1PL.POSS:SG	house	REL–see$_{VTI}$–3OBJ:IN–1PL.CONJ
POSS	N	V

yaeñk	**dañ**	**li**	**feu**
only	PREP	the:MASC:SG	fire
ADV	PREP	ART	N

'all we saw was our house on fire'
(Victoria Genaille, 2012–10–15)

In (15), we see that the noun *kuuvart* 'blanket' and its article are of French origin, while the verb *ee-kii-akotaachik* 'they hung s.t.' and the particle *akota* 'there' are of Cree origin. Likewise, in (16), the preposition *zhuus-a* 'until', the number *kat* 'four', and the noun *añ* 'year' are of French origin, while the particle *eekotee* 'there' and both verbs are derived from Cree. In (17), all elements of the sentence besides the verb are of French origin. As we will see below, however, there is leakage between these two systems, and elements of both major source languages are distributed throughout the Michif grammatical system.

The purpose of this chapter is to introduce the reader to basic features of Michif that are needed to interpret the examples provided in this dissertation, focusing in

particular on aspects of Michif structure that are relevant to an analysis of nominal classification. We also examine the properties of nominal classification for both of Michif's primary source languages, considering the kinds of constructions in which both animacy and gender are reflected, and identifying points of similarity and difference between these source languages and Michif. We begin with an overview of Michif's sound system in Section 4.1. This is followed by a description of nouns and nominal constructions in Section 4.2, and by a description of verbs and verbal constructions in Section 4.3.

4.1 Sound system

The analysis of Michif's phonological system is controversial and, as discussed in Chapter One, has received considerable attention in the literature. Under some analyses, Michif is considered to be comprised of two parallel phonological systems, one Cree and one French (e.g., Andrella 1983; Bakker & Papen 1997; Papen 2005a; Papen 2017). For these authors, the French-origin inventory consists of twenty-three consonants, while the Cree-origin inventory consists of ten consonants, as shown in Table 5–Table 6 (adapted from Papen 2005a: 80):

Table 5: French-derived consonants

	Bilabial	**Labiodental**	**Alveolar**	**Alveo-palatal**	**Velar**	**Glottal**
Stops	p b		t d		k g	
Fricative		f v	s z	ʃ ʒ		h
Affricate				tʃ dʒ		
Nasals	m		n	ɲ	ŋ	
Liquids			r l			
Glides	w			j		

Table 6: Cree-derived consonants

	Bilabial	**Alveolar**	**Alveo-palatal**	**Velar**	**Glottal**
Stops	p	t		k	
Fricatives			ʃ		h
Affricates			tʃ		
Nasals	m	n			
Glides	w		j		

While a large portion of these consonant inventories overlaps, there are a number of French phonemes which are not shared by Cree: /b/, /f/, /v/, /d/, /s/, /z/, /r/, /l/, /ʒ/, /dʒ/, /ɲ/, /g/, /ŋ/.

Meanwhile, the vowel inventory on the French side has eleven oral vowels and five nasalized vowels, while the Cree vowel inventory includes seven oral vowels (four of which are long) and two nasalized vowels. These are represented in Figure 1–Figure 2 (adapted from Papen 2005a: 80):

Figure 1: French-derived vowels

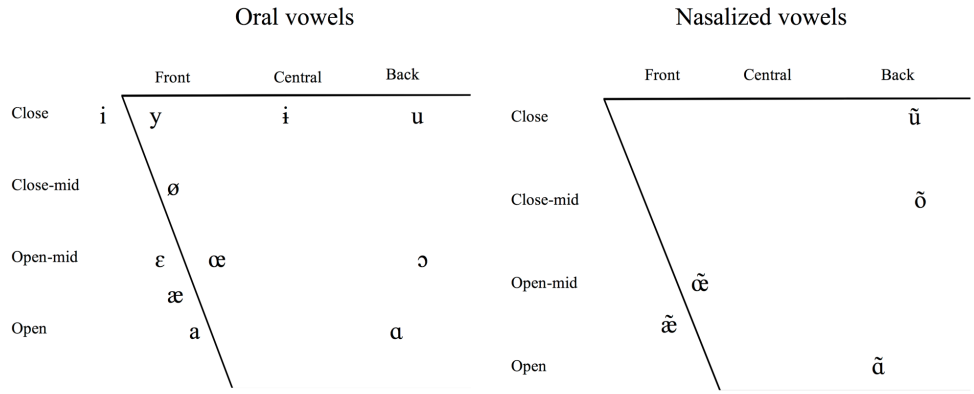
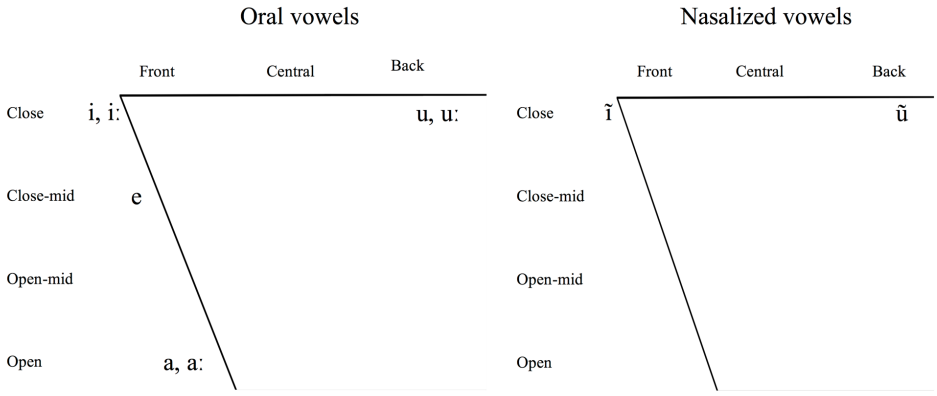


Figure 2: Cree-derived vowels[9]



Again, while there is some overlap between these two inventories, there are also several phonemes which are present in one source language but not the other. For example, the French vowels /y/, /ɨ/, /ø/, /ɛ/, /œ/, /ɔ/, /æ/, /ɑ/, /ũ/, /õ/, /æ̃/, /œ̃/, and /ã/ are not represented in Cree, whereas the Cree vowels /i:/, /u:/, /e/, /a:/, and /ĩ/ are not represented in French.

Other researchers (e.g., Prichard & Shwayder 2014; Rosen 2007) have argued that Michif is comprised of a single phonological system, and that categorization of the

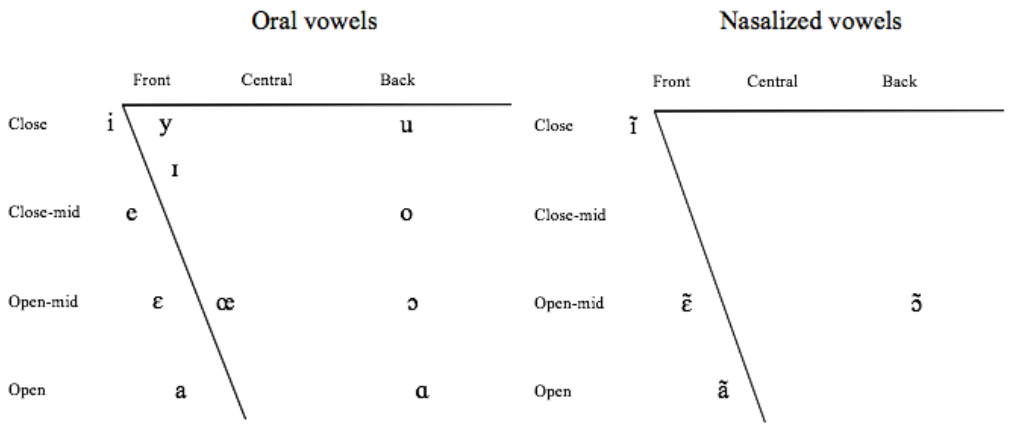
[9] Note that some Cree-origin forms appear in Michif with nasal vowels, although Plains Cree, unlike Ojibwe, is not generally considered to have a separate series of nasal vowels.

phonemic inventory according to source language is neither accurate nor necessary. The consonant and vowel inventories under such a single-system analysis are shown below:

Table 7: Michif consonants (adapted from Rosen 2007: 142)

	Bilabial	**Labiodental**	**Alveolar**	**Alveo-palatal**	**Velar**	**Glottal**
Stops	hp p b		ht t d		hk k g	
Fricatives		f v	s z	ʃ ʒ		h
Affricates				htʃ tʃ dʒ		
Nasals	m		n			
Liquids			r l			
Glides	w			j		

Table 8: Michif vowels (adapted from Rosen 2007: 166)



According to this analysis, Michif consists of a total of twenty-five consonants, including voiced and voiceless stops, affricates, fricatives, nasals, liquids, and glides. Note that this differs somewhat from a simple combination of the French and Cree consonant

inventories presented in Table 5–Table 6. For one, Rosen (2007) considers a class of preaspirated obstruents (/hp/, /ht/, /hk/, (/htʃ/) as single segments rather than consonant clusters. In addition, /ɲ/ and /ŋ/ are not found to be phonemes in the Michif consonant inventory. The Michif vowel inventory under this analysis consists of a total of fifteen vowels, eleven oral and four nasalized. Note that not all French-origin vowels seen in Figure 1 are represented here, including /ɨ/, /ø/, /æ/, /æ̃/, /ũ/, and /õ/. Moreoever, this analysis does not consider Michif to have long vowels, as in Cree. Instead, this surfaces in the merged phonology as a quality distinction (e.g., /i/ vs. /ɪ/ rather than /i/ vs. /i:/.

4.2 Nouns and nominal constructions

Most nouns in Michif are of French origin, though a handful of Cree nouns (e.g., *koohkom* 'grandmother', *tahkwahiminaana* 'chokecherries', *shikaak* 'skunk') are also in use. More recently, some English-origin nouns have become integrated into the lexicon as well (e.g., *aeñ pascher* 'pasture', *li giir* 'gear', etc.). Nouns are lexically classified as both masculine/feminine and animate/inanimate. As covert categories, these nominal classification values are not marked on nouns themselves, but are instead reflected through agreement in several constructions. In this section, we discuss various aspects of nouns and nominal constructions in Michif and its source languages, including nominal categorization (4.2.1), inflectional categories (4.2.2), constructions showing gender agreement (4.2.3), and constructions showing animacy agreement (4.2.4).

4.2.1 Nominal categorization

Michif has inherited two covert systems of nominal classification, each from a different source language—masculine/feminine gender from French, and animacy from Cree. In the variety of Michif studied here, all nouns are lexically classified for both of these categories, whether they are of French, Cree, or English origin.[10] In Michif, animacy and gender are both manifested through agreement with other forms with which the noun occurs. Evidence for the animate/inanimate distinction is shown through Cree-origin demonstrative and interrogative pronouns accompanying nouns (e.g., *awa*$_{AN}$ 'this', *ooma* 'this$_{IN}$'; *ookik*$_{AN}$ 'these', *ohiñ* 'these$_{IN}$'; *taana*$_{AN}$ 'which one', *taanima*$_{IN}$ 'which one'; *taaniki*$_{AN}$ 'which ones', *taanihi*$_{IN}$ 'which ones'), as well as through verbal agreement. Evidence for the masculine/feminine distinction is found in French-origin singular articles (e.g., *li* 'the$_{MASC}$' *la*$_{FEM}$ 'the'; *aeñ*$_{MASC}$ 'a/n', *en*$_{FEM}$ 'a/n'), possessive pronouns (e.g., *moñ* 'my$_{MASC}$', *ma* 'my$_{FEM}$'; *toñ* 'your$_{MASC}$', *ta* 'your$_{FEM}$'; *soñ* 'his/her$_{MASC}$', *sa* 'his/her$_{FEM}$'), and in some cases prenominal adjectives accompanying nouns (e.g., *gro*$_{MASC}$ 'big', *gros* 'big$_{FEM}$') (Bakker 1997: 102). Consider the following example:[11]

[10]In his study of a dialect of French/Cree spoken in Buffalo Narrows, Saskatchewan, Hogmen clearly demonstrates agreement for both animacy and gender (1981). However, he claims that masculine/feminine gender only affects French-origin nouns, and does not apply to the handful of Cree-origin nouns in the language (Hogmen 1981: 87). While this may be the case in the Buffalo Narrows variety, this is not the case in the variety of Michif examined here.

[11] This sentence arose during a toy game session. *Zwee* 'goose' is referring to a toy goose and *triangle* 'triangle' is a nonce borrowing from English.

(18) OK, moñ zwee awa, wiishta oota dañ li triangle ni-wii-ahaaw

OK	moñ	zwee	awa	wiishta
OK	1SG.POSS:MASC:SG	goose	DEM:AN:SG	him.too

oota	dañ	li	triangle
here	in	the:MASC:SG	triangle

ni–wii–ah–aa–w
1–VOL–put$_{VTA}$–NON3>3–3SG:AN

'OK, my goose, this one, I will put him in the triangle, too'

Mary Fleury; 2013–09–13

In (18), the noun *zwee* 'goose' agrees with neighboring elements, both for gender (masculine) and for animacy (animate). It appears with the possessive pronoun *moñ* 'my', indicating that it is both masculine and singular, as well as with the demonstrative *awa* 'this', indicating that it is animate and singular. Additionally, the verbal agreement suffixes indicate that the verb is transitive and that there is a first-person subject, while the suffix *-aa* also indicates that action is being done by a non-third person to a third person argument.

4.2.2 Nominal inflectional categories

Michif nouns inflect for two categories—number (singular, plural) and obviation (proximate, obviative). For French-origin nouns, singular and plural values are not marked on the nouns themselves, but are instead marked by agreement on the articles, possessives, and/or demonstratives that accompany the noun:

(19) pii li pchi shyaeñ wiishta awa nipaaw

pii	li	pchi	shyaeñ	wiishta	awa	nipaa–w
and	the:MASC:SG	little	dog	him.too	DEM:AN:SG	$sleep_{VAI}$–3SG:AN

'and this little dog is sleeping too'

(Mervin Fleury; 2013–07–16)

(20) deu lii shyaeñ oota niiya dayaawaawak

deu	lii	shyaeñ	oota	niiya	ni–t–ayaaw–aa–wak
two	the:PL	dog	here	1SG:PRN	1–t–$have_{VTA}$–NON3>3–3PL:AN

'I have two dogs here'

(Mary Fleury; 2013–09–13)

In (19), the noun *shyaeñ* 'dog' is singular, as indicated by the singular definite article *li* 'the', which also indicates masculine gender. This is in contrast to (20), where the plural definite article *lii* 'the' indicates that *shyaeñ* 'dog' is plural.

By contrast, while Cree-origin nouns are also marked for plurality by the article, possessive, or demonstrative, they may be additionally marked for number by the Cree-origin plural suffix *-a*:

(21) Yeah, pii…eñ laton moshonaan, eh. Lii pwer, lii takwaminaana, 'chiko keekwee.

yeah	pii	eñ	laton	moshow–naan	eh	lii	pwer
yeah	and	in	fall	$pick_{VAI}$–1PL	eh	the:PL	saskatoon.berry

lii	takwaminaan–a	nawachiko	keekwee
the:PL	chokecherry–PL	kind.of	all

'Yeah, in the fall…we picked, eh. Saskatoons, chokecherries, pretty much everything.'

(Mervin Fleury; 2012–10–13)

In (21), the Cree-origin noun *takwaminaana* 'chokecherries' is marked for plurality both by the French-origin definite plural article *lii* and also by the Cree-origin plural suffix *-a*. Michif articles, possessives and demonstratives are discussed in further detail below.

Michif nouns may also be marked for obviation, also referred to in the Algonquian literature as the fourth person or the minor third person. Obviation is used as a mechanism of differentiation in discourse spans involving more than one animate third person, as well as in possessive phrases involving third persons. There is no consensus as to what the pragmatic functions of obviation are. It has been said to involve speaker's point of view (Wolfart 1973: 17, citing Bloomfield 1962:38) and intentionality (Mühlbauer 2007; Mühlbauer 2008), among other things. A recent hypothesis proposes that in Michif, unlike in Cree, the obviative may be shifting from an information structure marker to a syntactic role marker, specifically that of a differential object marker (Antonov 2015). While an in-depth study of the discourse-level functions of obviation is beyond the scope of this dissertation, more information on this feature can be found in the Algonquianist literature for a number of languages (cf. Dahlstrom 1986; Dahlstrom 1991; Goddard 1984; Goddard 1990; Hasler & Wolfart 2002; Mühlbauer 2007; Mühlbauer 2008; Russell 1991; Thomason 2003; Weaver 1982; Weaver 1983; Wolfart 1978, among others).

In Michif, the more backgrounded or peripheral of two nouns (i.e., the obviative participant) receives obviative marking in the form of the suffix *-(w)a*, while the proximate participant remains unmarked. Obviative nouns are not marked for singular or plural number, and do not trigger plural agreement on the verb. The obviation status of the arguments of the verb is indicated by direction markers, where the obviative participant is glossed as 3':

(22) la fii kii-waapameew ohkoma

la	fii	kii–waapam–ee–w	o–hkom–a
the:FEM:SG	girl	PST–see$_{VTA}$–3>3'.IND–3SG:AN	3–grandmother–OBV

'the girl saw her grandmother'

(Victoria Genaille; 2015–07–17; elicited)

(23) ohkoma kii-waapamiko

o–hkom–a	kii–waapam–ikw–w
3–grandmother–OBV	PST–see$_{VTA}$–3'>3–3SG:AN

'her grandmother saw her'

(Victoria Genaille; 2015–07–17; elicited)

In (22), 'girl' is a proximate participant and is unmarked, while 'grandmother' is marked as obviative, and the participants 'girl' and 'grandmother' each trigger different agreement markers on the verb: the direct marker *-ee* on the verb indicates that a third person proximate (3) is acting on a third person obviative argument (3'), showing that the proximate argument 'girl' is the subject and the obviative argument 'grandmother' is the object. This is in contrast to (23), where the verb receives the inverse marker *-ikw*, indicating that an obviative participant (3') is acting on a proximate participant (3) and that the obviative argument 'grandmother' is the subject while the proximate argument 'girl' is the object.[12]

While Cree-origin nouns are frequently marked for obviation in Michif, especially in possessed form (Bakker 1997: 88; Bakker & Papen 1997: 343), this is less common for French-origin nouns. However, though it is rare, I have found examples in my data where it is possible for the obviative suffix to occasionally co-occur with French-origin nouns:

[12] In this example, *la fii* 'girl' was established within the discourse span, so the overt noun was not provided.

(24) eñ laatomobil kii-itohteew kiinaatam anima li shoebox anima pee-waapahtahaat anihi la faam-a

eñ	laatomobil	kii–itohtee–w	kii–naat–am
the:FEM:SG	car	PST–go$_{VAI}$–3SG:AN	PST–fetch$_{VTI}$–3OBJ:IN

anima	li	shoebox	anima
DEM:IN:SG	the:MASC:SG	shoebox	DEM:IN:SG

pee–waapaht–a–h–aa–t	anihi
come–see$_{VTI}$–a–CAUS–3>3'.CONJ–3SG.CONJ	DEM:AN:3'

la	faam–a
the:FEM:SG	woman–OBV

'he went to the car to go and get that shoebox to come and show that woman'

(Victoria Genaille; 2012–12–02)

As shown in the above example, the obviative marker *-a* is suffixed to the French-origin noun *faam* 'woman'. French-origin nouns can thus nevertheless be shown to exhibit obviation through both nominal and verbal agreement. This is also the case for the French-origin noun *fii* 'girl' in example (74) in Section 4.3.3 below. The obviation status of a nominal argument is reflected in verbal agreement and direction morphology, even in cases where it is not overtly marked on the noun itself. This is further discussed in Section 4.3.3 below.

Obviative marking is not consistently used in all contexts where it would be expected for Cree, and thus appears to be on the decline in Michif (Weaver 1983). It often appears to be optional on both nouns and verbs. Without diachronic evidence, it is not possible to know whether use of this inflection is on the decline (as has been reported for some Algonquian languages; cf. Frantz 2009: viii on Blackfoot) or was never a stable feature in Michif the first place. Since Michif differs from Algonquian languages in

having predominantly French-origin nouns, one hypothesis about this inconsistency of obviative marking has been that the replacement of Cree nouns with French nouns has led to a lack of obviative inflection transferring over to Cree-origin verbs in Michif (Weaver 1982). This position is supported in part by the observation that in Michif—unlike Cree, which maintains obviation much more consistently—much of the obviative marking has been lost in the noun phrase. Nevertheless, even if a noun does not take overt obviative marking, it is still possible for it to be identified as obviative through inflection on the verb that agrees with it:

(25) kanawaapameew anihi lii pear dañ li payiiñ ee-apiyit

kanawaapam–ee–w	anihi	lii	pear	dañ
look.at$_{VTA}$–3>3'.IND–3SG:AN	DEM:AN:3'	the:PL	pear	PREP

li	payiiñ	ee–api–**yi**–t
the:MASC:SG	basket	CONJ–be.there$_{VAI}$–**OBV.SUBJ**–3SG.CONJ

'he's looking at those pears that are in the basket'

(Norman Fleury; 2013–08–24–02)

(26) waapahtam anima la rosh anda kaa-apiyit

waapaht–am	anima	la	rosh	anda
see$_{VTI}$–3OBJ:IN	DEM:IN:SG	the:FEM:SG	rock	there

kaa–api–**yi**–t
REL–sit$_{VAI}$–**OBV.SUBJ**–3SG.CONJ

'he didn't see the rock that was there'

(Norman Fleury; 2013–08–24–01)

In the examples above, neither of the French-origin subjects (*lii pear* 'pears' or *la rosh* 'rock') are morphologically marked as obviative, even though the verb reflects their obviative status through the use of the suffix *-yi* in both cases.

In outlining additional features of obviation in Michif, Bakker states that though most marking is optional, proper names always receive obviative marking. After this, humans are the most likely to receive obviative marking, followed by animals. Inanimates are never marked as obviative (1997: 89). Antonov (2015) argues that, when obviation is used, it tends to follow Silverstein's animacy hierarchy (PROPER NAME > HUMAN > ANIMATE), and that the higher on this hierarchy an argument is, the more likely it is to be marked as obviative.

4.2.3 Gender agreement

As in French, agreement for gender is found in several contexts in Michif. Evidence for masculine/feminine gender is primarily found in articles and possessives. Definite and indefinite articles and possessive pronouns agree with their head nouns for gender. Unlike French, Michif adjectives do not consistently agree for gender with their references, and demonstratives are of Cree rather than French origin. This is illustrated in Table 9:

Table 9: Gender agreement in French and Michif (after Hogmen 1981: 84)

	French		**Michif**	
Construction	**Masculine**	**Feminine**	**Masculine**	**Feminine**
Definite article	*le livre* 'the book'	*la table* 'the table'	*li liivr* 'the book'	*la tab* 'the table'
Indefinite article	*un livre* 'a book'	*une table* 'a table'	*aeñ liivr* 'a book'	*en tab* 'a table'
Demonstrative	*ce livre* 'this book'	*cette table* 'this table'	*ooma li liivr* 'this book'	*ooma la tab* 'this table'
Possessive pronoun	*mon livre* 'my book'	*ma table* 'my table'	*moñ liivr* 'my book'	*ma tab* 'my table'
Definite article, adjective	*le livre vert* 'the green book'	*la table verte* 'the green table'	*li liivr ver* 'the green book'	*la tab ver* 'the green table'

Michif nouns are most often accompanied by either a definite or indefinite article. Articles are derived from French and mark both gender and number, indicating whether a noun is masculine or feminine, singular or plural. Singular articles also indicate definiteness (e.g., *li kok* 'the rooster$_{MASC}$' vs. *aeñ kok* 'a rooster$_{MASC}$'; *la meezoñ* 'the house$_{FEM}$' vs. *en meezoñ* 'a house$_{FEM}$'). Both Michif and French make no gender distinction in the plural, either with definite articles (e.g., *lii liivr* 'the books,' *lii tab* 'the tables), or with possessive pronouns (e.g., *mii liivr* 'my books,' *mii tab* 'my tables'), and no plural forms exist for the indefinite article (Hogmen 1981: 85). Michif articles, along with their French equivalents, are provided in Table 10 below:

Table 10: Michif articles

	Definite	**French Equivalent**	**Indefinite**	**French Equivalent**
Masculine singular	li /lɪ/	le	aeñ /æ̃/	un
Feminine singular	la /la/	la	en /ɛn/	une
Plural	lii /li/	les	-----	des
English gloss	'the'		'a, an'	

Possessive pronouns in Michif also agree in gender with their accompanying nouns. Three distinct possessive systems are attested in Michif—one based on the French possessive paradigm, one based on the Cree possessive paradigm, and one which is mixed (Sammons 2013d). Of these, both the French-based and mixed systems indicate gender, while the Cree-based system does not. The French-origin paradigm can be used for any Michif noun, regardless of source origin. In this system, the French-origin possessive pronouns indicate the gender and number of the possessum and number and person of the possessor:

(27) ma belmer
ma belmer
1SG.POSS:FEM:SG mother-in-law
'my mother-in-law'

(28) moñ frer
moñ frer
1SG.POSS:MASC:SG brother

(29) mii zañfañ
mii zañfañ
1SG.POSS:PL child
'my children'

In (27), the possessive pronoun *ma* 'my' indicates that the referent is both feminine and singular, with a first person singular possessor. Likewise, in (28), the possessive pronoun *moñ* 'my' indicates a masculine singular possessum, with a first person singular possessor. In (29), the possessive pronoun *mii* 'my' indicates a plural possessum and a first person singular possessor. Gender of the possessum is not distinguished in plural forms. The complete set of French-origin possessives are presented in Table 11:

Table 11: Michif possessives: French-origin system (after Gillon & Rosen 2018: 89; Rosen & Souter 2009a: 60)

	Singular		**Plural**	
Animate subject	**Masculine**	**Feminine**	**Masculine**	**Feminine**
1SG	moñ	ma	mii	
2SG	toñ	ta	tii	
3SG	soñ	sa	sii	
1PL.EXCL	not		noo	
1PL.INCL	not		noo	
2PL	vot		voo	
3PL	lœr		lœr	

The Cree-origin possessive paradigm may also be used for the handful of nouns in Michif that are of Cree origin. In Cree, most body part and kinship terms are inalienably possessed. According to Bakker & Papen, there are only a handful of Cree-origin nouns in Michif, most of which are used to refer to berries, plants, kinship terms, and household objects (1997: 324). In my data, Cree-origin nouns are also occasionally used for body part terms, as shown in (30):

(30) sii zoree kinwayiw, oshoy chakwayiw

sii	zoree	kinwa–yiw	o–shoy	chakwa–yiw
3SG.POSS:PL	ear	be.long$_{VII}$–3'SG:IN	3–tail	be.short$_{VII}$–3'SG:IN

'his ears (OBV) are long, his tail (OBV) is short'

(George Fleury; Louis Riel Institute 2013)

In the above example, we see the Cree-origin *oshoy* 'his tail' noun being used to refer to a body part. This noun is inflected with the prefix *o-* to denote a third person possessor. Note that another noun referring to a body part also occurs in this example, but is of French origin. In this case, *zoree* 'ear' is used with the French-origin possessive pronoun *sii* rather than a Cree-origin prefix. The complete Cree-origin possessive system is provided in Table 12 below:

Table 12: Michif possessives: Cree-origin system

Animate Subject	**Form**	**Example**
1SG	ni–	*nimoshoom* 'my grandfather'
2SG	ki–	*nimoshoom* 'your (sg.) grandfather'
3SG	o–a	*omoshooma* 'his/her grandfather'
1PL.EXCL	ni–inaan	*nimoshoominaan* 'our (excl.) grandfather'
1PL.INCL	ki–inaan	*kimoshoominaan* 'our (incl.) grandfather'
2PL	ki–iwaaw	*kimoshoomiwaaw* 'your (pl.) grandfather'
3PL	o–iwaawa	*omoshoomiwaawa* 'their grandfather'

Finally, a third system reveals inflectional mixing between the Cree- and French-derived components of Michif (Sammons 2013d). For example, in (31) and (32), first person plural exclusive possession is not marked with either *not*, as in Table 11, or *ni-*

inaan, as in Table 12, but instead represents a distinct combination of elements from both of these systems (with the French-origin *moñ* or *ma* serving essentially the same function as Cree-origin *ni-*):

(31) moñ liiv-inaan
moñ liiv-naan
1SG.POSS:MASC:SG book–1PL
'our (excl.) book'
(Grace Zoldy, 2012–07–13; elicited)

(32) ma meezoñ-inaan
ma meezoñ-naan
1SG.POSS:FEM:SG house–1PL
'our (excl.) house'
(Grace Zoldy, 2012–07–13; elicited)

(33) taandee tii soerinaanik?
taandee tii soer–naan–ik
where 2SG.POSS:PL sister–1PL–3PL:AN
'where are our (incl.) sisters?'
(Grace Zoldy, 2012–07–18)

In (31), the French-origin possessive pronoun *moñ* combined with the suffix *-inaan* indicates a first person plural exclusive possessor. The possessive pronoun *moñ* also indicates that the possessum is both masculine and singular. Similarly, in (32) the possessive pronoun *ma* combined with the suffix *-inaan* indicates a first person plural exclusive possessor. However, in this case the possessive pronoun *ma* additionally indicates that the possessum is both feminine and singular. In (33), the second person pronoun *tii* combined with the suffix *-naan* indicates a first person plural inclusive possessor. The pronoun *tii* combined with the plural suffix *-ik* additionally indicates that the possessum is plural. The full set of possessive markers within this mixed system is provided in Table 13:

Table 13: Michif possessives: mixed system (after Rosen & Souter 2009a: 60)

	Singular		**Plural**	
Animate subject	**Masculine**	**Feminine**	**Masculine**	**Feminine**
1SG	moñ	ma	mii	
2SG	toñ	ta	tii	
3SG	soñ	sa	sii	
1PL.EXCL	moñ–inaan	ma–inaan	mii–inaanik	
1PL.INCL	not–inaan	not–inaan	not–inaanik	
2PL	toñ–inaan toñ–inaawaaw	ta–inaan ta–inaawaaw	tii–inaanik tii–inaawaawik	
3PL	soñ–iwaaw	sa–iwaaw	sii–iwaawa	

It has been reported that the possessives listed in Table 11 and Table 13 can only be used with animate possessors. When a possessor is inanimate, the preposition *di* is used with the noun instead (Gillon & Rosen 2018: 89):

(34) a. la zhaañb di tab
la zhaañb di tab
the:FEM:SG leg PREP table
'the table's legs'

b. *la tab sa zhaañb
la tab sa zhaañb
the:FEM:SG table 3SG.POSS:FEM:SG leg
'the table's leg [sic]'

(Gillon & Rosen 2018: 90–91)[13]

Gillon and Rosen view this as offering further evidence of the important role that animacy plays in Michif (2018: 91). This distinction between constructions using the possessive paradigm and those using *di* with the noun is an area of research which merits

[13] Orthography and interlinearization adapted to be consistent with this work.

further investigation.

In general, it is unclear what factors determine when one possessive system is used over another in Michif. This is left for future research (Sammons 2013d). For the purposes of this study, however, only forms from the French-origin and mixed systems are considered, since they have clear indicators of masculine/feminine gender, while the Cree-origin possessives do not.

Adjectives also reportedly agree for gender in Michif. As in French, a lexically-restricted set of adjectives (e.g., *groo* 'big', *pchii* 'small, little', *boñ* 'good') appears before the nouns they modify, as in (35)–(36), while the remainder appear after, as in (37):

(35) aeñ groo pascher kii-oshiihtaawak

aeñ	groo	pascher	kii–oshiihtaa–wak
a:MASC:SG	big	pasture	PST–make$_{VAIT}$–3PL:AN

'they made a big pasture'

(Victoria Genaille; 2012–10–15)

(36) eekwa lii suuyii kii-ataaweew por lii pchii garsoñ

eekwa	lii	suuyii	kii–ataawee–w	por	lii	pchii	garsoñ
and	the:PL	shoe	PST–buy$_{VAI}$–3SG:AN	for	the:PL	little	boy

'and he bought shoes for the little boys'

(Victoria Genaille; 2012–10–15)

(37) li bwaa ver

li	bwaa	ver
the:MASC:SG	wood	green

'green wood'

(Norman Fleury; 2012–10–13)

It has been claimed that adjectives in prenominal position agree with the nouns they modify for masculine and feminine gender when pre-posed within attributive constructions, but not when post-posed (e.g., Bakker 1997: 102; Bakker & Papen 1997):

(38) aeñ gro garsoñ

aeñ	gro–Ø	garsoñ
a:MASC:SG	big–MASC	boy

'a big boy'

(Verna DeMontigny; 2013–08–06; elicited)

(39) en gros fii

en	gro–s	fii
a.FEM:SG	big–FEM	girl

'a big girl'

(Verna DeMontigny; 2013–08–06; elicited)

In example (38), the adjective *gro* 'big' is used with a masculine referent, *garsoñ* 'boy', and is inflected as such. This contrasts with example (39), in which *gros* 'big', the feminine form of this adjective, is used to refer to a feminine referent, *fii* 'girl'. In (40), we see an example of an adjective being post-posed to the noun:

(40) mischetaw maana wiiya nimaama lii gordõ ee-oshiihaat avik la galet shoo

mischetaw	maana	wiiya	ni–maama	lii	gordõ
many.times	used.to	EMPH	1–mother	the:PL	crackling

ee–oshiihaa–t	avik	la	galet	shoo
CONJ–make.AN$_{VTA}$–3SG.CONJ	with	the:FEM:SG	bannock	hot

'my mum, she used to make cracklings many times with hot bannock'

(Mervin Fleury; 2012–10–13)

In this example, the adjective *shoo* 'hot' appears after the noun it modifies, *galet* 'bannock'. Although the noun is feminine, the adjective is not inflected for feminine agreement, as it was in (39), where the adjective is pre-posed to the noun. Despite Bakker (1997: 102) and Bakker and Papen's (1997) claims noted above, however, adjective agreement was not found to be consistent in my data. This was also found to be the case in Rosen (2007: 27). For the purposes of this study, it is therefore not a reliable category for testing gender distribution in Michif.

4.2.4 Animacy agreement

As in Cree, a great deal of grammatical agreement in Michif revolves around the animate/inanimate distinction. One area in which the animacy category of Michif nouns becomes visible is when nouns are used in conjunction with Cree-origin demonstratives. These demonstratives agree with both the animacy and number of their nouns. Examples (41)–(42) illustrate this distinction:

(41)

a. awa lañfañ
DEM:AN:SG child
'this child$_{AN}$'

b. ookik lii zañfañ
DEM:AN:PL the:PL child
'these children$_{AN}$'

(42)

a. ooma li suuyii
DEM:IN:SG the:MASC:SG shoe
'this shoe$_{IN}$'

b. oñhiñ lii suuyii
DEM:IN:PL the:PL shoe–IN:PL
'these shoes$_{IN}$'

In (41)a, the demonstrative pronoun *awa* indicates that the noun *lañfañ* ‘child’ is both animate and singular. This is in contrast to (42)a, in which the demonstrative pronoun *ooma* indicates that the noun *suuyii* ‘shoe’ is inanimate and singular. A different set of demonstratives is used for plural nouns. In (41)b, the demonstrative pronoun *ookik* indicates that the noun *lii zañfañ* ‘children’ is both animate and plural, while in (42)b, the demonstrative pronoun *oñhiñ* indicates that *lii suuyii* ‘shoes’ is both inanimate and plural. In general, the use of Cree-origin demonstratives in Michif is exceptional, since the noun phrase is otherwise entirely derived from French (Bakker 1990: 30).

Demonstratives also indicate the distance of the noun in relation to the speaker (proximal, medial, and distal):

(43)

a. ookik lii garsoñ piikishkweewak

ookik	lii	garsoñ	piikishkwee–wak
DEM:AN:PL:PROX	the:PL	boy	speak$_{VAI}$–3PL:AN

‘these boys are talking (close)’

b. anikik lii garsoñ piikishkweewak

anikik	lii	garsoñ	piikishkwee–wak
DEM:AN:PL:MED	the:PL	boy	speak$_{VAI}$–3PL:AN

‘those boys are talking (further away)’

c. neekik lii garsoñ piikishkweewak

neekik	lii	garsoñ	piikishkwee–wak
DEM:AN:PL:DIST	the:PL	boy	speak$_{VAI}$–3PL:AN

‘those boys are talking (even further)’

(Rosen & Souter 2009a: 40)

In (43), three different demonstrative pronouns are used. These demonstratives are all animate and plural, and differ only in the degree of proximity of the noun to the speaker. In (43)a, the proximal demonstrative *ookik* indicates that the referent *lii garsoñ* ‘the boys’ is close to the speaker, while in (43)b, the medial demonstrative *anikik* indicates that the

referent is further away. In (43)c, the distal demonstrative *neekik* indicates that the referent is even further away.

Demonstratives in Michif must be accompanied by a French-origin article immediately preceding the noun:

(44)

a. awa li garsoñ (*awa garsoñ)
awa li garsoñ
DEM:AN:SG the:MASC:SG boy
'this boy$_{MASC:AN}$'

b. awa la fii (*awa fii)
awa la fii
DEM:AN:SG the:FEM:SG girl
'this girl$_{FEM:AN}$'

c. ooma li papyi (*ooma papyi)
ooma li papyi
DEM:IN:SG the:MASC:SG paper
'this paper$_{MASC:IN}$'

d. ooma la bwet (*ooma bwet)
ooma la bwet
DEM:IN:SG the:FEM:SG box
'this box$_{FEM:IN}$'

e. anima la meezoñ (*anima meezoñ)
anima la meezoñ
DEM:IN:SG the:FEM:SG house
'that house$_{FEM:IN}$'

f. neema li shañ (*neema shañ)
neema li shañ
DEM:IN:SG the:MASC:SG field
'that field$_{MASC:IN}$'

g. oñhiñ lii zafer (*oñhiñ zafer)
oñhiñ lii zafer
DEM:IN:PL the:PL business
'those businesses$_{IN}$'

(adapted from Bakker & Papen 1997: 328; asterisk examples added by author)

From the combination of articles and demonstratives in the examples above, we see that *li garsoñ* 'boy' and *la fii* 'girl' are animate, while the remaining nouns are inanimate. We also see that *li garsoñ* 'boy', *li papyi* 'paper,' and *li shañ* 'field' are masculine, while *la fii* 'girl,' *la bwet* 'box,' and *la meezoñ* 'house' are feminine. Furthermore, the fact that both animate and inanimate nouns may be either masculine or feminine suggests that there is no relationship between animacy and masculine/feminine gender in Michif—that these are two separate categories acting independently of each other (Rhodes 1977: 10). We will return to this point in Chapter Seven. Since articles agree with their respective nouns for masculine/feminine gender and for number (Bakker 1997: 108–109), while demonstratives agree for animacy, noun phrases containing demonstratives provide a particularly salient example of Michif nouns being coded for both animacy and gender.

Although not widely reported for Michif or Plains Cree (although see Gillon & Rosen 2018; Wolvengrey 2011a for recent exceptions), this study finds evidence that obviation also plays a role in the selection of the demonstrative *anihi* 'those'. Most descriptions analyze this demonstrative as being used only for inanimate plurals (see, e.g., Bakker & Papen 1997:328; Rhodes 1977:13–14; Rosen & Souter 2009:74 for Michif and Okimāsis 2004:24; Ratt 2016:63 for Cree), but the corpus data used in this dissertation provide numerous examples in which this demonstrative pronoun is additionally used with obviative animate singular nouns:

(45) uh, Johnny LeDoux, soñ maama anihi

uh	Johnny	LeDoux	soñ	maama	anihi
uh	Johnny	LeDoux	3SG.POSS:MASC:SG	mum	DEM:AN:3'

'uh, Johnny LeDoux, his mum, that one'

(Norman Fleury; 2012–10–13)

(46) aakoshi li pchi garsoñ anihi
aakoshi li pchi garsoñ anihi
and.so the:MASC:SG little boy DEM:AN:3'
'and so the little boy'
(Rita Flamand; 2013–08–13)

(47) eekwa soñ, soñ bicycle anihi eekwa a ter aheew
eekwa soñ soñ bicycle
and 3SG.POSS:MASC:SG 3SG.POSS:MASC:SG bicycle

anihi eekwa a ter ah–ee–w
DEM:AN:3' and on ground put.AN_{VTA}–3>3'.IND–3SG:AN

'and now he puts his bicycle on the ground'
(Norman Fleury; 2013–08–24)

In (45) and (46), the demonstrative *anihi* is used to refer to *maama* 'mum' and *garsoñ* 'boy' respectively, both nouns which are naturally as well as grammatically animate in Michif. In (47), the noun *bicycle* is used with the demonstrative *anihi*. The use of the direct suffix indicates that a third person proximate argument (he) is acting on a third person obviative (the bicycle) (see section 2.4.3 below). The complete set of Michif demonstratives is shown in Table 14:

Table 14: Michif demonstratives (after Gillon & Rosen 2018: 94)

	Animate			**Inanimate**	
	Singular	**Plural**	**Obviative**	**Singular**	**Plural**
Proximal	awa	ookik	oñhiñ	ooma	oñhiñ
Intermediate	ana	anikik	anihi	anima	anihi
Distant[14]	naha	neekik	neehi	neema	neehi

[14] Though attested in much of the descriptive work on Michif, the distant set of demonstrative pronouns rarely appear in my data. This may be a result of the particular recording contexts and situations I have encountered, or it may be evidence that the use of these demonstratives is on the decline. Peter Bakker also reports that these rarely occurred in the context of his elicitation work or in texts (p.c.).

French-origin demonstratives are reportedly used on occasion in Michif (Bakker & Papen 1997: 328; Bakker 1997: 116):

(48) la sañdr kii-mishipahkistin is praeñtaeñ
la sañdr kii–mishi–pahkistin is[15] praeñtaeñ
the:FEM:SG ash PST–greatly–it.falls$_{VII}$ DEM spring
'we had an ash fallout this spring'
(Bakker 1997: 109)

(49) wii-wiiwiw staton
wii–wiiwi–w st–aton
VOL–get.married$_{VAI}$–3SG:AN DEM–fall
'she'll get married this fall'
(Bakker 1997: 109)

However, these examples are rare, appear in limited contexts, and are likely to be fossilized units rather than being able to be productively applied as demonstratives with any noun, as they are in French.

Interrogative pronouns are another area in which agreement for animacy occurs with Michif nouns. In some cases, the same interrogative pronoun is used regardless of the animacy or number of the referent (e.g., *taandee* 'where'). However, in others, different pronouns may be used depending on whether the referent is animate or inanimate, singular or plural:

(50) taana lom kaa-nashkweehtak?
taana lom kaa–nashkweeht–ak
which:SG:AN man REL–answer$_{VTI}$–3SG.CONJ
'which man answered?'
(Laverdure & Allard 1983: 355)

[15] This is presumably derived from the French demonstrative *ce*.

(51) taanima li liiv?
taanima li liiv
which:SG:IN the:MASC:SG book
'which book?'

(Rosen & Souter 2009a: 120)

(52) taanihi li liiv kaa-otinamaan?
taanihi lii liiv kaa–otinam–aan
which:PL:IN the:MASC:PL book REL–take$_{VTI}$–1SG.CONJ
'which books shall I take?'

(Laverdure & Allard 1983: 355)

In (50), the animate singular pronoun *taana* 'which' is used to refer to the animate referent *lom* 'man', whereas in (51), the inanimate singular pronoun *taanima* 'which' refers to the inanimate referent *li liiv* 'book'. In (52), a different pronoun, *taanihi* 'which', is used to refer to the plural inanimate referent *lii liiv* 'books'. The full set of Michif interrogative pronouns is listed below (cf. Okimāsis 2004:22 for Cree):

Table 15: Michif interrogative pronouns

Number	**Animate**	**Inanimate**	**English**
Singular	taandee	taandee	'Where is s/he/it?'
	taana	taanima	'Which one?'
	aweena	----	'Who?'
	keekway	keekway	'What?'
Plural	taandee	taandee	'Where are they?'
	taaniki	taanihi	'Which ones?'
	aweeniki	----	'Who?'
	keekway	keekwaya	'What?'

4.3 Verbs and verbal constructions

Michif verbs, which are primarily of Cree origin, exhibit considerable morphological complexity. A great deal of this morphology is sensitive to animacy, as are other aspects of the Michif verb such as verb stem selection. This section provides a brief overview of verbs in Michif and describes how the classification of nouns according to animacy affects various aspects of the Michif verb. Section 4.3.1 discusses Michif verb orders, while Section 4.3.2 discusses inflectional classes. This is followed by Section 4.3.3, which discusses verbal agreement categories, and by Section 4.3.4, which discusses hierarchical alignment.

4.3.1 Verb orders

Michif has three agreement paradigms that are grouped into what are called "orders" in the Algonquian literature, each of which are associated with certain syntactic functions or environments. The expression of animacy and other agreement categories depends on the order in which the verb is inflected. The three orders include (1) the independent order, which is used in main clauses (e.g., *kinakamonaawaw* 'you (pl). sing'); (2) the conjunct order, which is used in subordinate clauses (e.g, *ee-nakamoyeek* 'as/that you (pl). are singing'); (3) and the imperative order, which is used for commands (e.g., *nakamok*! 'you (pl.) sing!'). Each order can then be inflected for different verbal modes. Independent clauses inflect for only the independent indicative mode. Subordinate clauses may inflect for either the conjunct indicative or the conjunct subjunctive mode. Imperative order clauses may inflect for either the immediate imperative or the delayed imperative modes. Michif verb orders and their respective modes are illustrated in Table 16 below:

Table 16: Michif verb orders and modes

Order	**Mode**	**Word-form**	**Gloss**
Independent	Independent indicative	**ki**nakamo**naawaw**	'you (pl) sing'
Conjunct	Conjunct indicative	**ee-**nakamo**yeek**	'as/that you (pl) are singing'
	Conjunct subjunctive	nakamo**yeeko**	'if you (pl) sing'
Imperative	Immediate imperative	nakamo**k**!	'you (pl) sing!'
	Delayed imperative	nakamo**hkeek**!	'you (pl) sing later!'

Minimally, the Michif verb consists of a stem followed by a person/number suffix. In the independent order, subject or object person is marked by the first prefix (depending on whether the verb is transitive or intransitive, direct or inverse), and tense by the second:

(53) ganakamon
ni–ka–nakamo–n
1–FUT–sing$_{VAI}$–NON3:SG
'I will sing'
(Verna DeMontigny; 2013–08–22; elicited)

In the conjunct order, a conjunct marker is used, followed optionally by a tense or mood marker. Person and number are indicated by portmanteau suffixes:

(54) ee-wiinakamoyaan
ee–wii–nakamo–yaan
CONJ–VOL–sing$_{VAI}$–1SG.CONJ
'that I will sing'
(Verna DeMontigny; 2013–08–22; elicited)

In the independent order, the inflectional prefixes represent first and second person and tense. Preverbal modifiers, known as preverbs in the Algonquian literature, may also be added to a verb to indicate properties such as aspect (e.g., *maachi-* 'begin') and modality

(e.g., *nohtee-* ‘want’). Suffixes are used to express direction, person, and number, as shown in (55):

(55) gii-nohtee-waapamimaanaanik
ni–kii–nohtee–waapam–im–aa–naan–ik
1–PST–want–see$_{VTA}$–OBV.OBJ–NON3>3–1PL–3PL:AN
‘we (excl.) wanted to see them’
(Verna DeMontigny; 2015–07–16; elicited)

In the above example, person marking appears in three places. The prefix *ni-* indicates the presence of a first-person argument. The obviative object marker *-im* appears immediately after the verb stem. Person is also expressed through the first-person plural marker *-naan*, followed by the third person plural marker *-ik*. Finally, the direct marker *-aa* immediately follows the stem, signaling that a non-third person argument, higher on the hierarchy, is acting on a lower-ranked third person argument.

In contrast with the independent order, conjunct order verbs typically begin with a conjunct marker, which acts as a sort of complementizer and is used to mark subordination, such as in relative clauses or purposive clauses. Person and number are expressed solely through suffixes on the verb. Other possible suffixes include those marking obviation and the subjunctive mode:

(56) ee-kiikakweenohteewaapamikoyeekok
ee–kii–kakwee–nohtee–waapam–ikw–eekw–ik
CONJ–PST–try–want–see$_{VTA}$–3>1PL/2PL.CONJ–2PL.CONJ–3PL:AN
‘as/that they wanted to try to see you (PL)’
(Verna DeMontigny; 2015–07–16; elicited)

In example (56), the conjunct marker *ee-* serves as a complementizer.[16] The past tense marker *kii-* appears immediately after the conjunct marker, while the inverse marker *-ikw* occupies the position immediately following the stem, signaling that a low-ranked third person argument is acting on a higher-ranked non-third person argument. The suffix *-eekw* from the conjunct order person paradigm appears after the inverse marker, indicating the presence of a second-person plural argument. Finally, the portmanteau suffix *-ik* appears in the final position, indicating the plurality of the third-person argument using conjunct-order morphology.

In the conjunct subjunctive mode, no initial conjunct marker is used. Person and number are exclusively suffixing, and the subjunctive marker appears in the final suffix position, as shown in (57):

(57) waapamimaayeeko
waapam–im–aa–yeekw–i
see_{VTA}–OBV.OBJ–1PL/2PL>3.CONJ–2PL.CONJ–SUBJ
'if you (PL) see him/her/NA (OBV)'
(Grace Zoldy; 2014–08–16; elicited)

In (57), the obviative marker *-im* appears between the stem and the direction marker. The obviative marker indicates the presence of an obviative argument, while the direct marker *-aa* signals that a non-third person argument is acting on a third person. This is followed by *-eekw*, which indicates a second person plural argument. The suffix *-i* appears in the final slot, indicating that this clause is in the subjunctive mode.

[16] Note that conjunct and tense markers are often offset with hyphens in the orthrography, particularly in instances when the verb begins with a vowel.

4.3.2 Inflectional classes of verbs

Perhaps the most striking aspect of the animacy property in Michif is the degree to which it governs verbal agreement, both in terms of verb stem selection and inflection.

Verbs in Michif follow the traditional Algonquian pattern, falling into four different inflectional classes based on different combinations of transitivity (transitive, intransitive) and animacy (either the animacy of the subject for intransitive verbs, or the animacy of the object for transitive verbs) (Wolvengrey 2011a: 18). These verb classes are summarized in Table 17 below:

Table 17: Michif verb types (after L'Hirondelle et al. 2001: 41–2)

Verb Type	**Subject**	**Object**
Inanimate Intransitive (II)	Inanimate	None
Animate Intransitive (AI, VAIt)	Animate	None
Transitive Inanimate (TI)	Animate	Inanimate
Transitive Animate (TA)	Animate	Animate

Each of these verb classes has a different set of inflections. For intransitive verbs, there are two separate verb paradigms, one for intransitive verbs with animate subjects (Animate Intransitive) and one for intransitive verbs with inanimate subjects (Inanimate Intransitive). In examples (58)–(59), we see the contrast between AI and II verb inflection:

(58) li pwii paashteew
li pwii paashtee–w
the:MASC:SG well be.dry$_{VII}$–3SG:IN
'the well is dry'
(Verna DeMontigny; 2015–03–19; elicited)

(59) la shayer paashow
la shayer paashw–w
the:FEM:SG pail be.dry$_{VAI}$–3SG:AN
'the pail is dry'
(Verna DeMontigny; 2015–03–19; elicited)

In the above examples, the verb stem and inflection indicate that the arguments *li pwii* 'well' is inanimate, while the arguments *la shayer* 'pail' is animate. In (58), the subject *pwii* 'well' is used with an II verb stem, *paashtee-* 'be dry', indicating that it is an inanimate noun. By contrast, in (59), the subject *shayer* 'pail' is used with an AI verb stem, *paashw-* 'be dry', indicating that it is an animate noun.

Transitive verbs agree for animacy with the object, rather than the subject. This contrast is shown below:

(60) sii zhveu paasham
sii zhveu paash–am
3SG.POSS:PL hair dry$_{VTI}$–3OBJ:IN
's/he is drying his/her hair'
(Verna DeMontigny; 2015–03–19; elicited)

(61) soñ nañfañ paashweew
soñ nañfañ paashw–ee–w
3SG.POSS:MASC:SG child dry$_{VTA}$–3>3'.IND–3SG:AN
's/he is drying his/her child'
(Verna DeMontigny; 2015–03–19; elicited)

In the above examples, we see that *zhveu* 'hair' is inanimate, while *nañfañ* 'child' is animate because of the different verb stems and agreement suffixes that are triggered for

each. In (60), the use of the suffix *-am* indicates that the object is inanimate, while in (61), the use of the direct marker *-aa* indicates that an animate third person is acting on an obviative argument, which is animate.

There is also a set of verbs which are syntactically transitive in that they have an overt or implied inanimate object, but which are morphologically inflected as AI verbs rather than TI verbs. These are referred to as VAIts, VAI+Os, or pseudo-transitive verbs. An example of a verb which is syntactically transitive but morphologically intransitive is provided in (62):

(62) kahkiyaw mii zañfañ kii-aapachihtaawak lii kosh eñ laeñzh

kahkiyaw	mii	zañfañ	kii–aapachihtaa–wak
all	1SG.POSS:PL	child	PST–use$_{VAIT}$–3PL:AN

lii	kosh	eñ	laeñzh
the:PL	diaper	PREP	cloth

'all of my children used cloth diapers'

(Verna DeMontigny; 2012–10–15)

In this example, the verb *kii-aapachihtaawak* 'they used them' has an inanimate object, *lii kosh* 'diapers'. However, rather than being inflected as a TI verb, with the inanimate object marker *-am*, it is instead inflected as an AI verb.

As in Plains Cree, there are a number of Transitive Inanimate and Transitive Animate verb stems which occur in pairs in Michif, depending on the animacy of the object (after Wolfart 1973: 22):

(63) VTI: *otin-* 'take in.'
VTA: *otin-* 'take an.'

VTI: *waapaht-* 'see in.'
VTA: *waapam-* 'see an.'

VTI: *potishk-* 'put s.t. on, wear in.'
VTA: *potishkaw-* 'put an. on, wear an.'

In some cases, the stem forms are identical and the only indicator that they are distinct is the use of different agreement suffixes with which they occur. In other cases, the forms of the stems are slightly different, as in the verb for 'see' (above). In still other cases, the verb stems are completely different:

(64) VTI: *miichi-* 'eat in. (inanimate)'
VTA: *mow-* 'eat an. (animate)'

The choice of verb stem thus depends on whether the noun it agrees with is animate or inanimate:

(65)

a. niwaapahten aeñ shapoo
ni–waapaht–en aeñ shapoo
1–see$_{VTI}$–NON3:SG.IND a:MASC:SG hat$_{IN}$
'I see a hat'

b. niwaapamaaw aeñ shyaeñ
ni–waapam–aa–w aeñ shyaeñ
1–see$_{VTA}$–NON3>3–3SG:AN a:MASC:SG dog$_{AN}$
'I see a dog'

In (65)a, the object of the clause is the inanimate noun *shapoo* 'hat', which requires a TI verb stem, in this case *waapaht-* 'see IN', as well as TI verbal inflection. This is in contrast to (65)b, in which the object *shyaeñ* 'dog' is animate, triggering the TA verb stem *waapam-* 'see AN' and accompanying agreement suffixes. More information

regarding Michif verb inflection can be found in Bakker (1997), Bakker & Papen (1997), Gillon & Rosen (2018: 183–5), and Rhodes (1977), among other sources.

4.3.3 Verbal agreement categories

Michif verbs obligatorily agree with the person (1, 2, 3), number (singular, plural), animacy (animate, inanimate), and obviation status (proximate, obviative) of their referents. These distinctions are summarized in Table 18:

Table 18: Michif distinctions for person, number, animacy, and obviation

Gloss	Category
1SG	First person singular
1PL.EXCL[17]	First person plural exclusive
1PL.INCL[18]	First person plural inclusive
2SG	Second person singular
2PL	Second person plural
3SG:AN	Third person proximate singular (animate)
3PL:AN	Third person proximate plural (animate)
3':AN	Third person obviative (animate)
3SG:IN	Third person proximate singular (inanimate)
3PL:IN	Third person proximate plural (inanimate)
3'SG:IN	Third person obviative singular (inanimate)
3'PL:IN	Third person obviative plural (inanimate)

In the independent order, these categories are expressed through prefixes and/or suffixes, while in the conjunct and imperative orders, they are expressed solely through suffixes, as shown in the following examples:

[17] This is also represented as '1P' or '1PL' in the Algonquian literature.
[18] This is also indicated as '21' in the Algonquian literature.

(66) kinipwaahkan
ki–nipwaahka–n
2–be.smart$_{VAI}$–NON3:SG
'you're smart'

(Verna DeMontigny; 2014–08–25)

(67) ee-atoshkeeyaan
ee–atoshkee–yaan
CONJ–work$_{VAI}$–1SG.CONJ
'as/that I'm working'

(Harvey Pelletier; 2013–08–29)

In (66), which is an independent order verb, the prefix *ki-* indicates the presence of a second person subject, while the suffix *-n* indicates that this argument is singular. In (67), the prefix *ee-* indicates that the verb is inflected in the conjunct mode, while the suffix *-yaan* indicates that the subject is first person singular.

There are only two person prefixes in Michif, *ni-* for first person, and *ki-* for second person, and these are only used in the independent order. The third person is indicated solely by suffixes in all orders. Michif also distinguishes between inclusive and exclusive in the first-person plural to indicate whether or not the hearer is included, as illustrated in (68) and (69):

(68) ki-atoshkaashonaan
ki–atoshkee–isho–naan
2–work$_{VAI}$–REFL–1PL
'we (incl.) work for ourselves'

(Verna DeMontigny; 2014–08–05)

(69) lii pwer giimoshonaan, lii pabinaa

lii	pwer	ni–kii–mosho–naan	lii	pabinaa
the:PL	saskatoon.berry	1–PST–pick$_{VAIT}$–1PL	the:PL	cranberry

'we (excl.) picked saskatoons, cranberries'

(Norman Fleury; 2013–09–26)

The inclusive-exclusive distinction is fairly consistent throughout the independent order, though it seems to have eroded in the conjunct order.

Much of Michif verbal marking is dependent on verb mode, and number is often included with person on verbal suffixes in the form of portmanteau morphemes. All Michif verbs agree in number with their subjects. The following examples show number agreement for third-person singular and plural subjects, respectively:

(70) atoshkeew
atoshkee–w
work$_{VAI}$–3SG:AN
's/he is working'
(Harvey Pelletier; 2013–08–29)

(71) atoshkeewak
atoshkee–wak
work$_{VAI}$–3PL:AN
'they are working'
(Harvey Pelletier; 2013–08–29)

In (70), the verb is inflected for a third person singular subject with the suffix *-w*, while in (71), the suffix *-wak* indicates that the subject is a third person plural.

While the suffixes on intransitive verbs indicate the number and person of the subject, as shown in examples (70)– (71), suffixes on transitive verbs indicate the number and person of the object, which may be first, second, or third person, as shown in (72)–(73):

(72) niwaapamaaw
ni–waapam–aa–w
1–see$_{VTA}$–NON3>3–3SG:AN
'I see him/her/NA'
(Verna DeMontigny; 2013–09–24; elicited)

(73) niwaapamaawak
ni–waapam–aa–wak
1–see$_{VTA}$–NON3>3–3PL:AN
'I see them'

(Verna DeMontigny; 2013–09–24; elicited)

When attached to a Transitive Animate verb, as in (72), the suffix *-w* indicates that the object is both third person and singular, while in (73), the suffix *-wak* indicates that the object is a third person plural.

Obviative nouns also trigger obviative agreement on the verb, whether or not the noun itself is overtly marked as such:

(74) kii-wiicheehameew sa fii-wa

kii–wiicheeham–ee–w	sa	fii–wa
PST–live.with.AN$_{VTA}$–3>3'.IND–3SG:AN	3SG.POSS:FEM:SG	daughter– OBV

'she lived with her daughter'

(Norman Fleury; 2012–10–13)

(75) Atoshkeew dañ lapitaal. Soñ vyeu miina dañ lapitaal atoshkeeyiwa.

atoshkee–w	dañ	lapitaal
work$_{VAI}$–3SG:AN	at	hospital

soñ	vyeu	miina	dañ
3SG.POSS:MASC:SG	husband	too	at

lapitaal	atoshkee–iyi–w–a
hospital	work$_{VAI}$–OBV.SUBJ–3SG:AN–OBV

'She works at the hospital. Her husband works at the hospital too.'

(Lawrance Fleury; 2013–07–24)

In (74), the noun *fii* 'girl' is marked as obviative through the suffix *-wa*. The direct marker *-ee* indicates a third person argument acting on a third person obviative argument. Meanwhile, in (75), the noun *vyeu* 'husband' is not overtly marked as obviative, though it

triggers obviative agreement on the verb through the use of the suffix *-iyi*, which indicates that the subject (in this case, *vyeu* 'husband') is obviative. Thus, like animacy, obviation is a covert category because it is not always overtly marked on the noun itself, yet it triggers verbal agreement.

Various verbal suffixes are used to express obviation. When the subject of a clause is obviative, the suffix *-iyi* is used, as seen in (75). Note that this suffix is reportedly reserved for instances in which an obviative subject is possessed by another third person (Bakker 1997: 235). However, in my data instances were found in which this morpheme is also used for unpossessed third-person subjects:

(76) moo taapwee itweewak piihtikweeyit dañ la meezoñ

moo	taapwee	itwee–wak	piihtikwee–yi–t
NEG	for.sure	say_{VAI}–3PL:AN	$enter_{VAI}$–OBV.SUBJ–3SG.CONJ

dañ	la	meezoñ
PREP	the:FEM:SG	house

'they said she couldn't go into the house'

(Victoria Genaille; 2012–12–02)

If the object of a clause is obviative, the obviative object is marked on the verb with the suffix *-im*:

(77) kaa-waapamimak sa jeeñg, gii-waashtahikaan

kaa–waapam–im–ak	sa	jeeñg
REL–see$_{VTA}$–OBV.OBJ–1SG>3.CONJ	3SG.POSS:FEM:SG	girlfriend

ni–kii–waashtahikee–n
1–PST–wave$_{VAI}$–NON3:SG

'when I saw his girlfriend (obv), I waved'

(Verna DeMontigny; 2014–08–14; volunteered)

In (77), the suffix *-ak,* which is only used with Transitive Animate verbs in the conjunct order, indicates that a singular first-person subject is acting on a third person argument. The suffix *-im* in the verb *kaa-waapamimak* 'when I saw her (obv)', indicates that the object of the verb is obviative.

4.3.4 Hierarchical alignment

Grammatical relations are head-marked in Michif, which, like Algonquian languages in general, uses a hierarchical system of alignment based on a (2 > 1 > 3PROX > 3OBV) person hierarchy (Siewierska 2013). This hierarchy is shown in (79) below:[19]

(79) Michif person hierarchy:
2 > 1 > 3 > 3'

Whenever a higher-ranked subject acts upon a lower-ranked object, a direct suffix is added to the verb, indicating that the participants comply with the person hierarchy. Conversely, whenever a lower-ranked A acts upon a higher-ranked P, an inverse suffix is

[19] While this has been the traditional characterization of hierarchical alignment in Algonquian languages, and has been maintained here for clearer comparison against descriptions in the wider literature, both Macaulay (2009) and Oxford (2014) offer critical reassessments of this analysis.

added to the verb, marking a violation of the person hierarchy. This opposition can be seen in the following examples:

(80) kiwaapamin
ki–waapam–i–n
2–see$_{VTA}$–NON3>NON3.DIR–NON3:SG
'you (sg.) see me'
(Verna DeMontigny; 2013–09–24; elicited)

(81) kiwaapamitin
ki–waapam–iti–n
2–see$_{VTA}$–NON3>NON3.INV–NON3:SG
'I see you (sg.)'
(Verna DeMontigny; 2013–09–25; elicited)

Examples (80)–(81) differ only in the direction marker used, producing two different readings. For both, the prefix *ki-* indicates the presence of a second person argument, while the suffix *-n* indicates the presence of a non-third person argument. In (80), however, the direct marker *-i* is used, indicating that a higher-ranking A (in this case, second person) is acting on a lower-ranking P (first person). This is in contrast to (81), in which the inverse marker *-iti* indicates that a lower-ranking A (first person) is acting on a higher-ranking P (second person).

The full set of agent-patient combinations for these direction markers is given in Table 19. Note that there are separate markers for the independent and conjunct orders.[20]

[20] In the 1>2 category, the *-iti* and *-it* might be treated as the same marker, with the final *-i* being epenthetic when appearing before consonants. Since the independent order suffixes which would follow this form all begin with a consonant and the conjunct order suffixes all begin with a vowel, there is not sufficient evidence to determine whether or not these are separate markers.

Table 19: Direction markers by order

	Independent	**Conjunct**
2>1 (direct)	-i	-i
1>2 (inverse)	-iti	-it
1/2>3 (direct)	-aa	
1SG>3 (direct)		-ak
2SG>3 (direct)		-at
1PL/2PL>3 (direct)		-aa
3>1/2 (inverse)	-ikw	
3>1SG (inverse)		-it
3>2SG (inverse)		-ishk
3>1PL/2PL (inverse)		-ikw
3>3' (direct)	-ee	-aa
3'>3 (inverse)	-ikw	-ikw

In most cases, these direction markers only indicate person and alignment between the participants. However, in the conjunct order, there is a small set of portmanteau morphemes which also indicate number for interactions between local and non-local participants (i.e., *-ak* 1SG>3.CONJ, *-it* 3>1SG.CONJ, *-at* 2SG>3.CONJ, *-ishk* 3>2SG.CONJ):

(82) ee-waapamishk
ee–waapam–ishk
CONJ–see$_{VTA}$–3>2SG.CONJ
'as/that s/he sees you (sg)'
(Verna DeMontigny; 2013–09–24; elicited)

Note that the affixes appearing in the table are almost all identical for the independent and conjunct orders, with the exception of 3>3':

(83) waapameew
waapam–ee–w
see_{VTA}–3>3'.IND–3SG:AN
's/he (prox) sees him/her (obv)'
(Verna DeMontigny; 2013–09–25; elicited)

(84) ee–waapamaat
ee–waapam–aa–t
CONJ–see_{VTA}–3>3'.CONJ–3SG.CONJ
'as/that s/he (prox) sees him/her (obv)'
(Verna DeMontigny; 2013–09–25; elicited)

Some scholars (e.g., Bakker & Papen 1997: 314; Rosen 2007: 46) do not distinguish between direct markers in the independent and conjunct orders for 3>3' cases, listing 3>3 as simply *-ee*. By comparison, Dahlstrom (1986) identifies the same pattern for Plains Cree, *-ee* in the independent and *-aa* in the conjunct order, that we also find for Michif (26).

These direction suffixes have been glossed in multiple ways by different authors. For example, Dahlstrom (1986) glosses *-i, -ee*, and *-aa* all as simply DIR. On the other hand, Rosen (2007) glosses *-i* as NON3DIR, *-a* as simply DIR, and *-ee* as DBL3, indicating that there are two third person participants. In this work, these direct markers are analyzed as follows:

Table 20: Direction marker glosses

	Participants	**Order**	**Morpheme**	**Gloss**
Direct markers	2>1	independent/conjunct	-i	NON3>NON3.DIR
	1/2>3	independent	-aa	NON3>3
	1SG>3	conjunct	-ak	1SG>3.CONJ
	2SG>3	conjunct	-at	2SG>3.CONJ
	1PL/2PL>3	conjunct	-aa	1PL/2PL>3.CONJ
	3>3’	independent	-ee	3>3’.IND
	3>3’	conjunct	-aa	3>3’.CONJ
Inverse markers	1>2	independent/conjunct	-it(i)	NON3>NON3.INV
	3>1/2	independent	-ikw	3>NON3
	3>1SG	conjunct	-it	3>1SG.CONJ
	3>2SG	conjunct	-ishk	3>2SG.CONJ
	3>1PL/2PL	conjunct	-ikw	3>1PL/2PL.CONJ
	3’>3	independent/conjunct	-ikw	3’>3

Chapter 5: Animacy and gender assignment in Michif and its source languages

Michif is primarily comprised of elements from two major source languages, Plains Cree (Algonquian) and French (Romance/Indo-European). Each of these source languages has a different type of nominal classification system, one based on a formal animate/inanimate distinction, and another based on a formal masculine/feminine distinction. One of the primary questions in this study is how animacy and gender are assigned to nouns in Michif, especially in relation to animacy and gender assignment patterns observed in Michif's primary source languages. Specifically, we are interested in determining whether Michif animacy and gender values are assigned based on (a) natural gender, (b) the source language of the noun (i.e., French, English, or Cree), or (c) inheritance from Cree and French (in which case Michif animacy and gender values would parallel those of their translation equivalents in Cree and French). In this chapter, we examine previous claims in the literature regarding animacy and gender assignment in Plains Cree and French, as well as what has been reported for Michif. We begin with an overview of animacy assignment in Plains Cree (Section 5.1), followed by an overview of gender assignment in French (Section 5.2). In Section 5.3, we discuss previous claims made regarding animacy and gender assignment in Michif, while Section 5.4 provides a brief summary of the chapter.

5.1 Animacy assignment in Plains Cree

In Plains Cree, and in Algonquian more generally, nouns are assigned one of two animacy values, either animate or inanimate. Broadly speaking, nouns referring to living entities such as human beings, animals, birds, and the like are classified as animate, while nouns denoting non-living entities are classified as inanimate. However, there are a number of exceptions to this generalization, in which nominals which are notionally or naturally inanimate are classified as grammatically animate. These exceptions include certain body parts and items of clothing, tobacco and related items, and some machines, household items, and natural objects, among others. Thus, animacy is not fully predictable, despite these larger patterns (Mithun 1999: 98). These generalizations, along with apparent exceptions and examples, are summarized in

Table 21: [21, 22, 23]

Table 21: Animate nouns in Plains Cree

Semantic category	**Example(s)**	
HUMAN BEINGS	*nāpēw*	‘man’
	iskwēw	‘woman’
	ayahciyiniw	‘enemy, esp. Blackfoot’
	cīpay	‘dead person’
ANIMALS (INCL. BIRDS, FISH, REPTILES, INSECTS)	*mistatim*	‘horse’
	mostos	‘buffalo’
	kinosēw	‘fish’
	mōswa	‘moose’

[21] Forms in this table are drawn from (Ahenakew 1987; Bloomfield 1946: 19; Gillon & Rosen 2018; Hogmen 1981; Okimāsis 2004; Rhodes 2013; Wolfart 1973; Wolfart 1996, and Wolvengrey 2011b).
[22] Cree examples and orthography here follow Wolvengrey (2001).
[23] Bloomfield (1946: 94) reports that the noun *sihkowin* ‘spittle’ is animate as well, though Wolvengrey (2001) lists this as an inanimate noun.

	kihiw	‘eagle’
SOME ANIMAL BODY PARTS	*mēkwan*	‘feather’
	ēskan	‘horn, antler’
ANIMAL HIDES AND GARMENTS MADE FROM THEM	*wāposwayān*	‘rabbit skin’
	mostoswayān	‘buffalo robe’
	maskwayān	‘bear skin’
MOST TREES	*sihta*	‘spruce’
	māyi-mītos	‘black poplar’
	mistik	‘tree’
CERTAIN PLANTS AND THEIR PRODUCTS	*mahtāmin*	‘maize’
	picikwās	‘apple’
	ayōskana	‘raspberry’ (but not strawberry)
	aski-pahkwēsikan	‘flour’
	pahkwēsikan	‘bannock’
	wīhkihkasikan	‘cake’
	okiniy	‘rose hip/tomato’
SPIRITS	*kisē-manitōw*	‘God’
	ātayohkan	‘spirit being’
SOME HUMAN BODY PARTS	*nisakitikom*	‘my braid’
	nitasiskitān	‘my calf of leg’ (but not thigh)
	nitihtikos	‘my kidney’
	nitīhiy	‘my shoulder blade’
	niyihk	‘my gland’
	maskasiy	‘nail (finger or toe)’
	mitohtōsim	‘breast/teat’
TOBACCO AND RELATED ITEMS	*ospwākan*	‘pipe’
	cistēmāw	‘tobacco’
SOME NATURAL OBJECTS	*asiniy*	‘rock/stone’
SOME PERSONAL / HOUSEHOLD ITEMS	*āhcanis*	‘ring’
	kotawānāpisk	‘stove’ (also NI)
	askihk	‘pail; kettle’

	asām	‘snowshoe’
	sōniyāw	‘money’
	ēmihkwān	‘spoon’
PHENOMENA OF THE NATURAL ENVIRONMENT / CELESTIAL BODIES	*pīsim*	‘sun; moon’
	acāhkos	‘star’
	kōna	‘snow’
SOME ITEMS OF CLOTHING	*astis*	‘mitten/glove’
	tāpiskākan	‘scarf/necktie’
	mitās	‘trousers’
	asikan	‘sock/stocking’
SOME MACHINES / MEANS OF TRANSPORTATION	*sēhkēpayīs*	‘car’
OTHER	*tēwēhikan*	‘drum’

The basis for the assignment of animate and inanimate values to nouns in Algonquian languages has long been a subject of linguistic investigation (see Goddard 2002 for an overview). While several analyses have been put forward to account for the apparent discrepancies seen above, no real consensus has emerged among them. For instance, some have claimed that animacy assignment in Algonquian is not semantically motivated, but completely arbitrary (e.g., Greenberg 1954; Wolfart 1996: 398), whereas others have claimed that it is not arbitrary at all, but instead rests on cultural patterns which are at times imperceptible to outsiders (e.g., Hallowell 1955; 1976).

According to Corbett’s typology of nominal classification systems, Algonquian is a predominantly semantic system (1991: 58–60)—that is, semantic principles account for animacy assignment in a significant proportion of the lexicon, but there is a certain degree of semantic residue or number of exceptions for which semantic criteria cannot explain the observed assignment of animacy. Most analyses of animacy in Algonquian

rest on the assumption that there is a semantic distinction between living and non-living entities which is reflected in the grammar, and numerous attempts have been made to identify a single feature which could account for some of the more exceptional cases of animacy assignment in Algonquian (Wolvengrey 2011a: 48). One hypothesis is that, rather than living vs. non-living, the core distinction between these animacy categories is based on power, with entities that possess power marked as grammatically animate, and those without power marked as grammatically inanimate (Black-Rogers 1982 for Ojibwe; Darnell 1991: 99; Darnell & Vanek 1976 for Cree; Straus & Brightman 1982 for Cheyenne). This notion of what is and is not powerful is supposedly rooted in an Algonquian cultural perspective:

> Our main conclusion, then, is that gender in Algonquian is semantically based, but that the semantics are rooted in a culture which is difficult for the outsider to grasp. Fluidity is an essential part of the world view, with the result that gender assignment too can vary. There is, however, a tendency for nouns to remain animate, even if the motivation for this gender is lost for particular nouns, so that sporadic exceptions occur, which are no longer motivated for present speakers. Thus there will be synchronic exceptions to the semantic assignment rules. (Corbett 1991: 24)

Straus & Brightman (1982) claim that English borrowings provide further evidence of a semantic assignment system for Algonquian, as speakers of Cheyenne are easily able to assign animate or inanimate values to English borrowings without reference to the linguistic form (100). However, even this notion of "power" to explain the animate/inanimate distinction in Algonquian is problematic, as the attribution of power to certain inanimate objects over others would seem somewhat arbitrary:

> Of course, in order for this to be a fully valid explanation in Plains Cree, Cree speakers would have to attest to a belief in the spiritually powerful nature of animate *asikanak* 'socks' and *ayōskanak* 'raspberries' in contrast to inanimate *maskisina* 'shoes' and *otēhimina* 'strawberries'. In the

> absence of this, such examples tend to be used to refute a pure equation of the animate class with "living things" or the "spiritually active or powerful". Nevertheless, the prevailing attitude has always been one in which there is something about the animate class that marks the nouns so designated as special… (Wolvengrey 2011a: 48)

In fact, Wolvengrey (2011a) goes so far as to claim that, because of the large number of exceptions, Plains Cree has a grammatical rather than semantic system of animacy assignment (48). He also proposes an alternative to the characterization of animacy as being associated with whether or not the referent is "living", instead focusing on the notion of "life". This would account for some of the more exceptional cases of grammatical animates, such as *cīpayak* "ghosts", *manitowak* "spirits", which can be considered as elements of spiritual life, body parts involved in bringing about life (e.g. *mispayowak* "ovaries", *mitisowayak* "testicles"), and items of clothing which at one time may have been used to preserve life from exposure during the winter (e.g., *asāmak* "snowshoes", *astisak* "mitts", *mitāsak* "pairs of pants" and *asikanak* "socks") (48).

Dahlstrom (1995) takes yet another approach, claiming that, while semantics does influence animacy assignment to some degree, it is not entirely semantically motivated. Adapting Lakoff's (1987) notion of a "radial category" in his treatment of Dyirbal, she proposes that the animate category is composed of central members, peripheral members, and arbitrary members. Central members of the animate category are based on a semantic feature which they all have in common, while peripheral members of the group are linked by semantic extension from the central members. Finally, some exceptional members belong to this category which are unmotivated, as they do not possess the semantic feature shared by the central members, nor do they have a connection to the central members by means of semantic extension. Under this analysis, spiritually powerful

entities represent one possible semantic extension from the central feature of animacy. Thus, while power may not be the determining factor in animacy assignment, it may play a prominent role (1995: 57).

Based on patterns such as those presented in

Table 21, animacy in Algonquian has previously been described as *absorptive*, meaning that a naturally inanimate noun may be classified as grammatically animate, but that the reverse is never the case (Hockett 1966: 62). However, several instances of notional animates receiving inanimate values have been noted in the literature. For example, in Cree, some terms for so-called living objects, such as *miskīsik* 'eye,' *miskāt* 'leg,' *mitēh* 'heart,' *misit* 'foot,' *micihciy* 'hand,' *wāpakwaniy* 'flower,' are classified as grammatically inanimate, while others are classified as animate (Okimāsis 2004: 6). In addition, it has been observed in other Algonquian languages, such as Cheyenne, that an animate noun can be intentionally reassigned to an inanimate value to "convey the loss of power or social deprecation" (Kilarski 2007: 336, citing Straus & Brightman 1982: 114). Straus & Brightman (1982: 133) also note several nouns in Cheyenne which refer to sacred objects that show up as inanimate within the context of dictionary elicitation, even though they typically appear as animate in conversation and other contexts. This provides further support for the position that the animate value is not absorptive in Algonquian.

In addition to the above examples of notional animates receiving inanimate values, some noun stems in Cree may have different interpretations, depending on whether they are animate or inanimate. For example, *mistik* means 'tree' when grammatically animate, but 'stick' when inanimate (Wolfart 1996: 399), as shown in (85):

(85)

a. mistikwa
   mistikw–a
   tree/stick–IN:PL
   ‘sticks’

b. mistikwak
   mistikw–ak
   tree/stick–AN:PL
   ‘trees’
   (adapted from Gillon & Rosen 2018: 81, citing Wolvengrey 2011b)

In (85), we see two different but homophonous nouns which belong to different noun classes. In (85)a, mistik$_{IN}$ means ‘stick’, whereas in (85)b, mistik$_{AN}$ means ‘tree’.

It should also be noted that differences in animacy assignment have been found across Algonquian languages, and even across different dialects of the same language. For example, *ciste·ma·w* ‘tobacco’ is animate in Plains Cree, while the Munsee equivalent, *kwšá·tay*, is inanimate (Goddard 2002: 200). Different dialects of Cree also do not necessarily assign the same animacy values to the same words. For instance, Wolfart (1973: 23) reports that the noun *sōniyāw* ‘gold, money’ consistently appeared as inanimate in a text from Fort Vermillion in northern Alberta, while in other varieties it typically appears as animate.

Though not widely discussed in the literature, instances of both intra- and interspeaker variability in animacy assignment have also been reported in various Algonquian languages. For example, Wolfart (1973) notes some variability found in Plains Cree animacy, but states that this phenomenon is not well understood (20–22). In terms of interspeaker variability, Straus and Brightman (1982) report that in Cheyenne, some nouns may be animate for some speakers, and inanimate for others (113). Similar instances are noted in Ojibwe (Black 1967) and Penobscot (Quinn 2001). Goddard (2002)

also provides a nonce example of Meskwaki *kemešoˑmesenaˑna*$_{\text{AN}}$ 'our grandfather' being used as inanimate *kemešoˑmesenaˑni*$_{\text{INAN}}$ when used to refer to a ceremonial pole in a Meskwaki text (211). Despite this assortment of scattered remarks across languages and decades, however, very little is known about this phenomenon of variable animacy in Algonquian. As Goddard notes, "we have virtually no information on such variation for any language" (2002: 215). Such work, both on a language-specific basis and at a comparative level, represents an important area for future investigation in Algonquian nominal classification.

5.2 Gender assignment in French

The assignment of gender to French nouns was long thought to be completely opaque and unpredictable, as expressed by Bloomfield (1933: 280): "There seems to be no practical criterion by which the gender of a noun in German, French, or Latin could be determined" (cited in Corbett 1991:7). While it may not always be possible to predict the gender of a word from its meaning, many attempts have been made to account for French gender assignment based on formal criteria. For example, Bidot (1925) proposes an analysis based on a combination of semantics and orthography, while Mel'čuk (1974) proposes an analysis based on phonological endings (Ayoun 2010: 121). Another study, Tucker, Lambert & Rigault (1977), was particularly instrumental in identifying major patterns of gender assignment in French. In their detailed account, Tucker, Lambert & Rigault find that French gender assignment can largely be predicted by the phonological form of the noun, specifically the final phone, or in some cases, the penultimate and/or

antepenultimate phone (Corbett 1991: 58–60). As an example, the phonological assignment rules for French nouns ending in /ɔ̃/ are summarized below:

1. Nouns [ending] in /ɛzɔ̃/, /sjɔ̃/, /zjɔ̃/, /ʒjɔ̃/ and /tjɔ̃/ are feminine;
2. Remaining nouns in /ɔ̃/ are masculine (Corbett 1991: 60).

Thus, nouns such as *maison* /mɛzɔ̃/ 'house', *action* /aksjɔ̃/ 'action', *persuasion* /pɛrsɥazjɔ̃/ 'persuasion', *contagion* /kɔ̃taʒjɔ̃/ 'contagion,' and *question* /kɛstjɔ̃/ 'question' are feminine, while nouns such as *jambon* /ʒɑ̃bɔ̃/ 'ham', *rayon* /rɛjɔ̃/ 'shelf', *camion* /kamjɔ̃/ 'truck', and *baton* /batɔ̃/ 'stick' are masculine. These rules account for approximately 98.2% of the data in Tucker, Lambert, and Rigault's sample (Corbett 1991: 60–61). Corbett states that "[P]honological rules are [thus] powerful predictors of gender", and despite previous claims that French gender is intractable and unpredictable (e.g, Bloomfield 1933: 280), gender assignment can be predicted for French nouns with a great deal of regularity (Corbett 1991: 58–60).[24]

Nevertheless, Tucker, Lambert, and Rigault's (1977) analysis is not without issue. There are numerous exceptions to the patterns they identify which must be accounted for with complicated rules. In addition, the existence of grammatical homonyms, homophones, and epicene nouns, which have the same phonological forms but different genders, all provide evidence that French gender assignment cannot be reliably accounted for based on phonological form alone. As Ayoun (2018: 119) concludes, "[i]t may thus be best to accept that grammatical gender assignment in French is based on a mix of

[24] Tucker, Lambert & Rigault also discovered that "[d]eaf children who learn to speak French do not learn to assign nouns to gender" because they "cannot hear the language and so cannot discover the assignment rules", further bolstering their argument as to the importance of phonological rules in the assignment of French gender (1977: 59; Corbett 1991: 58–60).

morpho-phonological and lexical rules in addition to spelling and semantics, rather than insist on "crazy rules" (Enger 2009) that look like post-factum rationalizations (Comrie 1999)."

According to Corbett, French has a formal system of gender assignment, based on a combination of phonological and morphological criteria. As described in Chapter Three, he maintains that even in formal systems such as this, however, semantic rules take precedence over formal ones (1991: 58–60). That is, nouns are first assigned their gender values based on semantic criteria, and the residue is handled by either phonological or semantic criteria, or a combination of both. The semantic assignment rules for French are as follows:

1. Sex-differentiable nouns denoting males are masculine.
2. Sex-differentiable nouns denoting females are feminine. (Corbett 1991: 57–58)

For example, *père* 'father' and *oncle* 'uncle' are masculine, while *mère* 'mother' and *tante* 'aunt' are feminine. The same follows for all sex-differentiable nouns, with the exception of some so-called hybrid nouns, which take more than one set of agreements depending on the agreement target (e.g., *sentinelle* 'sentry') (Corbett 1991: 58–60, 225), and nouns which may be assigned to either gender depending on the gender of the referent (e.g., *un/une catholique* 'a Catholic, m/f').

There is also a morphological rule at work in the assignment of gender to French nouns:

1. Compound nouns formed from a verb plus some other element are masculine (Corbett 1991: 58–60, citing Tucker, Lambert & Rigault 1977:19).

The example given here is *porte-monnaie* ‘a purse’, which means ‘a carry-money’. Although *monnaie* ‘money’ by itself is feminine, the compound is masculine. Thus, in contrast, to Cree, which has a predominantly semantic system of gender assignment, French gender assignment is based on a combination of semantic and formal (phonological and morphological) criteria.

French also has approximately 50 grammatical homonyms, in which a change in gender is the result of a change in meaning (Ayoun 2010: 120):

(86)

a.	un	livre
	a:MASC:SG	book
	‘a book’	
b.	une	livre
	a:FEM:SG	pound
	‘a pound’	
c.	un	moule
	a:MASC:SG	mold
	‘a mold’	
d.	une	moule
	a:MASC:SG	mussel
	‘a mussel’	
e.	un	manche
	a:MASC:SG	handle
	‘a handle’	
f.	une	manche
	a:FEM:SG	sleeve
	‘a sleeve’	

(Ayoun 2010: 120)

A number of homophones, which have identical pronunciations, but differ in both spelling and meaning, are also found in French:

(87)

a. /sɛl/ *sel*$_{\text{MASC}}$ ‘salt’, *selle*$_{\text{FEM}}$ ‘saddle’
b. /fwa/ *foie*$_{\text{MASC}}$ ‘liver’, *fois*$_{\text{FEM}}$ ‘time’, *foi*$_{\text{FEM}}$ ‘faith’
c. /ru/ *roux*$_{\text{MASC}}$ ‘redhead’, *roue*$_{\text{FEM}}$ ‘wheel’
d. /rɛn/ *renne*$_{\text{MASC}}$ ‘reindeer’, *reine*$_{\text{FEM}}$ ‘queen’
e. /po/ *pot*$_{\text{MASC}}$ ‘jar’, *peau*$_{\text{FEM}}$ ‘skin’

French also has epicene nouns, which can appear with either masculine or feminine gender depending on the gender of the referent:

(88)

a. un/e artiste ‘an artist’
b. un/e juge ‘judge’
c. un/e propriétaire ‘owner’
d. un/e camarade ‘friend’
e. un/e pensionnaire ‘boarder’
f. un/e malade ‘sick person’
g. un/e partenaire ‘partner’
h. un/e stagiaire ‘trainee’

(Ayoun 2010: 120–1)

While gender agreement in French is thought to be fairly consistent, instances of variable gender have been found in Canadian French, as shown in (89)–(92):

(89) il s'en va en bas puis il allume le cheminée, quand la cheminée elle était bien embrayée

il	se	en	va	en	bas	puis
3SG:MASC	REFL	PREP	go	PREP	down	and

il	allume	le	cheminée
3SG:MASC	light	the:MASC:SG	chimney

quand	la	cheminée	elle
when	the:FEM:SG	chimney	3SG:FEM

était	bien	embrayée
was	well	set.in.motion:FEM

'he goes downstairs and he lights the chimney/fireplace; when the fire was going well'

(Klapka 2002: 15, cited in Gillon & Rosen 2018: 79)[25]

(90) c'est des vies différents

c'est	des	vies	différents
it.is	some	$life_{FEM}$:PL	different:MASC:PL

'it's different lives'

(Klapka 2002: 19, cited in Gillon & Rosen 2018: 79)

(91) bien il y a tout le temps—il y a des—des petits choses de—tu sais…

bien	il	y	a	tout	le	temps
well	3SG:MASC	there	has	all	the:MASC:SG	time

il	y	a	des	des	petits
3SG:MASC	there	has	some	some	small:MASC:PL

choses	de	tu	sais
$thing_{FEM}$:PL	of	2SG	know

'well there is always—there are—small things—you know…'

(Klapka 2002: 20, cited in Gillon & Rosen 2018: 79)

[25] Interlinearization in examples (89)–(92) is mine.

(92) bien, elle dit [sic], écoute, les derniers journées là, tu vas avoir de la visite

bien	elle	dit	écoute	les	derniers	journées
well	3SG:FEM	say	listen	the:PL	last:MASC:PL	day$_{FEM}$:PL

là	tu	vas	avoir	de	la	visite
there	2SG	FUT	have	PREP	the:FEM:SG	visit

'well, she says, listen, those last days there, you will have visitors'

(Klapka 2002: 20, cited in Gillon & Rosen 2018: 79)

In (89), the noun *cheminée* 'chimney/fireplace' is referred to twice, once with the masculine definite article *le*, and once with the feminine definite article *la*. In (90), the adjective *différents* 'different' is inflected for masculine plural, rather than the expected form *différentes*, even though its referent *vies* 'lives' is feminine. In (91), the adjective *petits* 'small' modifies the feminine noun *choses* 'things' and yet is inflected for masculine plural. The expected form in prescriptive French would be *petites*. Finally, in (92), the adjective *derniers*, rather than the expected *dernières* 'last', is used to modify the feminine noun *journées* 'days'.

Several differences in gender assignment have also been noted between so-called Standard French and Canadian French. For example, many vowel-initial nouns are feminine in Canada but masculine in France, presumably due to a process of reanalysis through frequent co-occurrence with the masculine demonstrative *cet* [sɛt], which is homophonous with the feminine demonstrative *cette*. The gender values assigned to English borrowings may differ between France and Canada as well (e.g., *stéréo, job*) (Gillon & Rosen 2018: 80). It is reasonable to assume that some degree of the variability in masculine/feminine gender values described here may have been present in the forms

of French that contributed to the development of Michif, and may continue in the language today.

5.3 Animacy and gender assignment in Michif

Most Michif speakers today are not fluent in French or Cree, yet they assign animacy values to French-origin nouns, and both animacy and gender values to English loanwords. Moreover, the animacy and gender values are not always predictable for every noun. In this section, we explore general patterns of alignment between the gender of nouns in Michif and its source languages as reported in the literature. Section 5.3.1 focuses on general animacy assignment patterns in Michif, while Section 5.3.2 discusses gender assignment patterns. Finally, Section 5.3.4 presents several hypotheses regarding grammatical gender in Michif.

5.3.1 Animacy assignment

It has often been reported or assumed that the animacy values of Michif nouns are identical to those of their Cree equivalents. However, these claims are often made without careful examination of animacy assignment patterns in both languages, and certainly do not involve naturally occurring speech to any appreciable degree. Indeed, according to Gillon and Rosen (2018: 94), the details of how French-origin nouns are assigned animacy values in Michif remain unclear.

Bakker (1997) claims that Michif speakers simply assign the same animacy value to the French noun as the corresponding Cree term:

> The Michif French nouns have the same animate or inanimate gender as their Cree semantic equivalents. In short, all the verbs are used exactly as

> they would have been used in Cree or, to be more precise, in the Plains Cree dialect. (1997: 7).

He goes on to note that the same idiosyncrasies that are found in Cree animacy assignment can also be found in Michif. For example, terms for means of transportation are animate, terms for trees may be either animate or inanimate, and terms for some household items are animate, while others are inanimate (1997: 99). Consider the following examples:[26]

(93) kii-michimineew aatiht larzhañ

kii–michimin–ee–w	aatiht	larzhañ
PST–hold$_{VTA}$–3>3'.IND–3SG:AN	some	money

'he kept part of the money'

(Bakker 1997: 99)

(94) kii-michiminam aatiht la peey

kii–michimin–am	aatiht	la	peey
PST–hold$_{VTI}$–3OBJ:IN	some	the:FEM:SG	pay

'he kept part of the payment'

(Bakker 1997: 99)

In the examples above, the animacy of the nouns is apparent in the person agreement suffixes that occur with the verbs. In (93), we see that Michif *larzhañ* 'money' is animate, presumably because the corresponding Cree term *sôniyâs* 'money' is also animate. Meanwhile, in (94), Michif *la peey* 'payment' is inanimate because the Cree equivalent is inanimate, and so on (Bakker 1997: 99).

By comparison, Papen (2003a) offers a somewhat more qualified description. He compares the animacy of selected terms from two semantic fields—fruit and vegetables and clothing—in Michif to their equivalents in two Cree dialects. He limits this study to

[26] Interlinearization and orthography adapted.

these two semantic fields because they both have a combination of animacy values (animate and inanimate) represented within them, even though they are notionally inanimate objects. He compares the animacy of the Michif terms with their equivalents in two dialects of Cree (Plains Cree and "le cri du Québec," which is presumably a variety of East Cree), but for our purposes, we focus only on Plains Cree here. The results of his study for fruit and vegetable terms are summarized in Table 22:

Table 22: Michif vs. Cree animacy values for fruit and vegetable terms (after Papen 2003a: 134)

Concept	**Michif**	**Plains Cree**	**Same?**
walnut	animate	animate	✓
gooseberry	animate	animate	✓
strawberry	inanimate / animate	inanimate	variable
raspberry	inanimate / animate	animate	variable
orange	animate	animate	✓
berry (wild)	inanimate	inanimate	✓
apple	animate	animate	✓
apricot	animate	?	?
corn	animate	animate	✓
potato	inanimate / animate	inanimate	variable
turnip	inanimate	inanimate	✓
cabbage	animate	inanimate	X
carrot	animate	animate	✓
radish	inanimate / animate	inanimate	variable

As shown above, Papen found variable animacy values for a few items related to food – 'radish', 'strawberry', 'raspberry', and 'potato'. He claims that this uncertainty between

berries might stem from the fact that 'strawberry' is inanimate in Cree, while 'raspberry' is animate. Papen also notes that 'strawberry' is inanimate in Laverdure & Allard (1983), the most widely used Michif dictionary, but that Michif speakers with whom he worked in Manitoba indicated that both animacy values were acceptable (Papen 2003a: 135).

With respect to clothing terms, we can see in Table 23 that agreement for animacy between Michif and the equivalent Cree terms is practically complete:

Table 23: Michif vs. Cree animacy values for clothing terms (after Papen 2003a: 136)

Concept	**Michif**	**Plains Cree**	**Same?**
clothing item(s)	inanimate	inanimate	✓
coat	inanimate	inanimate	✓
dress	inanimate	inanimate	✓
(winter) hat/beanie/tuque	inanimate	inanimate	✓
jacket	inanimate	inanimate	✓
slippers/boots	inanimate	?	?
(wool) sweater	inanimate	?	?
stockings/socks	animate	?	?
socks	inanimate	inanimate / animate	variable
mittens/gloves	animate	animate	✓
scarf	animate	animate	✓
pants	animate	animate	✓

All of the terms in the above table align for animacy between Cree and Michif except for 'socks', which is claimed by Papen to vary between animate and inanimate values in Cree (2003a: 136). Papen also makes some notes about the Michif terms used. First, the term for 'clothes' in Michif is *bitaeñ*, which comes from a Canadian French term *butin*

used specifically for personal garments or clothing. The Michif word for 'scarf', *krimon*, comes from the Canadian French word designating a warm, thick scarf, and is not known outside of Canada. Finally, the term for 'sweater' *swetur* is borrowed from English, and has the same animacy value as its Cree equivalent (Papen 2003a: 136).

According to Papen, it can generally be assumed that a Michif noun has the same animacy value of the corresponding Cree noun (2003a: 132). However, this does not immediately account for animacy assignment involving nouns referring to items that were not present when Michif emerged in the 19th century. In these cases, Papen posits a process of semantic analogy, in which a new Michif word takes on the animacy of an existing word in the language with which it shares certain properties. Examples of apparent semantic analogy are provided in Table 24:

Table 24: Examples of semantic analogy in Michif (Michif and French data from Papen 2003a: 136–137)[27]

French	**English gloss**	**Michif**	**Proposed analogical base**	**Plains Cree**	**Same?**
casque (dur)	'helmet/hard hat'	inanimate	*chapeau* 'hat'	inanimate	✓
(chaussures de) tennis / espadrilles	'tennis shoes/ espadrilles'	inanimate	*souliers* 'shoes'	inanimate	✓
culotte de golf	'knickers'	animate	*pantalon* 'pants'	animate	✓
cravate	'tie'	animate	*écharpe/foulard*	animate	✓
combinaison de motoneige	'snowsuit'	animate	*pantalon* 'pants'	animate	✓
tuque, casque de motocyclette	'motorcycle helmet'	inanimate	*chapeau* 'hat'	inanimate	✓
casquette de baseball	'baseball cap'	inanimate	*chapeau* 'hat'	inanimate	✓
abricot	'apricot'	animate	*pomme* 'apple' / *pêche* 'peach'	animate / inanimate	?
laitue	'lettuce'	inanimate	*feuille* 'leaf'	inanimate	✓

From the table above, we see that words like 'hard hat', 'motorcycle hat', and 'baseball cap' are all inanimate in Michif, presumably by semantic analogy from *shapoo* 'hat',

[27] English glosses and Michif animacy values added by the author.

which is also inanimate. This, however, does not explain why certain terms, such as ‘avocado’ and ‘lettuce’ are considered inanimate (Papen 2003a: 137).

In general, Papen finds that, for the two semantic fields under study, Cree animacy values are systematically attributed to their French equivalents when integrated into Michif, though some exceptions are noted, either due to variation or to processes of semantic analogy being applied (Papen 2003a: 136–137). Additionally, both Bakker (1997: 104–105) and Bakker & Papen (1997: 325) find that all Cree-origin verbs in Michif that have been nominalized with the Cree-origin nominalizer *-win* are inanimate. While Bakker’s (1997) and Papen’s (2003a) hypotheses concerning Michif animacy assignment patterns are valuable, they are based on only small samples of the Michif lexicon (less than 30 words in the latter case), using primarily introspective data and dictionary examples. These claims have also not been confirmed to apply to English borrowings in Michif, or to Michif nouns outside of these two semantic fields. This leaves many aspects of the relationship between Cree animacy and Michif animacy in question which might be addressed by investigation of the kind pursued in this study.

As noted in Section 5.1, the interpretation of some nouns in Cree can differ depending on the animacy value attached to it (e.g., *mistik* means ‘tree’ when animate and ‘stick’ when inanimate). However, Gillon and Rosen (2018) report that no similar pattern exists in Michif:

(95)

a. li bwaa anima
the:MASC:SG wood DEM:IN:SG
‘that firewood$_{IN}$’ (cut wood)

b. li graañ bwaa anima
the:MASC:SG big wood DEM:IN:SG
'that forest$_{IN}$'

(Gillon & Rosen 2018: 97)

(96) en tramb ana
a:FEM:SG tree DEM:AN:SG
'that tree$_{AN}$'

(Gillon & Rosen 2018: 97)

In (95)a, the noun *bwaa* is inanimate and the interpretation is 'firewood'. In (95)b, the same noun is also inanimate, though the interpretation when co-occurring with *graañ* 'big' is 'forest'. In (96), rather than applying an animate value to the noun *bwaa*, a different lexical item altogether, *tramb*, is used to arrive at the interpretation of 'tree'.

While a change in animacy does not result in a change in interpretation as in Cree, it has been reported that adding a possessive to a noun can result in a change of animacy value, as shown in examples (97)–(98):

(97)

a. nikiiwaapahten li kor anima
ni–kii–waapaht–en li kor anima
1–PST–see$_{VTI}$– NON3:SG.IND the:MASC:SG body DEM:IN:SG
'I saw that/the body$_{IN}$'

b. nikiiwaapamaaw soñ kor ana
ni–kii–waapam–aa–w soñ kor ana
1–PST–see$_{VTA}$–NON3>3–3SG:AN 3SG.POSS:MASC:SG body DEM:AN:SG
'I saw his/her body$_{AN}$'

(Gillon & Rosen 2018: 97)

(98)

a. nikiiwaapahten lii maeñ
ni–kii–waapaht–en lii maeñ
1–PST–see$_{VTI}$– NON3:SG.IND the:PL hand
'I saw the hands$_{IN}$'

b. nikiiwaapamaaw sii maeñ
ni–kii–waapam–aa–w sii maeñ
1–PST–see$_{VTA}$–NON3>3–3SG:AN 3SG.POSS:PL hand
'I saw his/her hands$_{AN}$'

(Gillon & Rosen 2018: 97)

In (97)a, the noun *kor* 'body' is inanimate when unpossessed, as indicated by the inanimate demonstrative pronoun *anima*. In (97)b, however, *kor* 'body' appears as animate when possessed, as indicated through the use of the animate demonstrative pronoun *ana*. Similarly, in (98)a, the noun *maeñ* 'hand' is inanimate when unpossessed, since it is the object of a TI verb, which only takes inanimate objects. This is in contrast to (98)b, in which a TA verb is used with the possessed form of *maeñ* 'hand', indicating that the object is animate.

5.3.2 Gender assignment

Much like animacy, it has been reported in the literature that Michif nouns of French origin generally follow the same masculine/feminine gender patterns as those of their French equivalents, though some exceptions are noted (Bakker 1997: 103). However, Papen (1987a) identifies several instances in which the gender of a Michif word does not correspond to the gender of its equivalent term in either European or Québécois French

(250). Table 25 provides a brief summary of masculine/feminine gender dissimilarities between Michif nouns and their French equivalents as reported in several sources. [28, 29]

Table 25: Reported dissimilarities between Michif and French gender

Michif Word	**French Equivalent**
bol$_{\text{FEM}}$ 'bowl'	*bol*$_{\text{MASC}}$
beut$_{\text{MASC}}$ 'butte, hill'	*butte*$_{\text{FEM}}$
bwatoñ$_{\text{FEM}}$ 'stick'	*bâton*$_{\text{MASC}}$
grif$_{\text{MASC}}$ 'claw'	*griffe*$_{\text{FEM}}$
kũti$_{\text{FEM}}$ 'county'	*comté*$_{\text{MASC}}$
kwatoñ$_{\text{FEM}}$ 'cotton'	*coton*$_{\text{MASC}}$
lii maaryii nakatitowak$_{\text{FEM}}$[30] 'divorce'	*divorce*$_{\text{MASC}}$
maesk$_{\text{FEM}}$ 'mask'	*masque*$_{\text{MASC}}$
mosh$_{\text{MASC}}$ 'fly'	*mouche*$_{\text{FEM}}$
muulaeñ$_{\text{FEM}}$ 'mill'	*moulin*$_{\text{MASC}}$
shev$_{\text{MASC}}$ 'goat'	*chèvre*$_{\text{FEM}}$
shoo$_{\text{MASC}}$ 'whitewash'	*chaux*$_{\text{FEM}}$

In this table, we see instances of both French masculine nouns becoming feminine in Michif and arbitrary French feminine nouns receiving masculine gender in Michif. Papen posits that one potential explanation for this so-called "gender confusion" (i.e., where there are mismatches between Michif and French gender values) is that vowels in a

[28] Michif spellings in this table are adapted to the orthography used in this work. In some cases, Michif terms were added by author as they were not provided in the original sources. It should also be noted that the genders of Michif terms provided in these sources do not necessarily align with those found in my corpus data or other sources (cf. Fleury 2018).
[29] Forms in this table are drawn from (Bakker 1997; Bakker & Papen 1997; Papen 1987a; Papen 2003a)
[30] Lit: 'the married (ones) race one another' (Fleury 2018; Wolvengrey 2001).

centralized and unstressed position have a tendency to become laxed, making it difficult to distinguish the vowels in *li* and *la* (Papen 1987a: 251).

Michif only has a handful of Cree-origin nouns, most of which are used to refer to berries, plants, and animals, as well as certain kinship relationships and a handful of household objects (Bakker & Papen 1997: 324). Cree-origin nouns differ somewhat from French-origin nouns in that they typically do not appear with French-origin articles or possessive markers, though this does occur occasionally: *moñ kashkihchikeewin* 'my earnings', *la oshipeehikeewin* 'the writing', *li weepinikeewin* 'the garbage', *aeñ aamo* a bee', etc." (Papen 1987a: 250).[31] Rhodes (2013) claims that all Cree-origin nouns in Michif receive masculine gender, unless they refer to animates that are notionally feminine (102). The following examples show Cree nominalizations in Michif and their associated gender values:

(99)

a. li pakamahikan
   li pakamahi–kan
   the:MASC:SG hit$_{VTI}$–NOM
   'the striker$_{MASC}$' (cf. Cree *pakamahwêw* 's/he beats him/her')

b. li metawaakan
   li metawaa–kan
   the:MASC:SG play$_{VAI}$–NOM
   'the toy$_{MASC}$' (cf. Cree *mêtawêw* 's/he is playing')

c. li weepinikeewin
   li weepinikee–win
   the:MASC:SG throw.out$_{VTI}$–NOM
   'the garbage$_{MASC}$' (cf. Cree *wêpinikêw* 's/he throws things')

[31] Spellings regularized.

d. en pooyoshk[32]
en pooyo–shk
the:FEM:SG quit$_{VTI}$–NOM
'a quitter$_{FEM}$' (cf. Cree *pôyow* 's/he quits')

e. aeñ dotamaakeehk
aeñ nitotamaakee–hk
the:MASC:SG ask$_{VTA}$–INDEF.ACTOR
'a panhandling$_{MASC}$' (cf. Cree nitotamâkêw 's/he asks people for things')

(Bakker 1997: 104–105)[33]

In the examples above, we see several Michif nominalizations of Cree-origin verbs. Examples (99)a and (99)b both use the instrumental nominalizer *-kan*, while in (99)c the abstract nominalizer *-win* is used. Example (99)d provides an example of the *-shk* nominalizer, which has a meaning of "repetitive V-er", while (99)e illustrates the nominalizer *-hk*, used to indicate an action performed by an indefinite actor (Bakker 1997: 104). In all of these cases, the nominalized verb appears in conjunction with either a definite or indefinite article denoting masculine or feminine gender. Note that all the nouns listed here have masculine gender, except (99)d which has a feminine animate referent *pooyoshk* 'a (female) quitter'. While not consistently marked, this co-occurrence of Cree-origin nouns with articles denoting gender shows that they are in fact classified for gender, just as other types of nouns in the Michif lexicon. This goes against Hogmen's (1981) assertion concerning Northern Michif that "there is no evidence that [Cree-origin nouns] are marked for gender in any way (87)".

Michif reportedly has some epicene nouns which can take either masculine or feminine gender, depending on the gender of the referent:

[32] Note that this form differs from that reported by (Gillon & Rosen 2018) in example (100).
[33] Spellings regularized and interlinearization added.

(100)

a. aeñ pooyoosh
a:MASC:SG quitter
'a quitter$_{MASC}$'

b. en pooyoosh
a:FEM:SG quitter
'a quitter$_{FEM}$'

(Rosen & Gillon 2017: 4)

In (100)a, the noun *pooyoosh* 'quitter' is masculine, as indicated by the use of the masculine indefinite article, while in (100)b, the same noun is feminine, as indicated by the use of the feminine indefinite article. However, there are some nouns which can have either male or female referents, though the gender does not change:

(101)

a. aeñ nañfañ
a:MASC:SG child
'a child' (masc. or fem.)

b. *en nañfañ
a:FEM:SG child
'a child$_{FEM}$'

(Rosen & Gillon 2017: 4)

In (101)a, the gender of the noun *nañfañ* 'child' is masculine, as indicated by the masculine indefinite article, though it can refer to either a male or female child. As shown in (101)b, it is ungrammatical to use the feminine indefinite article with this noun. Though not widely attested, there have been brief mentions in the literature that Michif nouns may exhibit some variability in their gender values. For masculine/feminine gender, one of the earliest mentions of variability is offered by Papen (2003a: 131), who notes this phenomenon in passing, but does not provide any further elaboration. Papen also notes that some French-origin nouns have variable/uncertain gender: *en/enn sigaret*

'a cigarette' (2005b: 337). Neither Bakker (1997) nor Hogmen (1981) report on variability in their respective studies. Drawing on a subset of the data from the present dissertation, Gillon and Rosen (2018) observe variability between the feminine article *en* 'the' and masculine article *aeñ* 'the' for *gournoy* 'frog'. They also report instances of variable gender which arose through metalinguistic discussion with Michif speaker Verna Demontigny, in which both *en fleur* (fem.) and *aeñ fleur* (masc.) 'flower' and *en tahkweminaan* (fem.) and *aeñ tahkweminaan* (masc.) 'chokecherry' were found to be acceptable (2018: 104–105). Variability in gender is thus attested in the literature on Michif, but has not received much dedicated attention to date.

5.3.3 Animacy and gender assignment in English-origin nouns

Over time, English has come to exert increasing influence on Michif, as most Michif speakers today are minimally bilingual in English and Michif. For most contemporary speakers, English represents the language that they use most often in daily communication, a fact which has linguistic consequences for the use of Michif which merit further investigation. English-origin nouns are increasingly being integrated into the Michif lexicon and receive both masculine/feminine and animate/inanimate gender values in Michif, though it is not clear how these values are assigned. These English-origin forms have not inherited any animacy or gender values, since English does not classify nouns based on either animate/inanimate or masculine/feminine gender. An examination of how Michif assigns both animacy and gender values to nouns that are borrowed from English may provide some insight into the synchronic productivity of these systems.

Like French-origin nouns, English nouns that are borrowed into Michif must be accompanied by an article or possessive, which provides evidence of masculine/feminine gender. They are also classified as having animate/inanimate gender, as reflected in the forms of accompanying demonstratives and verbal agreement (Bakker & Papen 1997: 325). Consider the following example: [34]

(102) wiiya li trak pamineew por la meel

wiiya	li	trak	pamin–ee–w
EMPH	the:MASC:SG	truck	look.after$_{VTA}$–3>3'.IND–3SG:AN

por	la	meel
for	the:FEM:SG	mail

'he drives the mail truck'

(Bakker 1997: 105)

In (102), we see that the English borrowing *li trak* 'truck' is masculine because it is used with the masculine definite article *li* 'the', possibly corresponding in gender to the Canadian French equivalent *le char*, which is masculine. Meanwhile, we see that *la meel* 'mail' is feminine, as indicated by the use of the feminine definite article *la* 'the'. This may or may not be because the corresponding French word *la poste* is also feminine. In addition, a TA verb is used here, indicating that both the actor 'he' and goal 'mail truck' are animate. The animate status of 'truck' is likely due to influence from Cree, in which vehicles and other means of transportation are animate (Bakker 1997: 106).

Bakker claims that the fact that both of these categories of nominal classification are retained for English-origin Michif nouns in examples such as those above is proof that "double relexification" has taken place—first, from English to French, and then from

[34] Orthography and interlinearlization adapted to be consistent with this work.

French to a Cree syntactic framework (Bakker 1989b: 347). This hypothesis is perhaps plausible historically, and might be reasonable to assume synchronically if Michif speakers who are borrowing English words actually spoke French—but most do not. For most contemporary Michif speakers, the productive integration of English borrowings clearly cannot rely on personal knowledge of French or Cree equivalent forms to inform the assignment of animacy and gender values to nouns.

In general, the assignment of animacy and gender values to English nouns in Michif is an area which has received very little attention in the literature and has been only treated briefly in studies, if at all. Indeed, Bakker identifies the assignment of animacy and gender to English nouns as an area which is poorly understood (Bakker 1997: 105). The observations that have been made tend to focus on the alignment of English nouns with French gender, but generally say little about the alignment of Michif nouns of English origin with Cree animacy. In his study on the Buffalo Narrows variety of Michif (i.e., "Northern Michif"), Hogmen (1981) reports that nouns in this variety of Michif have inherited their animacy values from Cree. He provides a couple of examples (e.g., *liskif* 'skiff$_{\text{INAN}}$', cf. Cree *osi*$_{\text{INAN}}$ 'boat, canoe' and *libok*$_{\text{INAN}}$ 'book', cf. Cree *masinahikan*$_{\text{INAN}}$ 'book, letter, paper'), but does not elaborate any further (87). Indeed, Papen notes a degree of uncertainty as to whether or not the general alignment with Cree animacy observed for French-origin nouns in Michif would apply equally well to English-origin nouns, and identifies this as a future research direction (2003a: 139). I have found no other explicit mention of how animacy values for English-origin forms in Michif and their Cree counterparts compare in the literature to date.

The assignment of masculine/feminine gender to English borrowings in Michif has also received very little attention. Papen (2003a) even goes so far as to say "Pour le moment, nous n'avons aucune explication pour l'attribution du genre masculine ou féminin aux termes empruntés de l'anglais" (132), a point echoed by Bakker's (1997) observation that "[g]ender assignment of English nouns is not understood very well" (105). A common hypothesis has been that English loanwords in Michif simply assume the same gender of the equivalent French term that they replace (Bakker 1997; Hogmen 1981; Papen 2005b: 337). While this does hold true in some cases, Papen (2005b) further observes other instances in which the gender assigned to the English loanword does not align with that of its French equivalent (e.g., *la dam*$_{\text{FEM}}$ 'dam' from French *le barrage*$_{\text{MASC}}$; *li sutkees*$_{\text{MASC}}$ 'suitcase' from French *la valise*$_{\text{FEM}}$) (337). Instances of misalignment reported in various sources are provided in Table 26, in which we see examples of both masculine and feminine nouns being assigned the opposite gender in Michif:

Table 26: Reported gender of English borrowings and their French equivalents[35]

Michif Borrowing	**French Equivalent**	**Source**
aeñ baynder$_{\text{MASC}}$ 'binder'	*une moissoneuse*$_{\text{FEM}}$	(Bakker 1997: 105)
aeñ blaek ay$_{\text{MASC}}$ 'a black eye'	*un œil au buerre noir*$_{\text{MASC}}$	(Bakker & Papen 1996: 18)
la dam$_{\text{FEM}}$ 'dam'	*le barrage*$_{\text{MASC}}$	(Bakker 1997: 105; Papen 2003a: 131)
la faektrii$_{\text{FEM}}$ 'factory'	*l'usine*$_{\text{FEM}}$	(Bakker & Papen 1996: 18)
la fun$_{\text{FEM}}$ 'fun'	*le plaisir*$_{\text{MASC}}$*, le fun*$_{\text{MASC}}$	(Bakker 1997: 105)
li kuter$_{\text{MASC}}$ 'cutter'	*la vedette*$_{\text{FEM}}$	(Bakker & Papen 1996: 18)
la kuushin$_{\text{FEM}}$ 'cushion'	*le cousin*$_{\text{MASC}}$	(Papen 1987a: 250)
aeñ paañsyoñ$_{\text{MASC}}$ 'pension'	*une pension*$_{\text{FEM}}$	(Papen 2003a: 131)
li saydiñg$_{\text{MASC}}$ 'siding'	*le revêtement*$_{\text{MASC}}$	(Bakker & Papen 1996: 18)
la slee$_{\text{FEM}}$ 'sleigh'	*le traîneau*$_{\text{MASC}}$	(Papen 2003a: 131)
la stiim$_{\text{FEM}}$ 'steam'	*la vapeur*$_{\text{FEM}}$	(Bakker & Papen 1996: 18)
li staf$_{\text{MASC}}$ 'stuff'	*la matière*$_{\text{FEM}}$	(Papen 1987a: 250)
aeñ suutkees$_{\text{MASC}}$ 'suitcase'	*une valise*$_{\text{FEM}}$	(Papen 2003a: 131)

There is also an anecdotal hypothesis common among scholars that all English borrowings in Michif receive a masculine gender value by default. In addition to examples found in my own fieldwork, even a cursory glance at the table above shows that this is not the case. These points will be further explored in light of the results of this study in Chapter Seven.

[35] Spellings regularized. Also note that the gender values reported here do not necessarily align with those found in my corpus data or other sources (cf. Fleury 2018).

5.3.4 Hypotheses regarding grammatical gender in Michif

As Gillon and Rosen (2018) note, the historical emergence of Michif through intense language contact involving French and Cree creates the conditions for a range of potential linguistic outcomes. More specifically, given the presence of a sex-based system in French and an animacy-based system in Cree, Gillon and Rosen propose three broad categories of possible nominal classification systems that might result from this kind of contact situation (2018: 99):

1. A language consisting of only sex-based gender, in which animate and inanimate values are lost, mapping masculine and feminine gender values onto the few instances of Algonquian-origin nouns in the lexicon;
2. A language consisting of only animacy, mapping animate and inanimate values onto the French-origin nouns (e.g., all masculines are inanimate, and all feminines are animate, or vice versa); or
3. A language which maintains both animate/inanimate and masculine/feminine systems of nominal classification.

Additionally, I propose a fourth possibility:

4. A language which maintains neither animacy nor gender.

It might be possible that several of the central grammatical features of Michif introduced in Chapter Two may make some of these options more implausible than others. For one, it was noted that animacy plays a central role in the overall organization of Michif grammar, governing not only nominal and verbal agreement, but also often the selection of phonologically distinct verb stems. All other things being equal, it is unlikely that a feature so central to Michif grammatical structure would be abandoned. Options 1 and 4

therefore do not seem probable. Option 2, on the other hand, seems much more likely, given that sex-based gender is not as crucial to the workings of French, in addition to a general tendency in languages towards simplification of marked forms with low functional load (Gillon & Rosen 2018: 99). Option 3 would seem unlikely, as it is the most complicated of the possibilities.

Gillon and Rosen (2018) also make a number of predictions regarding nominal classification in Michif. In particular, they predict that the sex-based gender system will erode over time, while animacy will remain stable (101). Specifically, arbitrary feminine nouns will be lost and instead assume default masculine gender, while semantically-based features will remain stable (104).

5.4 Summary

In this chapter, we have seen that Michif inherited two distinct systems of noun classification, an animate/inanimate distinction inherited from Cree, and a masculine/feminine distinction inherited from French. Some of the literature suggests that both nominal classification systems are still active in Michif, especially given the reported assignment of both masculine/feminine and animate/inanimate values to borrowings from English, but further investigation is required to determine the nature of these assignment patterns, as several researchers have noted (e.g., Gillon & Rosen 2018; Stoltzfus & Boissard 2016). Nevertheless, the picture of Michif nominal classification that emerges is generally not one of simplification, but of retention of these categories from both of its primary source languages:

> Although both the French and the Cree components are slightly altered from the source languages, these alterations simplify both components

> only a tiny bit, not what one would expect in cases of trade languages. In fact, Michif seems to combine the most complex parts of both languages. The Cree verb is infinitely more complex than the French verb, and the French noun, with its arbitrary genders and definite-indefinite distinction, is more complex than the Cree noun. In trade contacts, the source languages tend to lose a large part of their bound morphemes and semantic distinctions such as gender and number. This did not occur in Michif. So, both for historical and linguistic reasons, Michif cannot have been a language used in trade contacts (Bakker 1997: 277–278).

That Michif maintains the complexity of nominal classification systems from both of its major source languages, despite most speakers not being able to speak either language, has been called "theoretically problematic" by some (Bakker & Papen 1997: 315–316). At the very least, the fact that Michif has largely preserved both of these systems shows that it did not undergo a process of simplification, and is potentially even more complex than its source languages (Bakker 1997: 14). In Chapter Seven, we examine the results of our quantitative investigation and compare them against many of the generalizations and predictions in the literature summarized here.

Chapter 6: Data and methodology

As seen in the preceding chapters, it has often been reported or assumed in the literature that the animacy and gender values of Michif nouns are identical to those of their Cree and French equivalents (e.g., Bakker 1997; Papen 2003a). While valuable, these studies leave open a range of questions concerning animacy and gender assignment in Michif. For one, it remains unclear whether different strategies are employed to assign animacy and gender values to French-origin, English-origin, and Cree-origin nouns in Michif, or how animacy and gender values have come to be assigned to nouns drawn from source languages where these categories are not present (Bakker 1997: 105; Gillon & Rosen 2018: 94). While some authors have additionally noted apparent mismatches between Michif animacy and gender values and those of their source language equivalents (e.g., Papen 2003a), it is unclear what might motivate these cases. Finally, several studies have made predictions as to the productivity of nominal classification in Michif and the directionality of predicted changes to these systems over time, arguing that certain aspects of Michif nominal classification may be fossilized or in a process of gradual erosion to grammatical default values and/or semantically motivated categorizations (Gillon & Rosen 2018; Stoltzfus & Boissard 2016). In addition, the claims made in this literature have been based on small, hand-selected samples of Michif nouns gathered through elicitation and/or from dictionary examples, and have not generally been tested empirically against observations of nominal classification drawn from broader and potentially more representative samples of Michif language in use.

In this study, I aim to address several of these outstanding questions by examining the extent to which both Algonquian-origin animacy and Romance-origin gender have been transferred and/or maintained in Michif, as well as whether or not these systems remain productive or have become fossilized or lost altogether. We also investigate the relationship between Michif animacy and gender values and those of their Cree and French translation equivalents to determine the extent to which these nominal classification systems align with those of their source languages. Finally, we examine how English borrowings in Michif receive their animacy and gender values, since they cannot be attributed to inheritance from Cree or French.

To do this, I will examine the distribution of animacy and gender values among all Michif nouns found in a corpus of spoken Michif. Data for this study come primarily from projects involving fieldwork conducted in Manitoba and Saskatchewan during the period of 2011–2016. One result of this fieldwork has been the creation of a corpus of annotated audio and video recordings of contemporary spoken Michif across multiple speakers, varieties, and genres. In this chapter, I discuss how this corpus was constructed (Section 6.1), how the Michif noun dataset upon which this study is based was developed (Section 6.2), and the classification of lemmas in the dataset (Section 6.3).

6.1 Corpus construction

The corpus used for this study, assembled in collaboration with Métis communities in western Canada, is based on a subset of approximately 60 hours of audiovisual recordings of connected Michif speech representing multiple speakers, communities, and genres. The material gathered during my fieldwork represents the first audiovisual corpus of

Michif, and includes speakers of three major Michif varieties (Northern Michif, Southern Michif, and Michif French). This corpus complements uncontrolled, naturalistic speech with a standard set of controlled linguistic tasks. In this section, I describe various aspects of the construction of this corpus, including the contributors (6.1.1), data collection (6.1.2), annotation (6.1.3), and data management (6.1.4).

6.1.1 Contributors

A total of 42 Michif speakers were consulted during the course of corpus development, representing a number of Michif-speaking communities in western Canada. These include Southern Michif speakers in Camperville, Manitoba, and southeastern Saskatchewan; Northern Michif speakers in Île-à-la-Crosse, Saskatchewan; and a speaker of Michif French in St. Ambroise, Manitoba. In Manitoba, speakers were consulted in Binscarth, Brandon, Camperville, Gambler Reserve, Russell, and St. Ambroise. In Saskatchewan, speakers were consulted in Île-à-la-Crosse, Saskatoon, and Yorkton. Because Michif speakers are geographically dispersed across western Canada and the northern United States, it was often necessary to travel long distances with one or more speakers from one community to another to facilitate conversation sessions. The sampling of speakers and recording sites was based on speakers' availability and interest in language work. While most of this fieldwork has been conducted with speakers of Southern Michif in southwestern Manitoba, work has also been done with speakers in other Métis communities.

The L1 Michif speakers represented in this corpus were born between 1926 and 1962, and can be grouped into three broad categories based on the variety of Michif that

they speak as their first language – Northern Michif, Southern Michif, and Michif French. Two L2 speakers of Southern Michif were also present for some of the recording sessions and occasionally appear in the corpus. The distribution of speakers represented in the corpus, including the two L2 speakers, is presented in Table 27:

Table 27: Distribution of Michif speakers in corpus

	Northern Michif	**Southern Michif**	**Michif French**[36]	**TOTAL**
Males	10	12	1	23
Females	8	13	0	21
TOTAL	**18**	**25**	**1**	**44**

6.1.2 Data collection

This corpus contains four main types of recordings: (1) spontaneous speech, (2) controlled tasks, (3) targeted elicitation, and (4) annotation. The corpus of nearly 75 hours of spontaneous naturalistic speech is complemented by recordings of speakers performing a standard set of controlled linguistic tasks. These recordings are further supplemented with recordings of linguistic elicitation sessions targeting specific grammatical topics (e.g., applicatives, vocabulary items, adjective agreement, reflexives, verb inflection, etc.). Development of the written transcripts of all of the recordings is ongoing, and extensive oral annotation of corpus materials has been provided to facilitate these efforts, as described in Section 6.1.3 below. Additional sessions with Michif speakers focusing on annotation (i.e., transcription, translation, and/or oral annotation of

[36] One speaker doubled as both a Southern Michif and a Michif French speaker. This individual is counted here in the Southern Michif column, as the majority of his contributions were in this variety.

documentary recordings) were also recorded, and form an additional component of the corpus. A summary of the Michif corpus is provided in Table 28:

Table 28: Michif corpus

	Northern Michif	**Southern Michif**	**Michif French**	**TOTAL**
Spontaneous speech	07h24m	65h59m	01h15m	**74h39m**
Controlled tasks	02h04m	20h20m	00h59m	**23h24m**
Targeted elicitation	14h54m	36h31m	00h00m	**51h25m**
Annotation	01h11m	161h38m	00h00m	**162h50m**
TOTAL	**25h35m**	**284h27m**	**02h14m**	**312h18m**

The Michif language projects through which these recordings were developed have focused primarily on documenting spontaneous interactions between Michif speakers and cover a range of speech situations, such as conversations, narratives, procedural descriptions, and prayers. These sessions were largely conducted in Michif by Michif-speaking interlocutors and took place in speakers' homes or familiar public spaces, such as the local Métis office. Sessions were recorded in accordance with recommendations from documentary linguistics for best practices in linguistic data management, thus ensuring that data are consistently recorded and processed in such a way that the materials produced may be easily shared and permanently archived (Conathan 2011; E-MELD 2006; Good 2011; OLAC 2018; Thieberger & Berez 2012). Conversation and narrative sessions were video-recorded using a Canon VIXIA HF S-30 digital video camera with a wide-angle lens in MPEG-4 AVC (H.264) format (at full high-definition, i.e., 1920x1080 resolution and 24 full frames per second). Audio was simultaneously recorded for these sessions using either a Marantz PMD 661 or a Zoom

H6 solid-state digital audio recorders in WAV format (48 KHz, 24-bit uncompressed LPCM) onto SD media, often with an Olympus LS-10 solid-state audio recorder serving as a secondary recording device for backup. To ensure the highest audio quality, a Countryman E6i head-mounted omnidirectional condenser microphone was used wherever possible, depending on the number and preferences of participating speakers. Otherwise, either Countryman B3 or Audio Technica AT831b lavalier microphones (for two speakers) or the Røde NT4 tabletop stereo condenser microphone (for three or more speakers) were used.

Once a recording session was completed, all recordings were transferred from the recording devices to a computer, where the audio and video tracks for each session were synchronized manually using Final Cut Pro, a non-linear video editor. This alignment of audio and video tracks typically replaced the audio recorded on any built-in microphones on the video camera with the higher-quality audio recorded by the dedicated audio recording devices and attached microphones, resulting in single recordings that contained the highest quality audio and video available for each session. These synchronized recordings were subsequently exported from Final Cut Pro in both full resolution (for archival preservation alongside the original, free-standing audio and video source tracks produced by the recording equipment listed above) and lower resolution (as smaller working copies that could be annotated more easily with documentary linguistic software tools). The latter versions were then annotated, as discussed in Section 6.1.3.

Wherever possible, contributors who were represented in the conversation sessions also completed a set of controlled linguistic tasks. These include Pear Film and Frog Story narrations (Chafe 1980; Mayer 1969), Toy Game sessions (McDonough &

Lachler 2012), Totem Field Storyboard narrations (Totem Field Storyboards 2013), and Topological Relations Picture Series descriptions (Bowerman & Pederson 1993). These guided tasks complement the spontaneous conversations and narratives included in the corpus, and represent Northern Michif, Southern Michif, and Michif French, allowing for future systematic comparison of speech between contributors of all three Michif varieties. Pear Film narrations have been completed for 27 of the 42 speakers represented in the corpus, while Frog Story narrations have been completed for 24 speakers. The breakdown of recordings of these controlled tasks is given in the following table, where we see that the majority are for the Southern Michif variety.

Table 29: Controlled tasks

Task	**Northern Michif**	**Southern Michif**	**Michif French**	**Total**
Pear Film Narrations	6	20	1	**27**
Frog Story Narrations	6	17	1	**24**
Topological Relations Picture Series	1	3	0	**4**
Toy Game Sessions	0	3	1	**4**
Totem Field Storyboards	0	7	0	**7**
Total	**13**	**50**	**3**	**66**

These sessions were audio recorded using the same selection of solid-state audio recorders and microphones noted in Section 6.1.2. The Toy Game sessions were additionally video recorded using the standards described above as well.

Another component of corpus development involved targeted elicitation of specific lexical and grammatical data. These sessions complement and often draw on the spontaneous component of the corpus and have been useful in filling in inflectional

paradigms for which only some elements are attested in spontaneous speech, exploring lexical derivation, and more. These sessions were always audio recorded at the same standards described above in Section 6.1.2, typically with a head-mounted microphone. Pen and paper notes were taken during these sessions in field notebooks. In the later stages of the project, notes were taken with a LiveScribe Smartpen and notebooks, producing digital PDF documents of handwritten notebook pages that contain embedded, time-aligned audio (Martinez 2014). Some elicitation also took place remotely via Skype. This process is described in Section 6.1.3.

There are several advantages to developing a corpus of this type. For one, no comparable resources currently exist for Michif, and, given the advanced state of language shift and loss affecting many Métis communities, the creation of a permanent collection of examples of proficient spoken Michif is particularly relevant for supporting ongoing efforts in both Michif linguistic research and language revitalization. From the perspective of linguistic research, much of what has been reported in the scientific literature on Michif has relied primarily on elicitation, invented example sentences drawn from a published Michif dictionary (Laverdure & Allard 1983), and/or translation tasks (Bakker 1997). While these sources of linguistic information certainly represent important contributions in their own right to our present understanding of Michif, it is clear that they offer only partial representation of the structure of the language, particularly as far as discourse-related and socially marked features of the language are concerned. A corpus consisting of connected, unprompted, interactional language in context is therefore helpful in drawing a distinction between what speakers *can* say and what they *do* say (Mithun 2007; Sammons 2015). As other researchers have noted (e.g.,

Mithun 2014), non-naturalistic sources of linguistic data may diverge from conventional patterns of language use in the speech community, being influenced both by the nature of the linguistic tasks at hand (e.g., transfer effects from the source language into the target language in bilingual elicitation) and by speakers' self-reporting of their own linguistic practices (e.g., underreporting instances of variability that are disfavored or socially marked within the speech community; cf. Labov 1996). For the present study, contextualized examples drawn primarily from observed, rather than elicited, language use in the present corpus thus provide an important source of counterbalance to other possible sources of linguistic information.

Annotation sessions form another significant component of this documentation. These sessions involved one-on-one meetings with speakers, much like elicitation sessions. However, rather than focusing on answering structural questions about the language, the emphasis here was on developing written transcriptions and free translations of the recordings of both spontaneous speech and the controlled tasks. Much as in elicitation sessions, these meetings often involved significant Michif language use, with careful repetitions of utterances from source recordings being a central and recurring component of these meetings, in addition to cultural and metalinguistic commentary. Accordingly, these sessions were recorded and treated as another element of the documentation of Michif assembled in these projects. While not the focus of analysis in this study, these materials are valuable as additional documentation of Michif.

6.1.3 Annotation

As is common practice in documentary linguistics (Austin 2010; Good, Myers & Nakhimovsky 2010), annotations of audiovisual materials were developed primarily in

ELAN, a standard documentary linguistic software tool for time-aligned annotation (Brugman & Russel 2004; Max Planck Institute for Psycholinguistics 2016). A standard template for ELAN transcripts was developed at the outset of this project. This template was structured to contain the following sets of tiers for each speaker present in the recording: (1) Michif orthographic transcription (of each utterance), (2) English translation (of each utterance), and (3) notes. The notes tier includes assorted information, such as clarification of speaker names (when nicknames are used) and place names, cultural context, lexical and grammatical explanations of utterances, dialect differences, notes on the recording situation (such as background noise), and notes on the annotation process, such as dates that annotation began and ended, and flags on annotations that need to be reviewed. Annotation was always completed in consultation with fluent Michif speakers. This involved first pre-segmenting recordings in ELAN into utterances, then meeting in person with speakers to review, transcribe, and translate each of these segments. These annotation sessions were also audio recorded using the procedures outlined in Section 6.1.3 and constitute an additional source of documentation.

While written annotation was the primary focus of this stage of corpus development, these processes faced challenges similar to those of many other documentary linguistic projects, where the amount of recorded material exceeds the capacity of the project to provide even minimal, bilingual written annotations (Boerger 2011; Woodbury 2003). Indeed, common estimates of bilingual written annotation of documentary linguistic recordings estimate that it takes at least four times as long as the original recording to transcribe it, potentially longer (Sakel & Everett 2012: 107).

However, even this estimate assumes some familiarity with the language, and for multi-speaker recordings, the length of time needed to transcribe a recording increases exponentially with each additional speaker present. In my experience, a single-speaker recording typically takes significantly less time to transcribe than a recording with three speakers present, where there are multiple instances of speech overlap. During the initial stages of the project, I found that, for multi-speaker recordings especially, it took approximately one hour to fully transcribe and translate one minute of speech, even with the help of a native speaker present the entire time. It quickly became apparent that it would not be feasible to transcribe and translate all, or even most, of the corpus materials at this rate. Given these estimates, it would take anywhere from 600–1,200 hours to arrive at a minimal first pass transcription and translation, before the review and correction phase. At this rate, for a corpus of over sixty hours of speech, a large portion of which involves multiple speakers, the transcription and translation phase would take approximately 3600 hours, or 3.5 years at a rate of four hours a day, five days a week. In contexts of extreme language endangerment such as this, it is not feasible or even desirable to spend this much time focused solely on transcription. Even if it were possible to employ a researcher with adequate knowledge of the language and transcription tools and methods for that amount of time, it is doubtful that it would be possible to employ a native speaker for that amount of time as well, both in terms of time and financial resources. Furthermore, speakers' time is extremely valuable, and they are also often tapped for multiple language-related roles in the community, providing an additional reason why this type of work is prohibitive.

With these things in mind, rather than leave a substantial portion of these recordings entirely unannotated, this project instead adopted oral annotation methods to allow for some of this basic interpretative apparatus to be provided in spoken form and offset some of the bottleneck associated with completing full written transcription (see, e.g., Cruz 2016; Rosenblum & Sammons 2014). In oral annotation workflows, a recording is reviewed in real time by a commentator, who occasionally pauses playback to offer spoken commentary. This commentary is recorded as an annotation directly associated with the corresponding segment of the original recording. Combined with current technologies for recording and managing such time-aligned information automatically, oral annotation methods offer several important features that complement traditional written annotation workflows:

- *Rapid development:* As real-time methods, oral annotation typically allows recorded materials to be annotated more rapidly than through written annotation alone, often significantly reducing the overall time required on the part of individuals contributing to annotation;
- *Tandem transcription:* In many cases, oral and written annotation are able to proceed in tandem, with oral annotators progressing through recorded materials at their own pace without the immediate constraint of slower written transcription;
- *Long-term accessibility:* With the additional information that they incorporate, orally annotated materials are typically more amenable to written transcription and analysis in the future. This can significantly improve the long-term accessibility of materials for which full, written transcription may not be immediately feasible.

According to Boerger (2011), oral annotation workflows can reduce the time commitment required on the part of native speakers by half compared to written annotation. However, although there are clear benefits, there are challenges in adapting this method to multi-speaker recordings (Cruz 2016; Rosenblum & Sammons 2014). An oral annotation workflow was developed for this project which focused on single-speaker or two-speaker recordings where there was not significant overlap between speakers. This project implements a software-assisted oral annotation workflow using SayMore software (SIL International) which manages associations between source segments and corresponding oral annotations. However, rather than recording with the lower quality, built-in microphone available on the computer that is running SayMore, which the program assumes as its default, these oral annotations were recorded using a Zoom H6 solid-state digital audio recorders in WAV format (48 KHz, 24-bit uncompressed LPCM) connected to the computer via USB. A Countryman E6i head-mounted omnidirectional condenser microphone was also used.

In addition to the oral and written annotation that took place in situ, I also developed and implemented a methodology for remote transcription for instances in which in-person meetings with speakers were not feasible (e.g., due to geographical distance from consultants). This involved using Skype video conferencing software to share a view of an ELAN transcript and the corresponding audio and/or video with a Michif consultant and continue transcription and translation in this way. These annotation sessions were recorded in WAV format on an Olympus LS-10 with an attached Audio Technica ATR3350 cardioid condenser lavalier microphone in headset configuration, producing high-quality, uncompressed WAV recordings of any repeated Michif phrases

or English commentary. These recordings were then shared via Dropbox. Backup recordings were made using Call Recorder software for Skype. Since time that can be spent working with speakers in person is limited, this allowed us to continue to make progress on annotation even while separated by distance.

6.1.4 Data management

All metadata for the corpus is stored in an Excel database, using fields that map directly onto OLAC metadata categories (OLAC 2018). All contributing speakers completed a questionnaire detailing sociolinguistic characteristics such as age, place of birth, origin of parents, other languages spoken, education, gender, places lived, and lengths lived in each location, which is included as part of the metadata of the corpus. Additional metadata provide details about each specific recording session (e.g., date, location, topics, as well as the recording equipment used and its placement in the physical space). All recordings, transcripts, and database contents are regularly backed up on several media, including two external hard drives (one stationary master-copy and a matching travelling copy), as well as on a remote network server that conducts regular tests for data integrity.

6.2 Dataset construction

Having introduced the corpus upon which this study is based, we now turn specifically to the construction of the dataset focusing on animacy and gender marking in Michif. To facilitate the discussion of how animacy and gender values were retrieved from the corpus, Section 6.2.1 provides a brief summary of Michif animacy and gender marking. It also describes how the corpus was queried to retrieve all nouns that might be expected to provide information on animacy and gender, as well as which items were subsequently

excluded upon further review. Section 6.2.2 discusses the coding of the dataset for features related not only to the observed animacy and gender values, but also concerning natural animacy and gender and the animacy and gender values of equivalent nouns in Michif's primary source languages.

The present study focuses on animacy and gender assignment in the Southern Michif subset of the corpus, rather than in Michif French or Northern Michif, as animacy and gender values either cannot generally be retrieved or are not as robustly attested in the latter varieties. In the case of Northern Michif, which shows considerably less influence from French than is found in Southern Michif, masculine/feminine gender values appear relatively infrequently in the few attested French-origin nouns (cf. Ahenakew 2009; Hogmen 1981). By contrast, Michif French does not exhibit animacy marking. In addition, as the present study focuses on native speakers' assignment of animacy and gender to Michif nouns, any tokens from L2 speakers were removed from the final dataset to avoid potential interference as a result of imperfect learning. As such, the dataset described below draws exclusively on the Southern Michif subset of the corpus.

6.2.1 Querying the corpus

As discussed in Chapter Two, animacy and gender are observable in several constructions in Michif. Animacy is primarily retrievable from verbal inflection and demonstratives, while gender is most consistently evident in singular definite and indefinite articles and possessive adjectives. This agreement for animacy and gender is exemplified in (103)–(104):

(103) oh, mitoni kawatishin ooma moñ, ma keu ooma

oh	mitoni	katawashishi–n	ooma
oh	really	be.beautiful$_{VII}$–NON3:SG	DEM:IN:SG

moñ	ma	keu	ooma
1SG.POSS:MASC:SG	1SG.POSS:FEM:SG	tail	DEM:IN:SG

'oh, my tail is really beautiful'

(Verna Demontigny; 2013–09–18)

In (103), we see that the noun *keu* 'tail' is both inanimate and feminine. It is inanimate because it agrees with the Inanimate Intransitive verb *kawatashin* 'it is beautiful', as well as with the Cree-origin inanimate singular demonstrative *ooma* 'this one'. This noun is feminine as shown through its agreement with the French-origin feminine possessive adjective *ma* 'my'. Note also that the speaker first uses the masculine possessive adjective, before correcting herself and using the feminine form instead. In contrast, the noun *pwasoñ* 'fish' in (104) below is both animate and masculine:

(104) li pwasoñ awa, nawahtahtam anima li kroshee

li	pwasoñ	awa	nawahtaht–am
the:MASC:SG	fish	DEM:AN:SG	bite$_{VTI}$–3OBJ:IN

anima	li	kroshee
DEM:IN:SG	the:MASC:SG	hook

'this fish here bites that hook'

(Verna Demontigny; 2013–09–18)

In this example, the noun *pwasoñ* 'fish' occurs with the Cree-origin animate singular demonstrative *awa* 'this one', showing that it is animate. It also occurs with the French-origin masculine singular definite article *li* 'the', showing that it is masculine. This is in contrast to the noun *kroshee* 'hook', which appears in the same utterance, but with the

inanimate singular demonstrative *anima* ‘this one’, indicating that it is inanimate rather than animate. As noted in Chapter Five, some sources additionally report agreement for gender in prenominal adjectives. However, this was not found to be a reliable indicator of gender in the corpus data considered here, and is not considered further.

Drawing on the contents of the bilingual Michif-English written annotations found in the ELAN transcripts in this corpus (see Section 6.1.3 above), I was able to identify and extract all possible occurrences of nouns in the corpus that may contain information about their animacy and gender. ELAN’s multi-transcript search function made it possible to perform an exhaustive search for all possible noun tokens in the transcribed corpus, even in the absence of prior part-of-speech tagging. The first step in extracting these tokens involved querying the corpus, using regular expressions to search across multiple transcripts. Since nouns in Michif must be accompanied by a definite article, indefinite article, or possessive marker, this query was divided into three separate searches:

1. Retrieve all annotations that contain an indefinite article: *aeñ/aen/en/une/un*
2. Retrieve all annotations that contain a definite article: *li/le/la/lii*
3. Retrieve all annotations that contain possessive adjectives: (a) Michif *soñ/son/sa/sii*; (b) English *my/your/his/her/its/our/their/’s*

The first search sought to identify all nouns in the corpus associated with the indefinite articles *aeñ* ‘a(n) (masc.)’ and *en* ‘a(n) (fem.)’. The second query searched for all possible instances and spellings of the Michif definite articles *li* ‘the (MASC:SG)’, *la* ‘the (FEM:SG)’, *lii* ‘the (PL)’ in the corpus. Finally, a search for all possessive forms was performed. This search consisted of two separate queries. I first ran a search over all the Michif lines in the corpus for the Michif possessive markers of French origin: *soñ*

'his/her/its (MASC:SG)', *sa* 'his/her/its (FEM:SG)', and *sii* 'their (PL)'. However, this search alone would not locate all possessive forms in the corpus. In the uncommon case where a Cree-origin possessive is used in Michif (e.g., *noohkom* 'my grandmother', *nimoshoom* 'my grandfather'), the associated noun typically does not co-occur with a French-origin possessive, so the first search for Michif possessive forms would not return any of these instances. In addition, a simple search for the Cree possessive marker *o-* would have returned too many false hits. To address this, I performed a second search on just the English tiers for any instance of the English possessive adjectives "my", "your", "his", "her", "its", "our", "their", and "'s". In this way, all instances of possessives from the corpus should have been returned, be they of French or Cree origin.

Once all of these searches were performed, the results were compiled into a single spreadsheet in Microsoft Excel. A total of 5,175 possible tokens were initially extracted from these searches. These data were then carefully examined, and 3,972 tokens were excluded from further analysis using the linguistic and extralinguistic criteria outlined below.

Duplicates

Any duplicate items returned by these three searches (e.g., any annotations that contained both a Michif-origin possessive form in its transcription and an English possessive in its translation, which would have thus been retrieved separately in both the first and third searches given above) were eliminated, such that each record in the spreadsheet represented one unique instance of a noun token in the Michif corpus. As oral annotation sessions formed part of the corpus being queried, some searches unintentionally retrieved

annotations that represented direct repetitions of part of another recording (e.g., a careful repetition of a sentence found in one of the spontaneous or structured speech event recordings). These duplicate annotations were removed, as the animacy and gender values associated with the original occurrences and the repetitions were the same.

False hits

In some cases, the corpus searches returned false hits, or forms that met the search criteria, but were not relevant to this study. In particular, the search for *la* retrieved several items that were homophonous with the definite article, but that were used in different ways. For example, in some cases, a search for the article *la* turned up the homophonous interjection:

(105) “ooh la la,” ihtweew, “li dibrii!”

ooh	**la**	**la**	ihtwee–w	li	dibrii
ooh	INTERJ	INTERJ	say_{VAI}–3SG:AN	the:MASC:SG	tripe

“ooh **la la**,” she said, “tripe!”

(Norman Fleury; 2012–10–13)

In other cases, the *la* found was the adverb *la* ‘now’, rather than the definite article:

(106) eh, li pchi garsoñ pee-itohteew la, avik soñ bike

eh	li	pchi	garsoñ	pee–itohtee–w
eh	the:MASC:SG	little	boy	come–go$_{VAI}$–3SG:AN

la	avik	soñ	bike
now	with	3SG.POSS:MASC:SG	bike

'eh, a little boy is coming **now** with his bike'

(Irene Fleury; 2013–08–27)

There is also an adverb *la* meaning 'there':

(107) 'ka dañ li grañdrii kii-ashtaawak la

eeka	dañ	li	grañdrii
NEG.CONJ	in	the:MASC:SG	granary

kii–ashtaa–wak	**la**
PST–put$_{VAIT}$–3PL:AN	there

'they didn't put [it] in the granary **there**'

(Mervin Fleury; 2013–07–17)

The search for definite articles also retrieved several instances in which *li* or *la* occurs within a verbal predicate, rather than being used as a definite article. This is common in cases where French and English nouns and adjectives are being "converted" into Michif verbs. In these cases, *li* and *la* function as a "dummy element" rather than as definite articles (Bakker 1997: 114–115):

(108) eekoshpii kii-shoohkee-li-Michifiwinaan

eekoshpii	kii–shoohkee–**li**–Michif–iwi–naan
at.that.time	PST–strong–the:MASC:SG–Michif–COP–1PL

'back then we were strong Michif people'

(Norman Fleury; 2012–10–13)

(109) lii, lii pear anikik, kiiyaapich li-veriwiwak

lii	lii	pear	anikik
the:PL	the:PL	pear	DEM:AN:PL

kiiyaapich	**li**–ver–iwi–wak
still	the:MASC:SG–green–COP–3PL:AN

'those pears, those ones, they're still green'

(Norman Fleury; 2013–08–24)

(110) ayish ooma kaa- [inc] la-vyee-iwiyen, la-vyee-iwiyen kahkiyaw (…)

ayish	ooma	kaa–	[inc]	**la**–vyee–iwi–en
because	DEM:IN:SG	REL	inc	the:FEM:SG–old– COP–2SG.CONJ

la–vyee–iwi–en	kahkiiyaw
the:FEM:SG–old– COP–2SG.CONJ	all

'because as you're getting older, you're all getting older (…)'

(Victoria Genaille; 2012–10–15)

On occasion, *li* also appears in the corpus as a sort of fossilized French copula:

(111) li pa tro vyoo, eh, li pa tro vyoo, li pa tro ana Louisa, eh

li	pa	tro	vyoo	eh	**li**	pa	tro	vyoo
COP	not	too	old	eh	COP	not	too	old

li	pa	tro	ana	Louisa eh
COP	not	too	DEM:AN:SG	Louisa eh

'she's not too old, eh, she's not too old, she's not too (old), that one, Louisa, eh'

(Mervin Fleury; 2013–07–24)

Only three instances of this construction were found in the corpus, all of which were produced by the same speaker, so this appears to be rare. These were excluded from the dataset as well.

Unusable data

Other instances were excluded due to the quality of the underlying recording or because of hesitations and false starts. For example, in some cases, even though it was clear that an article was being used with a noun, it was not possible to determine what the noun was due to overlap, rapid speech, and other effects of the recording situation such as background noise:

(112) waapahtam anihi, waapameew anihi lii, lii pear akota dañ li [inc]

waapaht–am	anihi	waapam–ee–w	anihi	lii
see_{VTI}–3OBJ:IN	DEM:IN:PL	see_{VTA}–3>3'.IND–3SG:AN	DEM:IN:PL	the:PL

lii	pear	akota	dañ	**li**	**[inc]**
the:PL	pear	there	on	the:MASC:SG	[inc]

'he saw it, he saw them, the pears there on **the [inc]**'

(Cecile Burroughs; 2013–07–30)

Other hits consisted of false starts, hesitations, and the like. These were also excluded:

(113) hockey kaa-ashteek boy li, mitoni kii-moshkineew anima la meezoñ, eh

hockey	kaa–ashtee–k	boy	**li**	mitoni
hockey	REL–$be.there_{VII}$–3SG:IN.CONJ	boy	the:MASC:SG	very

kii–moshkinee–w	anima	la	meezoñ	eh
PST–$be.full_{VII}$–3SG:IN	DEM:IN:SG	the:FEM:SG	house	eh

'when hockey was on, boy the house was really full, eh'

(Harvey Pelletier; 2012–10–13)

(114) kii-oshiihtaawak uh, li, uh, li vaeñ

kii–oshiihtaa–wak	uh	**li**	uh	li	vaeñ
PST–make$_{VAIT}$–3PL:AN	uh	the:MASC:SG	uh	the:MASC:SG	wine

'they made, uh, uh wine'

(Gracy Zoldy; 2012–07–18)

In (114) above, there are two occurrences of the definite article *li* 'the'. In cases such as this, only the article appearing closest to the noun was retained; all others were excluded.

In addition, there were several cases in which it was not possible to definitively determine the gender of the noun based on the surrounding context. For example, in some cases, the effects of rapid speech, unclear pronunciation, environmental factors, speech overlap, etc., made it difficult to determine with complete certainty whether an article was the masculine *aeñ* vs. feminine *en* 'a/n', or masculine *li* vs. feminine *la* 'the'. Any token for which this was the case was excluded, leaving only phonetically unambiguous cases in the dataset. With further analysis (e.g., by instrumental phonetic means), it may be possible to reintegrate some of these tokens into the dataset, although this has been reserved as a task for future research.

Animacy and/or gender irretrievable

Any tokens for which solid grammatical evidence for both animacy and gender were not directly observable were further excluded from the dataset. While it would be possible to consider the observed animacy and gender values in isolation (i.e., all tokens in which only animacy is observable, or all tokens in which only gender is observable), this study considers issues at the intersection of these two systems, which largely requires consideration of instances where both values are observable in the same utterance. In addition, the possibility of variability in animacy and gender assignment, both in the

speech of individual Michif speakers and between speakers, cannot be ruled out at the outset, making it difficult to combine animacy values observed without gender in one token with the gender value observed in another. Instead, this study focuses exclusively on tokens for which both animacy and gender values were retrievable, which allows for attention to be given to any potential variability in animacy and gender assignment.

In some cases, even though the noun was accompanied by an article or possessive adjective, it was not possible to retrieve the gender. This was especially true for plural nouns appearing in the corpus, since the article *lii* 'the' is used for both masculine and feminine nouns, obscuring the gender distribution that exists for the singular definite articles:

(115) dahor kii-akoteew dañ lii pchit brañsh

dahor	kii–akotee–w	dañ	**lii**	pchit	brañsh
outside	PST–hang$_{VII}$–3SG:IN	on	the:PL	little	branch

'it was hanging outside on little branches'

(Verna DeMontigny; 2013–07–16)

(116) pii lii dañs kaa-ayaachik

pii	**lii**	dañs	kaa–ayaa–t–ik
and	the:PL	dance	REL–have$_{VAIT}$–3SG.CONJ–3PL:AN

'and they had dances'

(Norman Fleury; 2012–10–13)

(117) aakoshi ee-peeshiiwiichihkot anihi kahkiiyaw lii pom …
ee-mooshaahkinaachik

aakoshi	ee–pee–ishii–wiichih–ikw–t	anihi
so	CONJ–come–thus–help$_{VTA}$–3'>3–3SG.CONJ	DEM:AN:3'

kahkiiyaw	**lii**	pom	…
all	the:PL	apple	…

ee–mooshaahkin–aa–t–ik
CONJ–pick$_{VTA}$–3>3'.CONJ–3SG.CONJ–3PL:AN

'so they came and helped him, they picked up the apples'

(Rita Flamand; 2013–08–13)

We cannot determine whether the nouns in the above examples are masculine or feminine because they appear with the plural definite article *lii* 'the'.

Similarly, a number of utterances containing the French-origin possessive adjectives *not* 'our', *vot* 'your', and *loer* 'their' were returned in the searches. As these do not distinguish gender, however, it is not possible to determine the gender values of the nouns they accompany, so they were excluded from the dataset as well:

(118) oshaam chahkoshiw loer keulot

oshaam	chahkoshi–w	**loer**	keulot
too	be.short$_{VAI}$–3SG:AN	3PL.POSS	pants

'their pants are too short'

(Victoria Genaille; 2012–10–15)

(119) pii noohkwaatam not faas

pii	noohkwaat–am	**not**	faas
and	lick$_{VTI}$–3OBJ:IN	1PL.POSS:SG	face

'and he licks our face'

(Verna DeMontigny; 2015–06–04)

(120) vot parañtii neetee kii-wiikiwak

vot	parañtii	neetee	kii–wiiki–wak
2PL.POSS:SG	relative	over.there	PST–live$_{VAI}$–3PL:AN

'your relatives lived over there'

(Norman Fleury; 2012–10–13)

Similarly, instances of French-origin liaison (e.g., *lom* 'man') were also excluded because it was not possible to read either masculine or feminine gender from them:

(121) aeñ tramb niwaapamaaw aakota li bor di likrañ

aeñ	tramb	ni–waapam–aa–w	aakota
a:MASC:SG	tree	1–see$_{VTA}$–NON3>3–3SG:AN	there

li	bor	di	**likrañ**
the:MASC:SG	side	PTV	cliff

'I see a tree there, by the side of that cliff'

(Norman Fleury; 2013–08–24)

Finally, there were cases in which it was clear that the natural gender of a noun was masculine, but if this could not be confirmed by the grammar, the token was excluded:

(122) aeñ moshwee roozh kikishkaweew ana lom

aeñ	moshwee	roozh	kikishkaw–ee–w
a:MASC:SG	handkerchief	red	wear.AN$_{VTA}$–3>3'.IND–3SG:AN

ana	**lom**
DEM:AN:SG	man

'the man is wearing a red handkerchief'

(Norman Fleury; 2013–08–24)

Unlike French, gender assignment in Michif is not consistently retrievable through means such as demonstrative or adjective agreement. Thus, all cases in which it was not possible to determine the gender of a noun from the surrounding linguistic context were excluded from the dataset.

Finally, several Cree-origin nouns such as those in (123)–(126) were excluded from the dataset:

(123) kiishpin dañ li magazaeñ noohkom chi-itohteet

kiishpin	dañ	li	magazaeñ	**ni–ohkom**
if	to	the:MASC:SG	store	1POSS–grandmother

chi–itohtee–t
PURP–go$_{VAI}$–3SG.CONJ

'if my grandmother went to the store'

(Grace Zoldy; 2012–07–18)

(124) moo 'kaat giiwiikinaan Selby Town apree nimoshoom kaa-nipot

moo	wiihkaat	ni–kii–wiiki–naan	Selby	Town
NEG	ever	1–PST–live$_{VAI}$–1PL	Selby	Town

apree	**ni–moshoom**	kaa–nipi–t
after	1POSS–grandfather	REL–die$_{VAI}$–3SG.CONJ

'we never lived in Selby Town after my grandfather died'

(Lawrance Fleury; 2013–07–24)

(125) nipaapa gishkishin maana la poschin ee-oshiihaat mischet

ni–paapa	ni–kishkishi–n	maana	la
1POSS–father	1–remember$_{VAI}$–NON3:SG	used.to	the:FEM:SG

poschin	ee–oshiih–aa–t	mischet
pudding	CONJ–make$_{VTA}$–3>3’	a.lot

‘I remember my dad used to make a lot of pudding’

(Mervin Fleury; 2012–10–13)

(126) yeah, tuu lii swer maana kii-kashkikwaashow maana, nimaama, you know

yeah	tuu	lii	swer	maana	kii–kashkikwaasho–w
yeah	every	the:PL	night	used.to	PST–sew$_{VAI}$–3SG:AN

maana	**ni–maama**	you	know
used.to	1POSS–mother	you	know

‘yeah, my mum used to sew every night’

(Verna DeMontigny; 2012–10–15)

Since these nouns typically appear with Cree possessive affixes instead of the French-origin possessive adjectives discussed here, and without any definite or indefinite articles, it was impossible to deduce their gender in all but a handful of cases. In other cases, the Cree-origin noun appeared with a French-origin plural article:

(127) kishkishin lii takwaaminaana anihi?

ki–kishkishi–n	**lii**	**takwaaminaan–a**	anihi
2–remember$_{VAI}$–NON3:SG	the:PL	chokecherry–PL	DEM:IN:PL

‘do you remember those chokecherries?’

(Harvey Pelletier; 2012–10–13)

Because masculine/feminine gender is not discernible from the plural article *lii* ‘the’, instances such as these were also excluded from the dataset.

However, in a few rare cases, the Cree-origin noun appeared with either a possessive adjective or an article of French origin:

(128) lii zañfañ, ayi, toñ ooshishima, mishikitiwak eektiween la

lii	zañfañ	ayi	**toñ**	**ooshishim–a**
the:PL	child	HES	2SG.POSS:MASC:SG	grandchild–PL

mishikiti–wak	eektiween	la
be.big$_{VAI}$–3PL:AN	I.guess	INTERJ

'the children, uh, your grandchildren, they must be big now, I guess'

(Mervin Fleury; 2013–07–24)

In (128), the Cree-origin noun *ooshishima* 'grandchildren' appears with the French-origin possessive adjective *toñ* 'your', indicating that the noun is masculine. We know that it is animate because it is the subject of the Animate Intransitive verb *mishikitiwak* 'they are big', which can only have animate subjects. Since both animacy and gender values were retrievable for this noun, it was one of the few Cree-origin nouns which were retained in this phase of the development of the dataset.

Animacy irretrievable

As for gender, any noun token for which there was no clear grammatical indicator of animacy was excluded from the dataset. In some cases, this occurred because the nouns appeared in short phrases without any verbs or demonstratives from which the animacy of the noun could be retrieved:

(129) soñ pchii shyaeñ aakota

soñ	pchii	**shyaeñ**	aakota
3SG.POSS:MASC:SG	little	dog	there

'his little dog there'

(Rita Flamand; 2013–08–13)

In example (129), there is no verb or demonstrative present to provide an indication of the animacy of the noun *shyaeñ* 'dog'.

Many other cases of exclusion due to irretrievable animacy values occurred when the noun was part of a prepositional phrase:

(130) li gaa, kii-pee-niihtakoshiiw oschi dañ li tramb

li	gaa	kii–pee–niihtakoshii–w		
the:MASC:SG	guy	PST–come–climb.down$_{VAI}$–3SG:AN		
oschi	dañ	li	**tramb**	
from	from	the:MASC:SG	tree	

'the guy came down from the tree'

(Cecile Burroughs; 2013–07–30)

(131) dañ li karoo oota

dañ	li	**karoo**	oota
in	the:MASC:SG	square	here

'in the square here'

(Mary Fleury; 2013–09–13)

(132) añbaa la log en pchit gournoy aakota

añbaa	la	log	en	pchit	**gournoy**	aakota
below	the:FEM:SG	log	a:FEM:SG	little	frog	there

'down below the log, a little frog is there'

(Norman Fleury; 2013–08–24)

In (130)–(132), the nouns *tramb* 'tree', *karoo* 'square', and *gournoy* 'frog' all appear within prepositional phrases. There are no discernible animacy values for any of these nouns, either through verbal inflection, or through agreement with demonstratives.

Referent not retrievable

There were other cases in which both animacy and gender were retrievable, but it was not possible to determine from discourse what the actual referent was. These bear some similarity to the epicene nouns discussed below, but differ in that their semantic range is much broader (e.g., *la mem* 'the same one' could potentially refer to any feminine noun in discourse and not be uniquely identifiable with a particular referent). These were excluded as well:

(133) la mem eekwaana kaa–kii–ashamit ana
la **mem** eekwaana kaa–kii–asham–in ana
the:FEM:SG same DEM:AN:SG REL–PST–feed.AN_{VTA}–? DEM:AN:SG
'the same that he came and fed me'
(Mervin Fleury; 2013–07–24)

(134) eekwa anima li meeyoer
eekwa anima li **meeyoer**
and DEM:IN:SG the:MASC:SG best
'and that's the best'
(Grace Zoldy; 2012–07–18)

(135) well, li premyii ana
well li **premyii** ana
well the:MASC:SG first DEM:AN:SG
'well, the first one'
(Victoria Genaille; 2012–10–15)

(136) ekwa miina la res ohiñ
ekwa miina la **res** ohiñ
and also the:FEM:SG rest DEM:IN:PL
'and again the rest of these'
(Victoria Genaille; 2012–12–02)

(137) ookik li res ka–kanaweeyimaawak

ookik	**li**	**res**	ka–kanaweeyim–aa–wak
DEM:AN:PL	the:MASC:SG	rest	FUT–look.after$_{VTA}$–NON3>3–3PL:AN

'you can look after **the rest** of them'

(Victoria Genaille; 2012–10–15)

There were also a handful of Michif tokens for which an English translation was not available (even after consultation with Michif speakers). In some cases (e.g., (138)), the translating speaker had a general impression of the meaning, but a precise word-for-word translation could not be arrived at:

(138) soñ traeñ chi-otinamahk

soñ	**traeñ**	chi–otin–am–hk
3SG.POSS:MASC:SG	?	PURP–take$_{VTI}$–3OBJ:IN–INDEF.ACTOR

'to take somebody's belongings'

(Grace Zoldy; 2012–07–18)

These tokens were also excluded.

No equivalent source language form

Another major category for which items were excluded from the dataset involves any lemmas for which the source language information was not readily available. Since part of this investigation involves comparing the Michif animate/inanimate and masculine/feminine values to those of their source language equivalents, this is a crucial criterion. Any tokens for which a source language form could not be identified were excluded. This especially applied to proper nouns. Since there was not a clear source language equivalent form that we could associate with the occurrence in the corpus, all such forms were excluded from the dataset:

(139) la vyee Betsy kii–wiikiw anda

la	vyee	**Betsy**	kii–wiiki–w	anda
the:FEM:SG	old	Betsy	PST–live$_{VAI}$–3SG:AN	there

'Old lady Betsy, she lived there'

(Lawrance Fleury; 2013–07–24)

(140) ayi wiishta peeyek sa faam aeñ Boucher, ayi

ayi	wiishta	peeyek	sa	faam
HES	her.too	one	3SG.POSS:FEM:SG	wife

aeñ	**Boucher**	ayi
a:MASC:SG	Boucher	HES

'…uh, her too, one of the Boucher's wives, uh'

(Lawrance Fleury; 2013–07–24)

(141) oh, oh li vyoo Felix dañ larmii wiishta kii-ayaaw?

oh	oh	li	vyoo	**Felix**	dañ
oh	oh	the:MASC:SG	old	Felix	in

larmii	wiishta	kii–ayaa–w
the.army	him.too	PST–be.there$_{VAI}$–3SG:AN

'oh, oh, old Felix, he was in the army, too?'

(Mervin Fleury; 2013–07–24)

(142) la vyee Madeleine, pii…

la	vyee	**Madeleine**	pii
the:FEM:SG	old	Madeleine	and

'old lady Madeleine, and…'

(Harvey Pelletier; 2012–10–13)

In other cases, (e.g., *duuzen* 'dozen', *eñvii* 'craving', English-origin *break* 'break'), no Cree equivalent forms were noted in any of the dictionaries consulted in this study (see Section 6.2.2). These were also excluded from consideration, as equivalent Cree animacy values were needed for the analysis.

Lemmas with epicene natural gender

Finally, a handful of lemmas (e.g., KOZAEÑ ‘cousin’, MET ‘master (of an animal)’) had no single natural gender value. The natural gender of these lemmas could be either masculine or feminine, depending on the gender of their referent, and thus they could not be coded as having a single natural gender value. This poses difficulties for comparing observed Michif gender and natural gender in individual lemmas: while the observed gender values of some tokens may align with the natural gender of their referents, others may not, and these cases cannot easily be distinguished at the lemma level. Since one of the goals of this study is to consider instances of alignment or misalignment between observed and natural gender, these lemmas were excluded from the final dataset.

In each of the cases described in this section, there was not enough information to make them useable for the purposes of this study. In general, the aim of these criteria was to be conservative in determining which items should remain in the dataset and which should be eliminated, erring on the side of exclusion in cases where gender or animacy could not be determined with certainty. After all of these items were excluded, a total of 1,261 tokens remained, which had been drawn from over 22 hours of recordings across 51 recordings with 15 speakers. As the following sections discuss in more detail, these tokens were then coded for additional features and lemmatized, producing a final set of 261 unique lemmas. As noted previously, further phonetic and/or morphological analysis may allow for some excluded tokens to be reintroduced to the dataset, although the number of such cases is likely not substantial. In general, the resulting sample of nouns is sufficiently extensive to serve as the basis of this study. A table presenting all the

lemmas, their frequencies, English glosses, and source languages is provided in Appendix 1.

6.2.2 Coding the dataset

After the unsuitable data were excluded, all remaining tokens were coded for the following:

1. Observed Michif gender (masculine/feminine)
2. Observed Michif animacy (animate/inanimate)
3. Source language of token (i.e., whether the form of this noun was derived from French, Cree, or English)
4. The equivalent of this noun in contemporary Plains Cree
5. The equivalent of this noun in contemporary Canadian French
6. Animacy of the Cree source/equivalent form (animate/inanimate)
7. Gender of the French source/equivalent form (masculine/feminine)
8. Natural gender (masculine/feminine/neuter)
9. Natural animacy (animate/inanimate)

Table 30 below illustrates the coding for a noun of French, Cree, and English origin:

Table 30: Example of coding for tokens in the dataset

#	Lemma	Eng. Gloss	Michif gender	Michif animacy	Source lang.	Cree equiv.	Cree anim.	French syn.	French gender	Natural gender	Natural animacy
3360	BILLY_GOAT	'billy goat'	m	a	ENGLISH	*wâpatihk*	a	*couc*	m	m	a
561	KOK	'rooster'	m	a	FRENCH	*nâpê-pâhkahahkwân*	a	*coq*	m	m	a
1403	POOSHIISH	'cat'	m	a	CREE	*pôsîs-* (Eng. borrowing)	a	*minou*	m	n	a

Plains Cree sources consulted to determine the equivalent Cree forms and their animacy values were primarily drawn from Wolvengrey (2001) and the Online Cree Dictionary (Wolvengrey 2015). Canadian French sources consulted to determine equivalent French forms and their gender values were primarily Robinson & Smith (1990) and wordreference.com (2018b).

In some cases, the Michif noun appeared to be derived from a French noun with a related but distinct sense. For example, while the Michif noun for 'squirrel' is *swis*, the French cognate *suisse* means 'chipmunk', while 'squirrel' is *écureuil*. In this case, it is clear that the Michif noun was derived from *suisse*, although the meaning may have shifted or been applied in a different way in the course of Michif's development. In the handful of cases such as these, I treated the closest phonetic cognate as the French synonym and coded accordingly, as these are the forms that are more likely to provide relevant information about source language gender.

Finally, as mentioned previously, several instances of English-origin words appeared in the corpus. In these cases, I consulted the Plains Cree and Canadian French sources mentioned above to determine the equivalent forms in both source languages, as

well as their animacy and gender. It was not always possible, however, to determine what the equivalent would be in one or more source languages (e.g., *backscratcher, break, Coke, dozen, gangrene, gear, guitar, hotel room, hydro (electricity), interview, Norwegian, relief (as in government assistance), circle, sandwich, square, stuff, tape, yellow (one),* etc.). These cases had no Cree and/or French equivalent forms, and so could not be coded for source language animacy and gender. They were, however, still coded for observed Michif animacy and gender, as well as natural animacy and gender wherever possible.

6.3 Classification of lemmas

The procedures described above yielded a large number of tokens of many different Michif nouns, along with their values for animacy and gender, which were then grouped together under their lemmas. For the purposes of this study, I follow Frawley *et al.* (2002) in assuming that a lemma represents an abstraction over a set of word forms that share both morphosyntactic and semantic attributes (Frawley, Hill & Munro 2002: 3), and, secondarily, some degree of phonological similarity. The animacy and gender values for each lemma in this study are derived from the corresponding tokens in the dataset. For instance, 11 tokens of *frer* 'brother' in the dataset are all observed to be animate and masculine. These tokens serve as the basis for treating the lemma FRER as being animate and masculine as well. In this way, observations of token-level animacy and gender assignment to nouns in the corpus inform the type-level categorizations of the corresponding lemmas.

Defining a lemma as having occurrences that all share the same morphosyntactic properties and range of meanings has direct implications for the treatment of certain

Michif lexical items encountered in the corpus. For example, in two cases in the dataset, nouns referring to animals, namely PUUL ‘chicken’ and ZWEE ‘goose’, appeared with inanimate marking. This is unexpected, as animals generally occur as animates in both Michif and its Algonquian source languages. However, closer inspection of the relevant concordance lines reveals that in these instances, these inanimate instances refer not to the living animals, but rather to their meat, as in the example for ‘chicken’ reproduced below:

(143) keekwee eekoshi ishi kii-miichichik, la puul

keekwee	eekoshi	ishi	kii–miichi–t–ik
something	thus	this.way	PST–eat$_{VAIT}$–3SG.CONJ–3PL:AN

la	puul
the:FEM:SG	chicken

‘something like that they ate, chicken’

(Victoria Genaille; 2012–12–02)

In (143), the form of the VAIt verb stem *miichi-* ‘eat s.t.’ indicates that the object of the verb is inanimate. The verb is also inflected by the affixes *-t* and *-ik*, indicating that the actor of the verb is animate. *La puul* ‘chicken’ is therefore treated as being grammatically inanimate because of the form of the verb with which it agrees. In both of these cases, this divergence in animacy values appears to be a result of two different meanings being associated with this phonological form, one of which refers to the animal, and one of which refers to its meat. These are therefore treated as separate, homophonous lemmas in the dataset. This is further supported by the observation that these are also separate lexemes in Cree (e.g., *pâhkahâhkwân*$_{ANIM}$ ‘chicken’ vs. *pâhkahâhkwâniwiyâs*$_{INAN}$ ‘chicken (meat)’), although it is worth noting that the Michif forms do not co-occur with *vyañd*

'meat' or any other reflex of Cree *wiyâs* 'meat', which differs from the Cree pattern. All nouns in the dataset which share the same phonological form but have different meanings were treated as belonging to distinct lemmas. This occurred for a total of seven forms in the dataset, as indicated in Table 31:

Table 31: Homophonous lemmas in the dataset

Lemma	**Animacy**	**Gender**
BWAA-1 'stick'	i	m
BWAA-2 'wood'	i	m
LIIÑG-1 'line'	i	m
LIIÑG-2 'fishing line'	i	f
PUUL-1 'chicken'	a	f
PUUL-2 'chicken meat'	i	f
SHAR-1 'car'	a	m
SHAR-2 'train'	i	m
VYEE-1 'old lady'	a	f
VYEE-2 'wife'	a	f
VYEU-1 'old man'	a	m
VYEU-2 'husband'	a	m
ZWEE-1 'goose'	a	m
ZWEE-2 'goose meat'	i	m

Note that even though some of these items share the same animacy and gender values (e.g., *bwaa* 'stick; wood' is inanimate and masculine for both meanings), the different senses are enough to warrant their treatment as distinct lemmas.

6.3.1 Variable classifications

In most cases, all of the tokens associated with a particular lemma shared the same animacy and gender values, and determining the animacy and gender of that lemma was thus straightforward. In a small number of instances, however, variability was observed in animacy and gender values among the different tokens for a given lemma. For example, of the five times that *moshwee* 'handkerchief' appears in the dataset, it is animate three times and inanimate twice:

(144) aeñ moshwee roozh kishkaweew, ana lom

aeñ	moshwee	roozh	kikishkaw–ee–w
a:MASC:SG	handkerchief	red	$wear_{VTA}$–3>3'.IND–3SG:AN

ana	lom
DEM:AN:SG	man

'the man is wearing a red $handkerchief_{AN}$'

(Norman Fleury; 2013–08–24)

(145) soñ moshwee keeshchinaa anima aeñ moshwee, si pans

soñ	moshwee	keeshchinaa	anima
3SG.POSS:MASC:SG	handkerchief	definitely	DEM:IN:SG

aeñ	moshwee	si	pans
a:MASC:SG	handkerchief	1SG	think

'That is definitely his handkerchief, a $handkerchief_{IN}$, I think'

(Harriet St. Pierre; 2013–08–19)

In (144), *moshwee* 'handkerchief' agrees with a Transitive Animate verb and is thus classified as animate. In (145), however, it appears with the inanimate demonstrative

anima, and is thus inanimate. Note that each of these instances is provided by a different speaker. This is thus an example of interspeaker variability in animacy assignment.

Individual Michif speakers are also found to show variability in their animacy decisions. For example, the lemma PAEÑ ‘bread’ appears as both animate and inanimate in the dataset, as shown in (146) and (147):

(146) li paeñ kii-paashowaat ohiñ li swer ooma

li	paeñ	kii–paashw–aa–t	ohiñ
the:MASC:SG	bread	PST–dry_{VTA}–3>3’.CONJ–3SG.CONJ	DEM:AN:OBV

li	swer	ooma
the:MASC:SG	evening	DEM:IN:SG

‘he dried the $bread_{AN}$ this evening’

(Victoria Genaille; 2012–10–15)

(147) moñ paeñ doshtamaashon, ma galet, kahkiiyaaw

moñ	paeñ	ni–oshihtamaasho–n
1SG.POSS:MASC:SG	bread	1–$make.for.myself_{VAIT}$–NON3:SG

ma	galet	kahkiiyaaw
1SG.POSS:FEM:SG	bannock	everything

‘I make my own bread, my $bannock_{IN}$, everything’

(Victoria Genaille; 2012–12–02)

In the above examples, we see that on one occasion, the speaker treats *paeñ* ‘bread’ as animate, while on another occasion, the same speaker treats *paeñ* ‘bread’ as inanimate. In (146), *paeñ* ‘bread’ is treated as animate because it occurs with animate inflection in the corresponding verb. Likewise, in (147), it is used with a VAIt, which only takes

inanimate objects. This is therefore evidence of intraspeaker variability in animacy assignment.

This variability in animacy and gender assignment has a direct impact not only on lemmatization, but also potentially on the overall results of this study, as it complicated the decision as to how to code a lemma for observed animacy and/or gender. In some cases, even though there was variability represented among the tokens, there were enough tokens which shared a majority value that it seemed logical to assume that the variable tokens were likely nonce speech errors, rather than evidence of true variability. In other cases, however, there was not enough evidence to make any such determination. This was especially true when the overall token numbers for a given lemma were very low. In the hopes of identifying additional tokens which might reveal a larger tendency, in these cases I performed a search among the tokens which had been excluded from the final dataset for reasons discussed in Section 6.2.1, but which still had discernible animacy or gender values. These additional tokens, which are included in the token counts in Table 32–Table 35 and Table 33–Table 36, served as evidence informing the decision as to how to code these lemmas.

To resolve these ambiguities, the following criteria for assigning values to the lemmas were then applied:

> *Criterion #1:* If variability between animate/inanimate and/or masculine/feminine values was noted for a given lemma and fewer than four tokens were available in the dataset, then the lemma was excluded from further consideration in this study. In such cases, the token counts were so low that it was not possible to conclusively detect any central trend or preference for one value over another

(e.g., a lemma with only two tokens, one animate and one inanimate). This eliminated five lemmas (four on the basis of variability in animacy, one on the basis of gender).

Criterion #2: If the token count for a given lemma was higher than four and the exceptional tokens made up 20% or less of the total number of tokens, the lemma was assigned the majority animacy or gender value.[37] In other words, if the lemma had a very small number of exceptions among a larger number of consistent forms, it seemed reasonable to treat these exceptions as potential speech errors or idiolectal variation, rather than offering compelling evidence of pervasive variability in the language. For example, the noun *fìi* 'girl' appears 45 times in the dataset, twice as masculine (4.4%) and 43 times as feminine (95.6%). The corresponding lemma FII 'girl' was therefore coded as feminine. This occurred in 14 cases (8 for animacy, 6 for gender).

Criterion #3: Any other instances of variability which did not fit into either of the above two categories were treated as having variable animacy and/or gender values. In these cases, there is evidence in the corpus of lemmas showing substantial variability affecting more than 20% of the tokens, which cannot be reasonably attributed to nonce speech errors or idiolectal variation. This affected 11 lemmas (2 for animacy, 9 for gender).

[37] Although the cut-off of 20% is arbitrary, it nevertheless serves as a useful heuristic in the case of lemmas with low token frequencies, which make up the majority of the dataset. A lemma with five tokens, four of which are assigned consistently to one animacy or gender and one of which to the other (i.e., 4 vs. 1), will be assigned to the majority animacy or gender value, while a more even balance in gender or animacy values (e.g., 3 vs. 2) will be treated as variable.

Overall, there were fifteen Michif lemmas in the dataset which showed variability in animacy values. These variable nouns are presented in the following table, along with their animate and inanimate token frequencies in the dataset:

Table 32: Michif lemmas with variable animacy tokens

Lemma	**Animate frequency**	**Inanimate frequency**	**Coding decision**	
			Value	**Consistency**
BARYER ‘barrier; gate’	1/4	3/4	v	25.0%
BASKET ‘basket’	1/12	11/12	i	91.7%
BICYCLE ‘bicycle’	12/14	2/14	a	85.7%
BITAEÑ ‘clothes’	1/18	17/18	i	94.4%
DARYER ‘behind; rear-end’	1/3	2/3	exclude (<3 tokens)	
FEU ‘fire’	2/13	11/13	i	84.6%
KLOSH ‘clock’	7/8	1/8	a	87.5%
MOSHWEE ‘handkerchief’	3/5	2/5	v	40.0%
MUNICIPALITY ‘municipality’	1/2	1/2	exclude (<3 tokens)	
PAEÑ ‘bread’	1/2	1/2	exclude (<3 tokens)	
PAYIIÑ ‘basket’	1/29	28/29	i	96.6%
ROSH ‘rock’	14/15	1/15	a	93.3%
TABLIIYII ‘apron’	1/2	1/2	exclude (<3 tokens)	
TRAMB ‘tree’	11/12	1/12	a	91.7%

Variability involving animacy is thus evident in the dataset, both in the speech of individual Michif speakers and between speakers. In most cases, however, this variability appears to be restricted to a few, low-frequency exceptions among a more robustly attested animacy value, or occurs in cases where there are too few tokens to discern between actual variability and nonce speech errors. In only two instances were lemmas

treated as being variable in their animacy. This represents a significant difference from variability in gender.

Variability involving masculine/feminine gender in Michif has not been widely reported in the literature, but a close inspection of the corpus provides evidence of both inter- and intraspeaker variability in masculine/feminine gender values for particular lemmas in the dataset. Examples (148)–(149) show variability within the speech of a single speaker:

(148) soñ gournoy kaa-kii-otinaat anihi li pchi garsoñ

soñ	gournoy	kaa–kii–otinaa–t	anihi
3SG.POSS:MASC:SG	frog	REL–PST–take$_{VTA}$–3SG.CONJ	DEM:3’

li	pchi	garsoñ
the:MASC:SG	little	boy

‘his frog$_{MASC}$ that he took, that little boy’

(Rita Flamand; 2013–08–13)

(149) ah, ‘chiko ee-teepwaatat anihi sa gournoy

ah	nawachiko	ee–teepwaat–aa–t
ah	kind.of	CONJ–yell$_{VTA}$–3>3’.CONJ–3SG.CONJ

anihi	sa	gournoy
DEM:3’	3SG.POSS:FEM:SG	frog

‘ah, he’s still kind of yelling at his frog$_{FEM}$’

(Rita Flamand; 2013–08–13)

In the examples above, the same speaker uses both masculine (148) and feminine (149) possessive markers for *gournoy* ‘frog’. This is therefore an example of intraspeaker variability that does not appear to be shared by all other Michif speakers.

While the preceding example provides evidence of intraspeaker variability across utterances in the same session, in some cases different gender markers may even be used in the same utterance. This is the case for *fii* 'girl' below, where it appears first with the masculine indefinite article *aeñ*, followed by the feminine definite article *la*:

(150) Nakishkaaweew **aeñ** pchit fii. Wiishta aeñ bicycle, awa **la** pchit fii.

nakishkaw–ee–w	**aeñ**	pchit	fii	wiishta
meet$_{VTA}$–3>3'.IND–3SG:AN	a:MASC:SG	little	girl	her.too

aeñ	bicycle	awa	**la**	pchit	fii
a:MASC:SG	bicycle	DEM:AN:SG	the:FEM:SG	little	girl

'He met a little girl$_{MASC}$. Her too, she had a bicycle, the little girl$_{FEM}$'

(Harriet St. Pierre; 2013–08–19)

In the following example, the speaker begins with the masculine definite article *li* 'the' and masculine form of the adjective *primyee* 'first', before correcting himself and using the feminine form of the adjective *primyer* 'first':

(151) **li primyee, primyer** parson en otomobil ee-kii-ayaawaat anda

li	**primyee**	**primyer**	parson
the:MASC:SG	first:MASC	first:FEM	person

en	otomobil	ee–kii–ayaa–t	anda
a:FEM:SG	car	CONJ–PST–have$_{VTI}$–3SG.CONJ	there

'the first$_{MASC}$, first$_{FEM}$ person that had a car there'

(Lawrance Fleury; 2013–07–24)

The sensitivity to gender assignment in cases such as these suggests that intraspeaker variation in gender assignment is not necessarily haphazard.

Finally, in (152), we see the French-origin *pwel* 'stove' being used with feminine article, and in the same utterance the speaker uses the masculine indefinite article *aeñ* for English-origin *cookstove* 'cookstove':

(152) ma tañt ana peeyak **en** gros pwel kii-ayaaweew, **aeñ** cookstove, eh

ma	tañt	ana	peeyak	**en**	gros	pwel
1SG.POSS:FEM:SG	aunt	DEM:AN:SG	one	a:FEM:SG	big	stove

kii–ayaaw–ee–w	**aeñ**	cookstove	eh
PST–have$_{VTA}$–3>3'.IND–3SG:AN	a:MASC:SG	cookstove	eh

'one of my aunts, she had a big stove$_{FEM}$, a cookstove$_{MASC}$, eh'

(Grace Zoldy; 2012–07–18)

Here, what could reasonably be assumed to be the same referent receives two different gender assignments, depending on the phonological form being used to refer to it. Thus, gender cannot always be assumed to be an inherent property of the referent itself.

Variability in gender is attested not only in the case of individual Michif speakers, but also between speakers. Examples (153)–(154) provide further examples of variability for gender noted between Michif speakers in the dataset:

(153) **la** nik anima kiipahkihtin

la	nik	anima	kii–pahkihtin–w
the:FEM:SG	beehive	DEM:IN:SG	PST–fall$_{VII}$–3SG:IN

'that beehive$_{FEM}$ fell down'

(Norman Fleury; 2013–08–24)

(154) a ter anima **li** nik

a	ter	anima	**li**	nik
on	ground	DEM:IN:SG	the:MASC:SG	beehive

'that beehive$_{MASC}$ [is] on the ground'

(Rita Flamand; 2013–08–13)

In (153) *nik* 'beehive' is used with the feminine definite article *la* 'the' by one speaker, while in (154), it is used with the masculine definite article *li* by another speaker. Thus, interspeaker variability extends not only to cases such as the one summarized in Table 32, where certain speakers demonstrate variability in gender assignment and others do not; but also to cases where individual speakers differ from one another consistently, with some favoring one gender value for a particular noun and others another.

Sixteen Michif lemmas were found to show variability for observed gender in the dataset:

Table 33: Overall variability for observed Michif gender, all lemmas

Lemma	**French Equivalent**	**Fem. Freq.**	**Masc. Freq.**	**Coding decision**	
				Value	**Consistency**
BARYER 'barrier; gate'	*barrière*$_{\text{FEM}}$	10/14	4/14	v	71.4%
FARO 'forest'	*forêt*$_{\text{FEM}}$	1/2	1/2	exclude (<3 tokens)	
FII 'girl'	*fille*$_{\text{FEM}}$	43/45	2/45	f	95.6%
GOURNOY 'frog'	*grenouille*$_{\text{FEM}}$	16/38	22/38	v	57.9%
KAAB 'rope'	*câble*$_{\text{MASC}}$	1/10	9/10	m	90.0%
KASKET 'cap'	*casquette*$_{\text{FEM}}$	5/7	2/7	v	71.4%
KEU 'tail'	*queue*$_{\text{FEM}}$	31/37	6/37	v	83.7%
KOK 'rooster'	*coq*$_{\text{MASC}}$	1/22	21/22	m	95.5%
LUUN 'moon'	*lune*$_{\text{FEM}}$	4/5	1/5	f	80.0%
MASHIN 'machine'	*machine*$_{\text{FEM}}$	4/6	2/6	v	66.7%
NIK 'beehive'	*nic / nique*$_{\text{MASC}}$	3/6	3/6	v	50.0%
PAAT 'leg'	*patte*$_{\text{FEM}}$	3/5	2/5	v	60.0%
PARSON 'person'	*personne*$_{\text{FEM}}$	3/4	1/4	v	75.0%
ROB 'dress'	*robe*$_{\text{FEM}}$	5/6	1/6	f	83.3%
SHEEZH 'chair'	*chaise*$_{\text{FEM}}$	7/8	1/8	f	87.5%
SWIS 'squirrel'	*suisse*$_{\text{MASC}}$ 'chipmunk'	3/6	3/6	v	50.0%

In contrast to the lemmas with variable animacy tokens, the majority of variable gender tokens were treated as instances of true variability, as the consistency rate of gender values among the tokens was less than 80% (Criterion #3). Several others showed a degree of consistency equal to or higher than 80% (Criterion #2), and thus were assigned a masculine or feminine value in accordance with the most frequent value. Only one instance was excluded due to low token frequency (Criterion #1).

A summary of the coding decisions for lemmas with both variable animacy and gender tokens is provided in the following table:

Table 34: Coding decisions for lemmas with variable animacy and gender tokens

	Animacy	**Gender**
Criterion #1: Excluded	4	1
Criterion #2: Assigned a value	8	6
Criterion #3: Coded as variable	2	9

In sum, the data used in this study are drawn from a corpus of Southern Michif consisting of narratives, multispeaker conversations, and semi-controlled linguistic tasks. Over 5,000 noun tokens were initially extracted from this corpus though the majority of those were excluded for a variety of factors. The remaining 1,261 noun tokens were lemmatized and coded for Michif animacy and gender, source language, translation equivalents in Cree and French and their corresponding animacy and gender values, and natural animacy and gender values. The final dataset, drawn from over 22 hours of recordings across 15 speakers, forms the basis for the quantitative investigation seen in the following chapter, in which animacy and gender are shown to be both independent and productive categories in Michif.

Chapter 7: Results and analysis

While preceding research has pointed to the existence of multiple systems of noun classification in Michif, the productivity of these systems and the relationship of these classifications to those found in Michif's primary source languages, as well as to the semantics of individual referents, has remained an open question. Through quantitative investigation of animacy and gender assignment patterns in our corpus, this study finds that Michif has two independent and productive grammatical categories of noun classification, each inherited from a different source language. These results show that there is no statistically significant association between animacy and gender in Michif. A strong association between Michif and Cree animacy is also found. In addition, a clear difference in gender assignment patterns is found between French- and English-origin lemmas, with French-origin lemmas largely aligning with their French equivalents and English-origin lemmas receiving primarily masculine gender values.

Among the 1,261 tokens that have recoverable gender, animacy, and source language information in the corpus, 261 unique lemmas are attested (see Appendix 1). The animacy and gender values for these lemmas are summarized in the following table:

Table 35: Observed animacy by observed gender

	Feminine	**Masculine**	**Variable**
Animate	32/98 (32.7%)	63/98 (64.3%)	3/98 (3.1%)
Inanimate	57/161 (35.4%)	99/161 (61.5%)	5/161 (3.1%)
Variable	0/2 (0.0%)	1/2 (50.0%)	1/2 (50.0%)

While masculine inanimate nouns are more frequent than the values of the other categories, all possible animacy-gender combinations are robustly attested, suggesting that the values of the animacy and gender categories are independent of one another. A Pearson's Chi-squared test was performed to assess whether or not there is any statistical association between animacy and gender. Lemmas with variable animacy and gender were removed before carrying out this test, since their cells in the above contingency table each contain counts of fewer than five instances, which is less than what is required to accurately estimate chi-squared values (cf. Gries 2009: 190).[38] No evidence of a statistically significant association was found (χ^2 (1) = 0.10397, p = 0.7471), suggesting that the two gender systems inherited from Michif's source languages have remained distinct and function independently of one another. This provides motivation for considering both sex-based and animacy-based classifications separately in the sections that follow.

[38] The results of a Fisher's exact test performed on the same data with variable lemmas retained show the same result (p = 0.1756).

Since potential differences in animacy and gender patterns by source language are explored in several of the sections below, it is relevant to first consider how the different source languages are represented among the lemmas in this dataset. The distribution of Michif nouns in this sample according to their source language is given in Table 36:

Table 36: Distribution of Michif lemmas by source language

Source language	**# Lemmas**
Cree	1/261 (0.4%)
English	45/261 (17.2%)
French	215/261 (82.4%)

These numbers confirm quantitatively what has been claimed in the literature on Michif largely based on impressionistic observation—that is, the majority of the nouns in Michif are of French origin (82.4% in this corpus; see Bakker 1997: 117 for an approximate proporation of source languages across Michif grammatical categories, based on the results of a translation task across multiple speakers). Less common, though also attested in this sample, are lemmas derived from English (e.g., *ambulance, ball, graveyard, school bus, triangle, washing machine*) and Cree (*pooshiish* 'cat', borrowed from English "pussy"; Cree Online Dictionary). Note that there are only four instances of this single Cree-origin noun that were retained for the final analysis, as others that were retrieved in searches did not occur in constructional contexts which allowed for their masculine/feminine gender values to be determined. Given this relatively low number, attention in the following sections will be given solely to the French and English-origin lemmas that are predominant in this dataset.

In the remainder of this chapter, I will investigate animacy and gender assignment patterns in Michif, comparing them to the relative animacy and gender values of their source languages. Section 7.1 finds a strong association between Michif and Cree animacy, demonstrating that the animacy of Michif nouns generally patterns in the same way as their Cree equivalents. Cases of misalignment between Michif and Cree are also explored, as well as the relationship between Michif animacy and natural animacy. Section 7.2 examines gender assignment patterns in both French- and English-origin lemmas, finding a significant difference between the two. The gender values of French-origin lemmas generally align with those of their French equivalents, while the majority of English-origin lemmas receive masculine gender values in Michif. Cases of misalignment between Michif and French gender values are also considered, with arbitrary French feminine nouns (i.e., nouns with no natural gender whose French equivalents are classified as being feminine) found to be more likely to regularize for both French- and English-origin nouns in Michif. This section also notes several exceptions to these patterns, such as the default French masculines discussed in Chapter Five receiving arbitrary feminine values in Michif. The relationship between Michif gender and natural gender is also explored in this section, while Section 7.3 summarizes these findings.

7.1 Animacy

In this section, we examine what determines the animacy of a Michif noun. Based on observations made in the preceding chapters, we focus our attention on investigating three hypotheses concerning animacy assignment:

- *Hypothesis 1: Michif animacy is determined by natural animacy.* According to this hypothesis, the animacy of a Michif noun follows from the natural animacy of its referent (i.e., non-living referents will be assigned inanimate grammatical animacy values, while living referents will be assigned animate grammatical animacy values).
- *Hypothesis 2: Michif animacy is determined by source language.* According to this hypothesis, Michif nouns follow different animacy assignment patterns based on their language of origin (e.g., French-origin nouns show evidence of different animacy assignment patterns compared to English-origin nouns).
- *Hypothesis 3: Michif animacy is determined by Cree translation equivalent.* On this view, Michif animacy is considered to be derived primarily by inheritance, with close correlations between Michif animacy values and those of their Cree translation equivalents being seen primarily as the result of Michif retaining the animacy values of Cree.

Each of these hypotheses is considered in the sections that follow.

7.1.1 Hypothesis 1: Michif animacy is determined by natural animacy

Our first hypothesis is that Michif animacy assignment simply follows natural animacy. While Michif animacy does appear to show a correlation with Cree animacy, as discussed in 7.1.3, another factor might show an even closer correlation. One reasonable candidate for this is the natural animacy of the referent, since this has more of a semantic basis than potentially arbitrary animacy values based on Cree. Table 37 presents the distribution of natural animacy and observed Michif animacy for all lemmas in the dataset. Of the 261

lemmas, observed Michif animacy and natural animacy are in alignment 84.7% of the time (221/261 lemmas). Exceptions make up roughly 14.6% of cases (38/261 lemmas):

Table 37: Natural animacy by observed Michif animacy

		Natural animacy	
		Animate	**Inanimate**
Michif animacy	**Animate**	60/98 (61.2%)	38/98 (38.8%)
	Inanimate	0/161 (0.0%)	161/161 (100.0%)
	Variable	0/2 (0.0%)	2/2 (100.0%)

A Fisher test performed on this contingency table reveals that Michif animacy and natural animacy are significantly associated ($p < 0.001$). Post hoc comparisons of these values reveal that only the association between non-variable Michif lemmas and natural animacy is statistically significant ($p < 0.001$).

In most of those cases, Michif agrees with Cree animacy rather than natural animacy, as seen above. There are 38 instances in which Michif animacy and natural animacy values were not in consistent alignment among the 261 Michif lemmas. Of these, 78.9% (30/38) follow Cree animacy rather than natural animacy, and are identical to the instances of Michif animacy alignment with Cree rather than with natural animacy provided in Table 42. Other instances of misalignment between Michif and Cree may indirectly be following the Cree pattern through a process of semantic analogy, as in Table 46. The remaining exceptions, where the animacy values of the Michif nouns deviate both from those of the Cree source language and from natural animacy, are listed in the following table:

Table 38: Mismatches for natural and Cree animates

Lemma	**Cree Equivalent**
COOKSTOVE$_{ANIM}$ 'cookstove'	*kotawânâpisk*$_{INAN}$
PWEL$_{ANIM}$ 'stove'	*kotawânâpisk*$_{INAN}$
SAVOÑ$_{ANIM}$ 'soap'	*kisîpêkinikan*$_{INAN}$
TREATY_CARD$_{ANIM}$ 'treaty card'	*iskonikanîwasinahikan*$_{INAN}$

These do not appear to form a natural semantic class, and further investigation is required to determine a possible cause for these instances of misalignment.

While Michif animacy does show a correlation with natural animacy, a similar correlation is also found between Cree animacy and natural animacy, as shown in Table 39:

Table 39: Natural animacy by Cree animacy

		Natural animacy	
		Animate	**Inanimate**
Cree animacy	**Animate**	60/100 (60.0%)	40/100 (40.0%)
	Inanimate	0/161 (0.0%)	161/161 (100.0%)

A Chi-squared test performed on this table also shows a statistically significant dependency between these two variables (χ^2 (1) = 122.07, p<0.001). Importantly, as Table 37–Table 39 make clear, neither Cree nor Michif correlate with natural animacy in all cases: in particular, arbitrary animates cannot easily be reduced to predictions of natural animacy. As discussed in 7.1.3, in the majority of these exceptional cases, Michif and Cree animacy are in alignment, suggesting that Michif animacy is more closely determined by the animacy values inherited from its Cree source language than from any

ongoing correlation with natural animacy. Thus, natural animacy is less predictive of observed Michif animacy values than the animacy of Cree translation equivalents is, disconfirming our hypothesis that Michif animacy is determined by natural animacy.

7.1.2 Hypothesis 2: Michif animacy is determined by source language

Quantitative analysis shows that there is no difference in the way that animacy is assigned to French-origin lemmas as compared to English-origin lemmas in Michif. That is, animacy is assigned in a similar fashion to nouns of any source language, thus demonstrating Hypothesis 2 to be false. Table 40 below summarizes the distribution of animacy values across English and French-origin lemmas in the dataset:

Table 40: Distribution of animacy values in English-origin and French-origin Michif lemmas

	Michif animacy		
Source language	**Animate**	**Inanimate**	**Variable**
English	17/45 (37.8%)	28/45 (62.2%)	0/45 (0%)
French	81/216 (37.5%)	133/216 (61.6%)	2/216 (0.9%)

The proportions in this table show that there are only slight differences in the distribution of animate and inanimate values by source language. For example, 37.8% (17/45) of English-origin nouns are assigned animate values, which is only slightly more than the 37.5% (81/216) of French-origin nouns that receive the same classification. A Fisher's exact test performed on the contingency table above confirms that these values are proportional to one another, and that there is no statistically significant association

between Michif animacy values and the source language of individual lemmas ($p = 1$).[39] Thus, language of origin does not influence or predict the animacy value of a given noun in Michif and the hypothesis that Michif animacy is determined by source language can be rejected. Since animacy assignment produces indistinguishable outcomes for English-origin nouns as compared to French-origin nouns, there is no need to analyze them separately in the discussion that follows.

7.1.3 Hypothesis 3: Michif animacy is determined by Cree translation equivalent

Our third hypothesis is that Michif animacy is determined by the Cree translation equivalent. This study finds that the animacy patterns of Michif lemmas are in alignment with their Cree equivalents in the vast majority of cases (92.7%, or 242/261 lemmas). Instances of misalignment between Michif and Cree animacy values are much rarer, constituting only 6.5% (17/261) of the lemmas in the dataset, while only 0.8% (2/261) of lemmas showed variability in their animacy values. The correspondences between observed Michif animacy and the animacy of equivalent Cree nouns in the dataset are presented in Table 41 below:

[39] In this study, a Fisher's exact test was performed instead of a Chi-squared test where the counts in cells in a contingency table were less than five, as this produces more precise p-values for tables with smaller counts (Baayen 2008: 113).

Table 41: Comparison of Michif and Cree animacy values

		Cree animacy	
		Animate	**Inanimate**
Michif animacy	**Animate**	90/98 (91.8%)	8/98 (8.2%)
	Inanimate	9/161 (5.6%)	152/161 (94.4%)
	Variable	1/2 (50.0%)	1/2 (50.0%)

A Fisher's exact test of independence reveals a statistically significant association between Michif and Cree animacy ($p < 0.001$), suggesting that these variables are in close alignment with one another. Post hoc comparisons of Michif and Cree animacy values further reveal that only the association between Cree and Michif animate and inanimate values is statistically significant (i.e., the lemmas with variable animacy values cannot be shown to have a significant association with either Cree animate or inanimate values).[40] Indeed, a closer inspection of these items reveals that Michif not only follows Cree animacy patterns in cases that align with natural animacy, but also largely maintains the "classic" Cree animacy exceptions discussed in Chapter Five, in which a semantically inanimate noun is treated as being grammatically animate. This occurred across 30 lexemes in the dataset, as shown in the following table:

40 These post hoc tests involved pairwise comparisons using Fisher's exact tests with the Bonferroni correction applied, and were carried out using the fisher.multcomp function provided by the RVAideMemoire library, version 0.9-69-3 in R 3.5.1 (Hervé 2018).

Table 42: Alignment between Michif and Cree animacy values rather than natural animacy

Lemma	**Cree Equivalent**
4X4$_{ANIM}$ '4 X 4'	*âwatâswâkan*$_{ANIM}$
ARAÑZH$_{ANIM}$ 'orange'	*osâwâs*$_{ANIM}$
BALL$_{ANIM}$ 'ball'	*pâkahatowân*$_{ANIM}$
BWATOÑ$_{ANIM}$ 'button'	*sakwâskwahon*$_{ANIM}$
CAKE$_{ANIM}$ 'cake'	*wîhkihkasikan*$_{ANIM}$
FARIIN$_{ANIM}$ 'flour'	*pahkwêsikan*$_{ANIM}$
GALET$_{ANIM}$ 'bannock'	*pahkwêsikan*$_{ANIM}$
GLAS$_{ANIM}$ 'ice'	*maskwamiy*$_{ANIM}$
JEANS$_{ANIM}$ 'jeans'	*iskwêwitâs*$_{ANIM}$
KATAEÑ$_{ANIM}$ 'doll'	*awâsisîhkân*$_{ANIM}$
KLOSH$_{ANIM}$ 'bell'	*sêwêyâkan*$_{ANIM}$
KOP$_{ANIM}$ 'penny'	*pîwâpiskos*$_{ANIM}$
KUULOT$_{ANIM}$ 'pants'	*-tâs*$_{ANIM}$
L/N/ARZHAÑ$_{ANIM}$ 'money'	*sôniyâw*$_{ANIM}$
L/N/ZITWEL$_{ANIM}$ 'star'	*atâhk*$_{ANIM}$
L/N/ZOTOMOBIL$_{ANIM}$ 'car'	*sêhkêw*$_{ANIM}$
LUUN$_{ANIM}$ 'moon'	*tipiskâwi-pîsim*$_{ANIM}$
NIIZH$_{ANIM}$ 'snow'	*kôna*$_{ANIM}$
POM$_{ANIM}$ 'apple'	*picikwâs*$_{ANIM}$
ROSH$_{ANIM}$ 'rock'	*asiniy*$_{ANIM}$
RUUBARB 'rhubarb'	*pikwânâhtik*$_{ANIM}$
SCHOOLBUS$_{ANIM}$ 'schoolbus'	*âwatawâsiswâkan*$_{ANIM}$
SHAR-1$_{ANIM}$ 'car'	*sêhkêw*$_{ANIM}$
SHORT$_{ANIM}$ 'shorts'	*kîskicâsis*$_{ANIM}$
SIMAÑ$_{ANIM}$ 'cement'	*asinîwipayihcikan*$_{ANIM}$
SLACKS$_{ANIM}$ 'slacks'	*nitâs*$_{ANIM}$
TABAA$_{ANIM}$ 'tobacco'	*ciscêmâs*$_{ANIM}$

Lemma	**Cree Equivalent**
TRAMB$_{ANIM}$ ‘tree’	*mistik*$_{ANIM}$
TRUCK$_{ANIM}$ ‘truck’	*âwatâswâkan*$_{ANIM}$
WART$_{ANIM}$ ‘wart’	*micîhcîkom*$_{ANIM}$

Here, we see that Michif animacy values tend to align themselves with Cree equivalent forms, even in cases where this is in opposition to natural animacy: all of the items in the above table consistently appear as animate in Michif, as they do in Cree, despite the fact that they refer to inanimate objects. This alignment is largely in line with observations made by other scholars and confirms our hypothesis that Michif animacy is determined by Cree translation equivalent.

Even with this predominant alignment between Michif and Cree nouns for animacy, however, several exceptions are found. The following table shows the 17 instances in which the animacy values of Michif forms consistently differ from those of their Cree equivalents:

Table 43: Misalignment between observed Michif animacy and animacy of Cree equivalent

Lemma	**Cree Equivalent**	**Mismatch Type**
AMBULANCE$_{ANIM}$ ‘ambulance’	*ahkosiwtapan*$_{INAN}$	2
BAA$_{INAN}$ ‘sock’	*asikan*$_{ANIM}$	1
BICYCLE$_{ANIM}$ ‘bicycle’	*nîsokâcis*$_{INAN}$	2
BIKE$_{ANIM}$ ‘bike, bicycle’	*nîsokâcis*$_{INAN}$	2
COOKSTOVE$_{ANIM}$ ‘cookstove’	*kotawânâpisk*$_{INAN}$	3
KOSH$_{INAN}$ ‘diaper’	*âsiyân*$_{ANIM}$	1
KRIIYOÑ$_{INAN}$ ‘pencil’	*masinahikanâhcikos*$_{ANIM}$	1
PEN$_{INAN}$ ‘pen’	*masinahikanâpiskos*$_{ANIM}$	1
PLEM$_{INAN}$ ‘pen’	*masinahikanâpiskos*$_{ANIM}$	1
PLOT$_{INAN}$ ‘ball’	*pâkahatowân*$_{ANIM}$	1
PWEL$_{ANIM}$ ‘stove’	*kotawânâpisk*$_{INAN}$	3
SAVOÑ$_{ANIM}$ ‘soap’	*kisîpêkinikan*$_{INAN}$	3
SEUK$_{INAN}$ ‘sugar’	*sôkâw*$_{ANIM}$	1
SHAR-2$_{INAN}$ ‘train’	*iskotêwitâpân*$_{ANIM}$	1
TREATY_CARD$_{ANIM}$ ‘treaty card’	*iskonikanîwasinahikan*$_{INAN}$	3
UNDERWEAR$_{INAN}$ ‘underwear’	*atâmayiwinis*$_{ANIM}$	1
WASHING_MACHINE$_{ANIM}$ ‘washing machine’	*kisîpêkinikâkan*$_{INAN}$	3

These mismatches can be grouped into three different categories. Type 1 mismatches appear to be cases were Michif animacy follows natural animacy rather than Cree animacy (nine lemmas). For example, writing utensils (KRIIYOÑ ‘pencil’, PEN ‘pen’, and PLEM ‘pen’) appear as inanimate in the Michif dataset, even though their Cree equivalent *masinahikanâpiskos* is grammatically animate. Type 2 mismatches appear to be the result of semantic analogy from Cree (three lemmas). For example,

AMBULANCE ‘ambulance’ is treated as animate in the dataset even though the Cree equivalent *ahkosiwtapan* is inanimate. This is likely an extension of the classification of many means of transportation in Cree as animate. Type 3 covers mismatches where the motivation is not immediately apparent (five lemmas). This is summarized in Table 44:

Table 44: Types of mismatches between Michif and Cree animacy

Mismatch Type	**Number**
Type 1: Mismatches due to natural animacy	9
Type 2: Mismatches due to semantic analogy from Cree	3
Type 3: Miscellaneous	5
Total:	17

The majority of cases in Table 43 represent mismatches due to natural animacy (Type 1). These are what Gillon and Rosen (2018: 99) refer to as “arbitrary animates” in the source language—that is, the grammatical animacy of the Cree equivalent is animate, even though it refers to an inanimate object. For lemmas such as BAA_{INAN} ‘sock’, $\text{KOSH}_{\text{INAN}}$ ‘diaper’, $\text{KRIIYOÑ}_{\text{INAN}}$ ‘pencil’, $\text{PLEM}_{\text{INAN}}$ ‘pen’, $\text{PLOT}_{\text{INAN}}$ ‘ball’, $\text{SEUK}_{\text{INAN}}$ ‘sugar’), Michif appears to be following natural animacy rather than the animacy of the corresponding Cree terms. The distribution of cases where Michif and Cree animacy differ is shown in Table 45:

Table 45: Alignment between Michif and Cree animacy values across Cree arbitrary animate nouns vs. all other nouns

	Is Cree equivalent an arbitrary animate?	
	No	**Yes**
Cree animacy ≠ Michif animacy	8/17 (47.1%)	9/17 (52.9%)
Cree animacy = Michif animacy	212/242 (87.6%)	30/242 (12.4%)

A Fisher's exact test on these proportions indicates that Cree arbitrary animates are indeed more likely to show misalignment for animacy than other kinds of nouns in Michif (i.e., default inanimate, natural animate, natural inanimate), appearing as inanimate when their Cree equivalents are animate ($p<0.001$). While it is worth emphasizing that, in the majority of cases, Cree arbitrary animates still do align with Michif, Cree arbitrary animates are statistically more likely to regularize in Michif than other kinds of nouns.

For the Type 2 mismatches in Table 43, it appears that Michif may be applying a process of semantic analogy by following the same pattern as Cree in treating many items of clothing as animate (e.g., JEANS$_{\text{ANIM}}$ 'jeans', SLACKS$_{\text{ANIM}}$ 'slacks'). Similarly, many means of transportation are animate (e.g., 4X4$_{\text{ANIM}}$ '4 X 4', SCHOOLBUS$_{\text{ANIM}}$ 'schoolbus'), as they generally are in Cree (e.g., *sêhkêw*$_{\text{ANIM}}$ 'car'; *iskotêwitâpân*$_{\text{ANIM}}$ 'train'; *mistatimotâpânâskw*$_{\text{ANIM}}$ 'sled, sleigh, wagon, horse-drawn vehicle'; *âwatâswâkan*$_{\text{ANIM}}$ 'truck'). Michif appears to extend this generalization beyond what is noted in Cree, treating most vehicles as animate even when the equivalent Cree term is inanimate:

Table 46: Instances of Michif semantic analogy from Cree involving modes of transportation

Lemma	**Cree Equivalent**	**Anim. Freq.**	**Inan. Freq.**
AMBULANCE$_{ANIM}$ 'ambulance'	*ahkosiwtapan*$_{INAN}$	2/2	0/2
BICYCLE$_{ANIM}$ 'bicycle'	*nîsokâcis*$_{INAN}$	4/6	2/6
BIKE$_{ANIM}$ 'bike, bicycle'	*nîsokâcis*$_{INAN}$	1/1	0/1

Thus, even in certain cases of misalignment between Michif and Cree animacy values, Michif appears to be extending Cree animacy patterns beyond what is found in the Cree source language.

The possible motivations for the Type 3 misalignments noted in Table 43 are less apparent. In these cases, Michif assigns an animate animacy value, even though both the equivalent Cree form and the natural animacy are inanimate (e.g., Michif SAVOÑ$_{ANIM}$ vs. Cree *kisîpêkinikan*$_{INAN}$ 'soap'). One possible motivation for the assignment of an animate value to PWEL$_{ANIM}$ 'stove' may be linked to the fact that the Cree equivalent *kotawânâpisk* 'stove, oven' is listed in Wolvengrey (2001) as being inanimate, with a secondary note that animate occurrences are also attested. However, no information is provided that would distinguish between these two uses in Cree. Michif may therefore be following Cree in treating both COOKSTOVE and PWEL as animate. Finally, it is unclear why SAVOÑ$_{ANIM}$ 'soap' would be treated as being grammatically animate, since it is an inanimate object and does not appear to be associated with any related terms which would be animate. It should be noted, however, that only one token of this lemma appeared in the dataset, making it difficult to establish on the basis of corpus data alone whether this is merely a nonce speech error or speaks to a larger pattern. Even when exceptions such as these are taken into consideration,

however, Michif animacy values remain in alignment with Cree animacy values in the vast majority of cases.

7.1.4 Summary of findings concerning animacy

In this section, we have seen that there is a non-significant association between Michif animacy values for English-origin lemmas as compared to French-origin lemmas. Thus, on the whole, animacy assignment in Michif patterns in the same way, regardless of whether or not the noun was inherited from French or English. This disconfirms our second hypothesis that Michif animacy is determined by source language. In addition, Michif animacy values show a predominant alignment with the animacy values of their corresponding Cree terms (92.7%, or 242/261 lemmas), and show a statistically significant association for animacy with their Cree equivalents. This is true even in cases where the grammatical and natural animacy values are in opposition (e.g., $\text{ROSH}_{\text{ANIM}}$ 'rock'). Thus, Cree animacy is a better predictor of Michif animacy than is natural animacy, confirming that Hypothesis 1 is less explanatory than Hypothesis 3. One possible explanation for the close alignment between Michif and Cree animacy values is that the retention of animacy classification is linked to the high functional load that animacy bears in Michif, as it does in Algonquian languages. Almost all forms of predication in Michif involve either demonstratives or verbal inflection, both of which indicate the animacy of the referent. Even in cases of misalignment, Cree animacy still influences many instances of Michif animacy assignment through a process of semantic analogy, with English-origin lexical items such as AMBULANCE 'ambulance' and BIKE 'bike' now appearing to be included among classes of animacy exceptions where they are not found in Cree. In general, Cree arbitrary animates are more likely to show

misalignment for animacy than other kinds of nouns in Michif. The combined sum of this evidence demonstrates that Cree animacy assignment patterns exert an exceptionally strong influence on Michif, even in cases where animacy values are not in complete alignment. We now turn to an examination of masculine/feminine gender in Michif.

7.2 Gender

We have now seen that Michif animacy values generally correspond closely to those of equivalent Cree nouns, even in cases where this is in opposition to natural animacy. In this section, we similarly consider the relationship between the masculine/feminine gender values associated with Michif nouns and their corresponding French forms, testing the following hypotheses:

- *Hypothesis 1: Michif gender is determined by natural gender.* On this view, the gender values of Michif nouns would be expected to correspond to the natural gender properties of their referents (i.e., naturally masculine nouns receive masculine gender in Michif, while naturally feminine nouns receive feminine gender in Michif).
- *Hypothesis 2: Michif gender is determined by French translation equivalent.* On this view, the gender of a given Michif noun is due primarily to the classification that it receives in French, with these values having been retained in Michif.
- *Hypothesis 3: Michif gender is determined by source language.* According to this hypothesis, nouns from different source languages may show evidence of different gender assignment patterns (e.g., French-origin nouns may show one distribution of gender values, while English-origin nouns show another).

Each of these hypotheses is considered in the sections that follow. In Section 7.2.1, we consider Michif's relationship to natural gender (Hypothesis 1). This is followed by a discussion of the relationship between Michif gender and gender of the French translation equivalent (Hypothesis 2) in Section 7.2.2, an investigation into whether Michif gender is determined by source language in Section 7.2.3, and finally a summary of our findings concerning Michif gender in Section 7.2.4.

7.2.1 Hypothesis 1: Michif gender is determined by natural gender

In the case of animacy, all nouns can be considered to be either naturally animate or inanimate. Thus, all nouns have an inherent quality which fits within the animacy system, even if this is not followed grammatically. In the case of gender, however, many nouns have no masculine or feminine natural gender. Thus, unlike animacy, natural gender cannot explain the majority of gender assignments in Michif, since most nouns are not either naturally masculine or feminine. Thus, the hypothesis that Michif gender is determined by natural gender cannot be viable.

Nevertheless, an examination of gender assignment patterns for those nouns with no natural gender potentially offers a clearer view on how Michif gender is assigned when assignment patterns are not influenced by the masculine or feminine attributes of their referents. The distribution of the genders assigned to French-origin

lemmas in the dataset with no natural gender suggests that masculine gender is favored over feminine gender in Michif:

Table 47: Gender assignment to French-origin nouns of no natural gender

	Feminine	**Masculine**	**Variable**
Michif lemmas	72/186 (38.7%)	105/186 (56.5%)	9/186 (4.8%)
French lemmas	88/186 (47.3%)	98/186 (52.7%)	0/186 (0.0%)

A Fisher's exact test shows that the differences in the proportions of gender values assigned to sexless nouns in each language are statistically significant ($p = 0.00209$). Moreover, Michif shows a different distribution of gender values from French in nouns with no natural gender. While the French distribution of masculine and feminine gender values to nouns with no natural gender in the dataset also demonstrates a slight preference for masculine gender over feminine, this imbalance is slightly less pronounced than in Michif. Thus, Michif and French differ in how they treat naturally sexless nouns, with Michif assigning more of these nouns masculine gender than French. As Gillon and Rosen (2018) note, one motivation for this preference for masculine gender assignment in sexless French-origin forms may stem from the fact that masculine is considered to be the default gender in French (84). This distribution is therefore potentially indicative of a weak preference for masculine gender assignment among sexless French-origin nouns. This tendency is confirmed by the other findings of this study as well, as discussed below.

7.2.2 Hypothesis 2: Michif gender is determined by French translation equivalent

In this section, we test the hypothesis that the gender of a Michif noun is determined by its French translation equivalent. With this hypothesis confirmed, we can then test the sub-hypothesis that instances of misalignment between Michif and French gender values will generally assign default masculine values, as Gillon and Rosen (2018) suggest (104). We first examine the gender patterns for the French-origin subset of the dataset, followed by the English-origin subset.

In the majority of cases in this dataset (88.4%, or 190/215), the gender of a Michif lemma is identical to that of its French equivalent. Thus, as was the case with animacy, gender in Michif nouns also appears to have largely the same distribution across lexical items as in their source language equivalents. The correspondences between observed Michif gender and the gender of equivalent French nouns in the dataset are presented in Table 48:

Table 48: French gender by observed Michif gender in French-origin lemmas

		French gender	
		Feminine	**Masculine**
Michif gender	**Feminine**	82/85 (96.5%)	3/85 (3.5%)
	Masculine	13/122 (10.7%)	108/122 (89.3%)
	Variable	7/9 (77.8%)	2/9 (22.2%)

Lemmas showing misalignment between Michif and French gender values in the above table constitute 7.4% (16/215) of the French-origin lemmas, while variable lemmas constitute 4.2% (9/215) of the French-origin dataset. A Fisher's exact test reveals a significant association between Michif gender and French gender ($p<0.001$), suggesting

that, for French-origin lemmas, this close alignment between Michif and French gender in the dataset cannot easily be attributed to chance.

Despite large-scale agreement between the gender values of Michif lemmas and those of their French equivalents, however, there are several lemmas that do not follow this pattern. Mismatches between the observed Michif gender and the gender of the French equivalent noun were found in 16 lemmas in the dataset. Among these cases, 12 represent instances of arbitrary French feminines receiving a default masculine value in Michif, as shown in Table 49:

Table 49: Misalignment in French-origin lemmas of arbitrary feminine gender

Lemma	**French equivalent**
BWASOÑ$_{\text{MASC}}$ 'alcohol'	*boisson*$_{\text{FEM}}$
CHIM_DI_ZHVOO$_{\text{MASC}}$ 'team of horses'	*time de chevaux*$_{\text{FEM}}$[41]
ISHEL$_{\text{MASC}}$ 'ladder'	*échelle*$_{\text{FEM}}$
KOÑFICHEUR$_{\text{MASC}}$ 'preserves'	*confiture*$_{\text{FEM}}$
KUUVECHOER$_{\text{MASC}}$ 'roof; cover'	*couverture*$_{\text{FEM}}$
LAVEUZ$_{\text{MASC}}$ 'tub; washer'	*laveuse*$_{\text{FEM}}$
LIIÑG-1$_{\text{MASC}}$ 'line'	*ligne*$_{\text{FEM}}$
LUUMYER$_{\text{MASC}}$ 'light'	*lumière*$_{\text{FEM}}$
TAAB$_{\text{MASC}}$ 'barn'	*étable*$_{\text{FEM}}$
WIL$_{\text{MASC}}$ 'oil'	*huile*$_{\text{FEM}}$
ZWEE-1$_{\text{MASC}}$ 'goose'	*oie*$_{\text{FEM}}$
ZWEE-2$_{\text{MASC}}$ 'goose meat'	*oie*$_{\text{FEM}}$

41 In a section on farm-related vocabulary, Robinson & Smith (1990: 51) list timefem, teamfem, and spannefem as informal Québec variants of the more formal attelage 'team'.

All of these instances represent cases where the French noun has feminine grammatical gender even though the referent has no natural gender. Consider the distribution of cases where Michif and French gender differ shown in Table 50, which distinguishes French arbitrary feminine nouns from all other nouns in the dataset for French-origin lemmas:

Table 50: Gender of French-origin arbitrary feminine nouns vs. all other nouns

	Other	**Arbtrary feminine**
Mismatched gender	6/25 (24.0%)	19/25 (76.0%)
Matched gender	121/190 (63.7%)	69/190 (36.3%)

A Chi-squared test on these proportions indicates that, for the French-origin lemmas in the dataset, French arbitrary feminines are indeed more likely to show misalignment for gender than other kinds of nouns in Michif (i.e., default masculines, natural feminines, natural masculines), appearing as masculine when their French equivalents are feminine (χ^2 (1) = 12.796, $p < 0.001$). Thus, for the French-origin lemmas, arbitrary French feminine nouns are more likely to receive default masculine gender values in Michif than other kinds of nouns. This supports our sub-hypothesis that, where Michif and French differ in gender assignment, Michif will generally assign default masculine values.

The remaining cases of misalignment do not appear to follow either natural gender or French gender:

Table 51: Misalignment between observed Michif gender and both natural and French gender, French-origin lemmas

Lemma	**French Equivalent**	**Fem. Freq.**	**Masc. Freq.**
GRAÑDRII$_{FEM}$ ‘granary’	*grenier*$_{MASC}$	2/2	0/2
MAAMA$_{MASC}$ ‘mom’	*maman*$_{FEM}$	0/2	2/2
PWEL$_{FEM}$ ‘stove’	*poêle*$_{MASC}$	1/1	0/1
RABABUU$_{FEM}$ ‘stew; rabbit soup’	*ragoût*$_{MASC}$	1/1	0/1

The Michif gender values here are in opposition to what one would expect based on both natural gender and the equivalent French gender values. In all but one of these instances, the Michif lemma is feminine despite the French source being masculine. The lemma MAAMA$_{MASC}$ ‘mom’ is particularly unusual, as this appears twice as masculine in the dataset, despite the referent being naturally feminine:

(155) li maama pii li paapa anikik
li maama pii li paapa anikik
the:MASC:SG mom and the:MASC:SG father DEM:AN:PL
‘the mom and the dad, those ones’
(Norman Fleury; 2013–08–24)

(156) uh, Johnny LeDoux, soñ maama anihi
uh Johnny LeDoux soñ maama anihi
uh Johnny LeDoux 3SG.POSS:MASC:SG mom DEM:AN:3’
‘uh, Johnny LeDoux, his mom, that one’
(Norman Fleury; 2012–10–13)

In (155), *maama* ‘mom’ appears with the masculine definite article *li* ‘the’ rather than the expected feminine article *la* ‘the’. Similarly, in (156), *maama* ‘mom’ appears with the masculine singular possessive adjective *soñ* ‘his/hers’, as opposed to the expected feminine possessive adjective *sa* ‘his/hers’.

Note that there are only 1–2 total tokens for each of the lemmas for which this unexpected gender assignment is found, leaving open the possibility that these may be nonce speech errors rather than a coherent pattern in gender assignment. However, as noted in Chapter Five, similar discrepancies between French and Michif gender values have been identified in the literature (see, e.g., Papen 2003a). While some of these reported instances of misalignment correspond to arbitrary feminine nouns in French becoming masculine in Michif (e.g., *beut*$_{\text{MASC}}$ 'butte, hill' from French *butte*$_{\text{FEM}}$, *grif*$_{\text{MASC}}$ 'claw' from French *griffe*$_{\text{FEM}}$, *mosh*$_{\text{MASC}}$ 'fly' from French *mouche*$_{\text{FEM}}$), others pattern similarly to those listed in Table 51, where a French masculine noun becomes feminine in Michif (e.g., *bwatoñ*$_{\text{FEM}}$ 'stick' from French *bâton*$_{\text{MASC}}$, *kwatoñ*$_{\text{FEM}}$ 'cotton' from French *coton*$_{\text{MASC}}$, *muulaeñ*$_{\text{FEM}}$ 'mill' from French *moulin*$_{\text{MASC}}$). These instances of unexpected divergence from the French equivalent identified by previous researchers may also be nonce speech errors, but this is difficult to assess since no frequency data or explicit information about how these results were reached were provided by those authors. Identifying additional tokens of these lemmas in discourse and/or consulting with native speakers may provide an indication as to whether or not these unexpected Michif gender values are more than coincidental. Although further investigation of such forms would no doubt be beneficial, the presence of French masculine nouns receiving feminine gender in Michif in the corpus and in the wider literature appear to challenge Gillon and Rosen's claim that "…arbitrary feminine gender assignment can be replaced by default masculine, but default masculine cannot be replaced by arbitrary feminine" (2018: 105). Despite these discrepancies, however, these results generally suggest that,

for lemmas of French origin, Michif gender is generally in line with the gender values of French translation equivalents.

Unlike the French-origin nouns discussed above, the gender of French translation equivalents does not appear to have any appreciable bearing on Michif gender assignment to English-origin lemmas. Of the 45 English-origin lemmas, Michif observed gender and French-equivalent gender values are in alignment 66.7% of the time (30/45 lemmas):

Table 52: French gender by observed Michif gender in English-origin lemmas

		French gender	
		Feminine	**Masculine**
Michif gender	**Feminine**	2/4 (50.0%)	2/4 (50.0%)
	Masculine	13/41 (31.7%)	28/41 (68.3%)

A Fisher's exact test applied to the above distribution does not find any significant association between Michif and French gender values ($p = 0.5913$) among English-origin lemmas in this dataset. There is thus no evidence in this corpus from which to conclude that French gender has an appreciable influence on Michif gender assignment in the case of English-origin lemmas.

While the preceding discussion has focused on the gender values associated with English-origin nouns found in the corpus, gender values for some English borrowings in Michif have also been reported in several of the sources surveyed in Chapter Five. Table 53, reproduced from Chapter Five, presents additional examples of misalignment between Michif and French gender:

Table 53: Reported misalignment of gender values of English borrowings between Michif and French

Michif Borrowing	**French Equivalent**	**Source**
aeñ baynder$_{\text{MASC}}$ 'binder'	*une moissoneuse*$_{\text{FEM}}$	(Bakker 1997: 105)
la dam$_{\text{FEM}}$ 'dam'	*le barrage*$_{\text{MASC}}$	(Bakker 1997: 105; Papen 2003a: 131)
la fun$_{\text{FEM}}$ 'fun'	*le plaisir*$_{\text{MASC}}$*; le fun*$_{\text{MASC}}$	(Bakker 1997: 105)
li kar$_{\text{MASC}}$ 'car'	*l'automobile*$_{\text{FEM}}$	(Hogmen 1981: 86)
la kuushin$_{\text{FEM}}$ 'cushion'	*le cousin*$_{\text{MASC}}$	(Papen 1987a: 250)
aeñ paañsyoñ$_{\text{MASC}}$ 'pension'	*une pension*$_{\text{FEM}}$	(Papen 2003a: 131)
la slee$_{\text{FEM}}$ 'sleigh'	*le traîneau*$_{\text{MASC}}$	(Papen 2003a: 131)
li staf$_{\text{MASC}}$ 'stuff'	*la matière*$_{\text{FEM}}$	(Papen 1987a: 250)
aeñ suutkees$_{\text{MASC}}$ 'suitcase'	*une valise*$_{\text{FEM}}$	(Papen 2003a: 131)

Five of these nine cases of misalignment have French equivalent forms with an arbitrary feminine value which is reassigned to a masculine gender value in Michif. This follows the same pattern as seen in Table 49. However, other cases are less straightforward in comparison to the patterns seen above. These include *la dam*$_{\text{FEM}}$ 'dam', *la fun*$_{\text{FEM}}$ 'fun', *la kuushin*$_{\text{FEM}}$ 'cushion', and *la slee*$_{\text{FEM}}$ 'sleigh', all of which are masculine in French, but receive feminine gender in Michif. These examples parallel instances such as *bwatoñ*$_{\text{FEM}}$ 'stick', *kwatoñ*$_{\text{FEM}}$ 'cotton', and *muulaeñ*$_{\text{FEM}}$ 'mill' seen above in that both involve what would be masculine nouns in French appearing with feminine gender in English, suggesting that this pattern is not restricted to French-origin nouns.

Of the English-origin lemmas in the dataset, 91.1% (41/45) appeared with a masculine gender marker, while a mere 8.9% (4/45) were used with a feminine gender marker. Although this general preference for masculine gender assignment to English-origin nouns is noteworthy, the fact that some nouns are used with feminine gender suggests that not all nouns borrowed from English may necessarily default to masculine gender. The instances in which English-origin lemmas are coded as feminine appear in the following table:

Table 54: English-origin lemmas with feminine gender

Lemma	**French Equivalent**
$FARM_{FEM}$ 'farm'	*ferme*$_{FEM}$
FUN_{FEM} 'fun'	*fonne*$_{MASC}$ / *fun*$_{MASC}$
$GAME_{FEM}$ 'game'	*game*$_{FEM}$ / *jeu*$_{MASC}$
$SALOP_{FEM}$ 'slop'	*bouette*$_{FEM}$

As evident from this table, in some cases the Michif observed gender value was identical to that of the French equivalent ($FARM_{FEM}$ 'farm', $SALOP_{FEM}$ 'slop'), while in other cases, the observed Michif gender and French gender did not align (FUN_{FEM} 'fun', $GAME_{FEM}$ 'game'). In the case of $FARM_{FEM}$ 'farm' and $SALOP_{FEM}$ 'slop', it is likely that speakers are simply mapping their knowledge of the gender of the equivalent French form onto the English-origin form, and are thus applying French gender assignment rules to the English-origin forms. A possible explanation for the gender of $GAME_{FEM}$ 'game' is that the gender corresponds to Canadian French *la game* (f.) (an English borrowing into Canadian French with feminine gender, which typically refers to hockey), rather than to *le jeu* (m.). This may therefore be another case of alignment with the French source

language.[42] Thus, FUN_{FEM} 'fun' is the only English-origin form in the dataset which has feminine gender in Michif but not in French.

With only four lemmas attested in the dataset and low token counts for each, the status of English-origin nouns with feminine gender values in Michif such as these might be nonce speech errors. It is worth noting, however, that other examples of English-origin nouns with feminine gender have arisen in the course of elicitation and other language-focused tasks as well. Several of these are reproduced in (157) below:

(157) English-origin words with feminine gender:

la gum$_{FEM}$ 'chewing gum' (**li*) (cf. French *la gomme*$_{FEM}$)[43]
la kuushin$_{FEM}$ 'cushion' (**li*) (cf. French *le coussin*$_{MASC}$)
la light$_{FEM}$ 'light' (**li*) (cf. French *la lumière*$_{FEM}$)

(Verna DeMontigny, p.c.)

Importantly, according to Michif speakers, these forms are consistently feminine, and it is ungrammatical to use masculine gender markers with them. For *la gum* 'chewing gum' and *la light* 'light', it is likely that the feminine gender value of the original French form is being mapped onto the English label. Similarly, Bakker (1997) reports *la meel* 'mail' as being feminine in Michif, which is in accordance with French *la poste* 'mail', and

42 It is an open question whether the source form for GAME 'game' should be French-origin jeu, English-origin game, or French-origin game (where game represents an English borrowing that has been integrated into Canadian French). This usage of game in Canadian French has a limited semantic range, typically only being used to refer to hockey games. The Michif lemma GAME in the dataset here is not used in this restricted sense, so there is reason to treat this lemma as being of English origin, although it is not possible to know for certain. Further etymological work would be needed on the history of the Canadian French borrowing game to determine whether or not it would have entered the language before, during, or after the formative period when Michif was in heavy contact with French to draw any definitive conclusions. Barring this, we provisionally treat this as an English borrowing in this study.

43 As *gum* and *gomme* are cognate between English and French, this form in Michif may be of either French or English origin.

would seem to follow this same pattern (105). *La kuushin* 'cushion' falls into the same class of exceptions seen for both French and English-origin nouns above (*dam* 'dam', *fun* 'fun', *slee* 'sleigh'; *bwatoñ* 'stick', *kwatoñ* 'cotton', *muulaeñ* 'mill'), where default French masculines receive feminine gender in Michif.

In Chapter Six, it was noted that many nouns in the corpus were removed from the final dataset because animacy values were not retrievable for them. However, many of these forms do have retrievable gender values, and may offer an additional source of information about English-origin nouns which are assigned feminine gender in Michif, especially since the sample reproduced in Table 54 is so small. These additional forms are provided in the following table, along with their French equivalent forms:

Table 55: Additional English-origin forms with feminine gender in the corpus

Lemma	**French Equivalent**	**FEM. Freq.**
BANK$_{FEM}$ 'bank'	*banque*$_{FEM}$	1
BLUFF$_{FEM}$ 'bluff (group of trees)'	*bosquet*$_{MASC}$	1
BUNCH$_{FEM}$ 'bunch'	*groupe*$_{MASC}$ | *bande*$_{FEM}$	1
FAMILY$_{FEM}$ 'family'	*famille*$_{FEM}$	1
GLASS$_{FEM}$ 'glass'	*verre*$_{MASC}$ | *tasse*$_{FEM}$	1
LOG$_{FEM}$ 'log'	*rondin*$_{MASC}$ | *bûche*$_{FEM}$	6
PEAR$_{FEM}$ 'pear'	*poire*$_{FEM}$	2
POCKET$_{FEM}$ 'pocket'	*poche*$_{FEM}$	1
RELIEF$_{FEM}$ 'relief' (government assistance)	*aide*$_{FEM}$ | *assistance*$_{FEM}$	1
SHED$_{FEM}$ 'shed'	*abri de jardin*$_{MASC}$ | *cabane*$_{FEM}$	1
SKATING_RINK$_{FEM}$ 'skating rink'	*patinoire*$_{FEM}$	1
SLEIGH$_{FEM}$ 'sleigh'	*bacagnole*$_{FEM}$	1
TOILET$_{FEM}$ 'toilet'	*toilette*$_{FEM}$	1
WAR$_{FEM}$ 'war'	*guerre*$_{FEM}$	2

In all but one of these cases (i.e., BLUFF 'bluff (group of trees)'), the corresponding French term is feminine. Thus, while the majority of nouns borrowed from English into Michif are masculine, this clearly does not apply to all nouns. Although most English nouns that receive feminine gender in Michif typically have a feminine French equivalent, there were two instances in the corpus that challenge this generalization, *lot*$_{FEM}$ 'lot' and *ride*$_{FEM}$ 'ride':

(158) yeah, koñbaeñ lii meezoñ kii-shashkahamwak, en lot?

yeah	koñbaeñ	lii	meezoñ		
yeah	how.many	the:PL	house		
	kii–shashka–h–am–wak			en	lot
	PST–ignite$_{VTI}$–CAUS–3OBJ:IN–3PL:AN			the:FEM:SG	lot

'yeah, how many houses did they set fire to, a lot?'

(Lawrance Fleury; 2013–07–24)

(159) "kiiyaapit folee ma ride tipahamaan," ihtweew

kiiyaapit	folee	ma	ride	
still	must	1SG.POSS:FEM:SG	ride	
	tipah–am–aan			ihtwee–w
	pay$_{VTI}$–3OBJ:IN–1SG.CONJ			say$_{VAI}$–3SG:AN

'"I still have to pay for my ride," he said'

(Lawrance Fleury; 2013–07–24)

Examples (158)–(159) show two cases of nonce borrowings from English which do not correspond to nouns in French. Rather, the noun LOT$_{FEM}$ 'lot' in *en lot* 'a lot' corresponds to the adverb *beaucoup* 'a lot' in French, while the noun RIDE$_{FEM}$ 'ride' in '(pay for) my ride' corresponds to a verb in French (e.g., *conduire qqn. quelque part*). It is unusual that these forms would receive feminine gender, when masculine gender appears to be the default (Gillon & Rosen 2018: 99). As these forms have no clear-cut corresponding nominal equivalents in French, instances such as these provide evidence that not all gender values in Michif are necessarily inherited from the source language, and that other factors may influence the assignment of these gender values.

In sum, the gender of English borrowings in Michif attested in the dataset is overwhelmingly masculine (91.1%), although a smaller number of feminine forms are

also attested. The majority of English forms that are feminine in Michif have French equivalent forms with feminine gender, suggesting a possible process of double relexification (cf. Bakker 1989b), whereby the gender of the French form is mapped onto the English form that replaces it. There are, however, a few exceptions to this pattern, where the French source form is masculine and yet receives feminine gender in Michif (e.g., FUN$_{FEM}$ 'fun' from French *fonne*$_{MASC}$ / *fun*$_{MASC,}$ DAM$_{FEM}$ 'dam' from French *le barrage*$_{MASC}$, etc.). Some of these may be speech errors, but the fact that others, such as FUN$_{FEM}$, are both attested in my data and also reported elsewhere (cf. Bakker 1997: 105) suggests that this pattern may be more consistent than the low frequencies associated with many of these forms in the corpus might otherwise lead us to believe. There are also the two somewhat more problematic cases of nonce borrowings from English which receive feminine rather than masculine gender (LOT$_{FEM}$ 'lot', RIDE$_{FEM}$ 'ride'). In these examples, the borrowed English nouns in these expressions have neither received default masculine gender nor inherited their feminine value from French (since they do not have a nominal French equivalent). As with the English-origin Michif nouns noted above that show consistent feminine gender values despite having masculine equivalents in French, these instances suggest that processes other than inheritance are at work in the assignment of gender to words from English in Michif.

Despite these discrepancies, the fact remains that the overall tendency is for English nouns in Michif to receive masculine gender. Gillon and Rosen (2018) claim that this shift of arbitrary feminines towards default masculine gender is indicative of an incipient weakening of sex-based gender (i.e., masculine/feminine noun categorization) in Michif, which would be in line with the predictions of their model of the Michif

determiner phrase (104). Although the authors are not explicit on this point, this framing appears to assume French gender values as the basis of comparison, with Michif gender values losing arbitrary classifications that French maintains. However, it is difficult to see how regularization such as this can be interpreted as undermining the overall viability of the masculine/feminine gender system in Michif. As this study has shown, all Michif nouns continue to receive one gender value or the other and, with the exception of a handful of lemmas with variable gender, these categorizations are not in free variation, even if they may differ from their French equivalents. If anything, the fact that only natural feminines retain feminine gender values could be seen as strengthening the semantic basis of sex-based gender classification. Bringing grammatical gender values into closer alignment with natural gender only serves to make this part of the language more like grammatical animacy, where the alignment of Michif nouns with natural animacy values is much more pronounced.

Even if one assumes that sex-based gender is in the process of "weakening" in Michif, it remains unclear why some nouns appear to have retained the feminine gender of their French equivalents, or why others would be assigned feminine gender even though the French counterpart is either masculine or non-existent. It is possible that some of the English nouns that display feminine gender values are of particularly high frequency in Michif discourse, although this would be difficult to confirm on the basis of a relatively small corpus such as the one used in this study. Overall, these results show that, unlike French-origin lemmas, French gender does not determine Michif gender for English-origin lemmas.

7.2.3 Hypothesis 3: Michif gender is determined by source language

Our third hypothesis is that gender is determined by source language—that is, that gender is assigned differently to French-origin lemmas as compared to English-origin lemmas. Table 56 summarizes the distribution of Michif gender values across English and French-origin lemmas.

Table 56: Distribution of gender values in English-origin and French-origin Michif lemmas

	Michif gender		
Source language	**Feminine**	**Masculine**	**Variable**
English	4/45 (8.9%)	41/45 (91.1%)	0/45 (0%)
French	85/215 (39.5%)	121/215 (56.3%)	9/215 (4.2%)

As opposed to animacy, where the distribution of animate and inanimate values was relatively even across lemmas representing different source languages, the above table makes clear that gender does not follow such a pattern. Performing a Fisher's exact test on this contingency table confirms that gender values are not distributed in the same way across English-origin and French-origin lemmas ($p<0.001$). Post hoc tests show that only the association between masculine and feminine gender values and source languages is statistically significant. That is, no association was found involving variable gender lemmas. In addition, a binomial test for French-origin lemmas found that the observed proportion of masculine and feminine values (0.41, 85/206) was not evenly distributed (0.5; $p = 0.01455$). This is no different from French, which has also been reported to have a skew towards masculine values in its lexicon (Ayoun 2018: 115, citing Séguin 1969). In the dataset, 44.8% of French translation equivalents are feminine, while 55.2% are

masculine, which reflects this same trend. An additional binomial test indicates that gender assignment in French-origin Michif lemmas (0.41, 85/206) is consistent with the distribution of masculine and feminine values found in the French translation equivalents ($p = 0.3269$). Similarly, a binomial test found that the observed proportion of masculine and feminine values for English-origin Michif lemmas (0.089, 4/45) was not consistent with a random, evenly-proportioned distribution of masculine and feminine gender values ($p<0.001$). Thus, the apparent preference for masculine gender assignment in English-origin lemmas cannot be attributed to chance. Crucially, these results imply that (1) French-origin lemmas appear to have the same overall distribution of gender values as in Canadian French, suggesting a close relationship between gender assignment and source language in this subset of the lexicon; and (2) gender assignment behaves differently in English- and French-origin lemmas, with English-origin forms overwhelmingly favoring masculine values. These results suggest that gender is assigned differently to French- and English-origin nouns in Michif, confirming our hypothesis that Michif gender is determined by source language (Hypothesis 3). There are several possible reasons why English-origin nouns in Michif tend to be masculine. For one, English borrowings in French also tend to default to masculine values. Another possibility is that phonological factors may play a role in gender assignment in Michif, similar to what Tucker, Lambert, and Rigault (1977) have observed for French. Though beyond the scope of this study, an investigation into possible phonological predictors of gender assignment in Michif may reveal that English-origin nouns are more likely to receive masculine gender than French-origin nouns simply due to their phonological composition. In addition, masculine gender is the typologically unmarked value, and is also slightly more common in the French

lexicon than feminine gender. Michif's apparent preference for masculine gender is therefore not necessarily exceptional in light of these other potential factors. We will return to this idea in Section 7.3 below.

7.2.4 Summary of findings concerning gender

In this section, we have seen that our hypothesis that Michif gender is determined by natural gender (Hypothesis 1) is not tenable because we find masculine/feminine gender values being assigned to nouns which do not have inherent natural masculine or feminine properties. In fact, this is the case for the majority of lemmas in the dataset. In addition, a significant difference was found between Michif and French regarding gender assignment of nouns with no natural gender in that Michif assigns more of these nouns masculine gender values than does French. We have also seen that the observed Michif gender values of French-origin nouns largely agree with the gender values of their source language equivalents (88.4%, or 191/216), confirming our second hypothesis that Michif gender is determined by French translation equivalent. Finally, we found a statistically significant difference between gender assignment patterns in French-origin lemmas compared to English-origin lemmas, confirming our third hypothesis that Michif gender is affected by source language.

On the surface, it might not seem surprising that French gender would have been maintained in Michif along with the French phonological forms associated with the majority of Michif nouns, but it is worth noting that this is counter to what is typically found in other contact situations. Other contact languages that have inherited their lexicon from French (e.g., Haitian Creole, Martinican Creole, Seychelles Creole, among others) have not maintained gender categorization (Michaelis et al. 2013), nor have other

mixed languages lexified from languages with sex-based gender, such as Media Lengua (Stewart 2015, cited in Gillon and Rosen 2018: 169). Good (2012) additionally notes the difficulty of transferring noun classes in creole formation in general. This is even more striking when one considers that, unlike animacy, gender bears a relatively minimal functional load in Michif, as it is only apparent in the selection of definite and indefinite articles, some possessives, and prenominal adjectival agreement. Further, both sets of definite and indefinite articles are phonetically quite close to one another—*aeñ* and *en, li* and *la*—unlike most indicators of animacy, whose forms are generally more distinctive and thus harder to confuse. With the possible exception of these definite and indefinite articles, none of these constructions in which gender is marked are nearly as prevalent in actual usage as predication, where we find animacy being maintained.

Despite this overall alignment, however, several instances of misalignment with the French source are also attested. Arbitrary French feminines were found to be more likely to become masculine than all other kinds of nouns (i.e., default masculine, natural masculine, natural feminine), thus confirming our sub-hypothesis that exceptions to the alignment of Michif and French gender are cases where a default masculine gender is applied. Three quarters of the mismatches between Michif and French gender represent cases where the French translation equivalent is arbitrarily feminine in French, but receives masculine gender in Michif. The remaining instances of misalignment consist of three lemmas for which the French source form is masculine, but which receive feminine gender in Michif; and one which has feminine natural gender, but receives masculine gender in Michif. While several of these forms may be possible nonce speech errors, there are reports of similar phenomena in previous studies of Michif. The overall

frequency counts for these forms are very low, making it difficult to make any generalizations, but these instances suggest that it may be possible for nouns that have an arbitrary masculine value in the French source language to receive an arbitrary feminine value in Michif.

Unlike French-origin Michif nouns, whose gender values can generally be compared directly against the corresponding gender values for equivalent lexical items in the source language, English-origin nouns offer no source language gender value against which to compare. No significant association was found between Michif and French gender values for English-origin lemmas, which align with their French equivalents in only 66.7% of cases, thus disconfirming the hypothesis that Michif gender is determined by French translation equivalent (Hypothesis 2). English-origin lemmas which correspond to arbitrary French feminines were found to be much more likely to show misalignment than other kinds of nouns. This tendency for arbitrary French feminines to regularize to masculine gender in Michif was found to be even more pronounced for English-origin lemmas than for French-origin lemmas.

Several additional observations about gender in English-origin lemmas can also be made. For one, by far the majority (91.1%) of English-origin nouns in the dataset receive masculine gender, although feminine gender values are also attested in the corpus. In addition, most English-origin nouns that do receive feminine gender are found to correspond to a French source form which is also feminine, which would be compatible with a process of double relexification (i.e., with an original French-origin lexeme being replaced by an English equivalent, thereby mainining its original French feminine gender; cf. Bakker 1989b: 347). Further, for many of these items, additional

consultation with first-language speakers of Michif, as well as support from the descriptive literature in some cases, reveals them to be *consistently* feminine, rather than one-off occurrences or nonce speech errors. There are also instances in which English borrowings receive feminine gender in Michif even though the corresponding form is masculine in French. The anecdotal claim that all English-origin nouns in Michif receive masculine gender is thus not accurate when these cases are considered.

The animacy and gender values observed in the corpus for lexical items such as these strongly suggest that Michif speakers have extended the application of both noun classification systems to nouns which were not previously part of the language. Although a skew toward masculine gender for English-origin forms is noted, this is in line with gender assignment patterns for English borrowings noted in French and other languages cross-linguistically (see, e.g., Haugen 1969 for Norwegian; Kilarski 1997 for several Scandinavian languages; Őrsi 2012 for French). The existence of novel English forms with feminine gender rules out the possibility that masculine/feminine gender is completely unproductive in Michif, and the extension of animacy exceptions to English-origin nouns discussed above again suggests that the animacy values assigned to these items are based on more than natural animacy alone. It might also be the case that the phonological predictors of gender assignment in French (see Tucker, Lambert & Rigault 1977) may account for some of the instances of misalignment between Michif and French gender values, as well as the assignment of feminine gender to English-origin nouns. While further, dedicated investigation of speakers' intuitions concerning animacy and gender assignment to novel lexical items would be invaluable in this regard (e.g., through nonce-word animacy and gender assignment tests), the corpus-based results of this study

point to cases where attributing observed animacy and gender values solely to historical retention is implausible.

7.3 Michif animacy, gender, and inheritance

The results presented in the preceding two sections have shown that Michif nouns exhibit an overwhelming alignment in both their animacy and gender values with those of their French and Cree sources. Over 88% of the masculine/feminine gender values of French-origin Michif lemmas correspond in classification with their French sources, while over 92% of animate/inanimate gender values observed for lemmas in the entire dataset correspond to the classification of their sources in Cree. This holds even in cases of words such as *bwatoñ*$_{\text{ANIM}}$ 'button', *fariin*$_{\text{ANIM}}$ 'flour', and *tramb*$_{\text{ANIM}}$ 'tree', where Cree grammatical animacy does not align with natural animacy. Patterns such as this provide evidence that Michif has maintained both an animacy category inherited from Cree, as well as a gender category inherited from French, rather than aligning its nominal classification systems with more transparently grounded semantic categories such as natural animacy and gender. In addition, this close alignment between Michif and Cree animacy values, even in cases where this is in opposition to natural animacy, may suggest that the French-derived elements of the Michif lexicon are much less like borrowing than would be expected, given that borrowings are often assigned either a default or a natural value of a nominal classificatory category (Haugen 1969; Kilarski 1997; Őrsi 2012). This retention of nominal classification systems from both Cree and French, despite the vast majority of Michif nouns now having French forms, is relevant to current typologies of nominal classification, as we discuss in the following chapter.

Although Michif animacy and gender values are predominantly similar in their overall level of alignment with their source language forms, notable differences are found as well. In fact, this study reveals multiple instances in which animacy and gender values of Michif nouns do not correspond to those of their source language counterparts (e.g, Michif *kosh*$_{\text{INAN}}$ vs. Cree *âsiyân*$_{\text{ANIM}}$ 'diaper', Michif *seuk*$_{\text{INAN}}$ vs. Cree *sôkâw*$_{\text{ANIM}}$ 'sugar'; Michif *grañdrii*$_{\text{FEM}}$ vs. French *grenier*$_{\text{MASC}}$ 'granary', Michif *fii*$_{\text{MASC}}$ vs. French *fille*$_{\text{FEM}}$ 'girl'). While both systems of nominal classification present examples of mismatches that appear to be influenced by natural gender or animacy, these do not represent the majority of instances of misalignment in either system. In the case of animacy, mismatches between Michif and Cree animacy values are often the apparent result of a process of semantic analogy, whereby Michif lexical items were treated as belonging to an existing class of arbitrary animate nouns that is less extensive in Cree (e.g., vehicles, with Michif including several more nouns referring to self-propelled modes of transportation in this class of arbitrary animate nouns than Plains Cree does). Notably, these kinds of analogically motivated exceptions are not found in the sex-based gender system of Michif. Instead, the majority of instances of misalignment between Michif and French gender values are the result of arbitrary French feminines shifting to masculine values in Michif.

With the exception of the few forms of semantic analogy affecting animacy noted above, instances of levelling and apparent simplification of arbitrary classifications of Michif nouns are considerably more prevalent for sex-based gender than for animacy. This may be due to the different semantic bases that underlie both systems of classification, with animate/inanimate values resulting in fewer arbitrary classifications to

begin with than masculine/feminine values when applied to sexless referents. Some of these discrepancies may also point to language attrition and/or language change (e.g., since a decrease in Michif language use may provide fewer opportunities for exceptions to natural animacy and gender falling outside of well-established classes to be maintained). This would be in line with other languages for which attrition has been identified as a factor in the reduction and/or restructuring of noun classification systems, such as Young People's Dyirbal and East Sutherland Gaelic (Aikhenvald 2000: 390). Establishing this, however, would require comparison with records of Michif as spoken by previous generations. These differences in animacy and gender between Michif and its source languages could also be the result of internally-driven language change, with this overall quite complex system of noun categorization moving towards simplification, although this would be difficult to demonstrate on the basis of the evidence presently available. In general, the presence of misaligned animacy and gender values in Michif means that it is not possible to predict the animacy and gender values of every noun in Michif based solely on the corresponding French or Cree values. While these differences deserve attention, it should be borne in mind that misalignment on the whole is essentially the exception rather than the norm, with only 6.5% of Michif animacy values and 14.4% of Michif gender values found to differ from their equivalents in Michif's source languages.

In addition, lexical items of both French and English origin show essentially the same patterns of animacy assignment, with no significant differences noted between these forms; both regular animacy values (i.e., those aligned with natural gender) and arbitrary animacy values are associated with Michif nouns from both English and French. This

does not appear to be the case with gender, where markedly different assignment patterns are noted for French and English-origin nouns. As noted previously, Michif nouns of French origin predominantly follow their French equivalents in their gender values, while the majority of English-origin nouns are assigned masculine gender values. These differences in how nouns from different source languages are categorized again distinguishes animacy from gender in Michif, and provides further evidence that these two systems of classification are distinct.

The assignment of animacy and gender to English loanwords noted in this dataset also offers potential insights into the synchronic behavior of noun classification in Michif. In comparison to Cree- and French-origin nouns, English borrowings constitute a later addition to the Michif lexicon, and are not generally considered to have played a significant role in the historical language contact situation which gave rise to Michif as a distinct language (see Chapter Two). The patterns noted here show that the application of animacy to nouns did not end after this period of contact. Moreover, the animacy and gender values that these nouns receive synchronically cannot generally be assumed to have entered Michif exclusively through inheritance, suggesting that both animacy and gender categories are still active in Michif.

On the whole, the corpus evidence presented in this chapter suggests that Michif has maintained two distinct and active systems of noun classification, with little evidence of either extensive innovation or simplification of these systems. The strong correlation between the animacy and gender values of Michif nouns and those of their French and Cree equivalents suggests that the basic attributes of nominal classification in Michif have been retained from both source languages, while the ongoing assignment of

animacy and gender values to newer lexical items provides additional evidence that both systems remain active in the language. In the following chapter, we discuss the implications of these findings, particularly as they relate to ongoing research on Michif and to typologies of nominal classification.

Chapter 8: Discussion and conclusions

Both cultural and cognitive aspects of classification are arguably relevant in situations of language contact, where nominal classification systems may be observed to be transferred, in part or in whole, between varieties in the development of lexicons in contact languages. Michif presents a particularly interesting case for the study of nominal classification, having inherited two systems of nominal classification from its source languages—Romance-derived gender (masculine/feminine) and Algonquian-derived animacy (animate/inanimate). As previous research has shown, the maintenance of multiple systems of nominal classification is typologically uncommon, as is the transfer and maintenance of noun class systems in language contact situations (Corbett 1991; Good 2012; see also Gillon & Rosen 2018; Stoltzfus & Boissard 2016, who claim that the Michif gender system is either weakening or fossilized). Accordingly, this dissertation has focused on investigating animacy and gender in Michif to determine the extent to which these nominal classification systems remain independent and productive. Through quantitative investigation of animacy and gender assignment patterns based on a documentary corpus of over 60 hours of spontaneous spoken Michif, this study finds that Michif indeed has two independent and productive grammatical categories of noun classification, each inherited from a different source language.

Quantitative analysis of the Michif dataset finds that the animacy and gender values of Michif nouns align with those of its Cree and French source languages in the overwhelming majority of cases, indicating that these systems have largely been inherited in their full complexity in Michif. Statistical tests also find no signs of interaction

between animacy and gender systems, motivating a treatment of Michif as having two separate, co-existing systems of nominal classification, rather than a single, merged system. Finally, a statistically significant difference is found between gender assignment patterns in French-origin lemmas as compared to English-origin lemmas, while no such difference is found in animacy assignment patterns.

This study also finds that arbitrary French feminines and arbitrary Cree animates are more susceptible to regularization than other kinds of nouns. The majority of instances of misalignment between Michif and French gender values represent cases of French arbitrary feminines being assigned default masculine values in Michif, although it is found that animacy is also not immune to regularization to a default grammatical value. The synchronic results of this study are thus partially compatible with the hypothesis that French-derived gender will be less stable than animacy over time (cf. Gillon & Rosen 2018).

Finally, the results of this study demonstrate that animacy and gender remain productive categories in Michif, rather than appearing only as fossilized elements in nominal constructions, as has been claimed by some (cf. Stoltzfus & Boissard 2016). This is supported by the observations that a) every lemma in the language must have values for animacy and gender, as indicated by the mandatory nature of grammatical agreement for these categories; b) with few exceptions, these values are stable and shared by speakers; and c) these values are always assigned to new lexical items brought into the language, even when the resulting classifications cannot be easily attributed to inheritance, as in the case of English borrowings. These results have implications for

Michif language description, for typologies of nominal classification systems, and for Michif language documentation and lexicography, as discussed in the following sections.

8.1 Implications for Michif language description

As the first full-length study of nominal classification in Michif, this dissertation brings attention to an area of Michif grammar that has only recently begun to be the subject of dedicated, systematic investigation. Few, if any, of the claims made in the literature concerning Michif nominal classification have been compared against corpus data, with these studies generally relying instead on elicitation, metalinguistic discussion, and/or examples drawn largely from introspection-based dictionaries such as Laverdure & Allard (1983). While results based on these kinds of data are valuable, they also leave open the possibility of discrepancies between the self-reported, careful speech found in such sources and conventional patterns of usage found among a broader range of speakers, communities, and contexts of use. This study attempts to address this gap by examining Michif gender as it appears in connected, spontaneous discourse, while aiming to be as transparent as possible about the nature of the underlying documentation and how it has been brought into this analysis. The corpus-based methodology adopted by this study has also revealed both inter- and intra-speaker variability for both animacy and gender categories. Such variability has rarely been documented or addressed in previous research and challenges the common characterization of animacy and gender as being entirely fixed and invariable in Michif.

This study also brings attention to other aspects of Michif noun classification that have escaped systematic attention in much of the literature to date. Only a handful of

authors make any reference to the assignment of grammatical gender to English-origin nouns in Michif, and those who do tend to focus only on masculine/feminine gender, leaving animate/inanimate gender effectively unaddressed (e.g., Hogmen 1981; Papen 2003a). Indeed, the assignment of gender to English nouns in Michif is identified as an area in need of research by both Bakker (1997: 105) and Papen (2003a: 139). In addition, animacy in Michif is typically treated only in passing, with the suggestion that it corresponds to Cree. This study presents the first empirical test of the claim that animacy in Michif generally corresponds to Cree, and largely supports it.

Finally, the results of this study speak directly to a number of recent claims made in the literature regarding Michif nominal classification. As mentioned above, based on the central role of animacy in Michif grammar, Gillon and Rosen (2018: 101) predict that sex-based gender will erode over time, while animacy will remain stable. Gillon and Rosen (2018) seem to suggest that this loss of marked sex-based gender distinctions is largely internally driven (i.e., by the low functional load placed on sex-based gender marking in Michif and a tendency towards unmarked values in cases of gender assignment without clear semantic motivations), rather than solely by processes of attrition. While this study is unable to comment on possible diachronic changes in Michif due to a lack of historical recordings or longitudinal data, the synchronic findings of this work at least partially corroborate this hypothesis. In Section 7.2, we note a number of arbitrary French feminines receiving default masculine values in Michif for English-origin nouns, and many of the instances of misalignment between Michif and French for the French-origin nouns reflect this same pattern as well. This, together with attested variability in gender assignment as well as a lack of consistent agreement in other aspects

of the grammar (e.g., in prenominal adjectives; cf. Section 4.2.3), might be interpreted to suggest that sex-based gender is in the process of weakening in Michif. However, the gender values of French-origin nouns nevertheless show alignment with those of their French translation equivalents in the majority of cases, and even though a preference is noted for masculine gender in English-origin nouns, this same pattern has been identified in French, where gender has not been claimed to be weakening. It should also be noted that the same data also show variability in animacy assignment, as well as a tendency for some arbitrary Cree animates to receive default inanimate values in Michif, which, by the same reasoning, would point to the possibility of weakening of animacy as well, which Gillon and Rosen (2018) do not predict. These results thus suggest that animacy is not immune to this kind of leveling, despite its importance to Michif grammar.

8.2 Implications for typologies of nominal classification

The findings of this study provide evidence which suggests that Michif is best analyzed as having two separate, co-existing systems of nominal classification, rather than a single combined system. For one, this study finds no statistically significant association between animacy and gender, suggesting that the two systems of nominal classification inherited from Michif's source languages have remained distinct and function independently of one another. Thus, this study finds no signs of interaction between animacy and gender which would compellingly motivate a merged analysis. Moreover, it appears that Michif speakers are applying knowledge particular to each system of classification to Michif nouns in different ways, with animacy and sex-based gender assignment each having their own dynamics and following their own particular patterns. Animacy assignment in Michif generally follows the same overall pattern as Cree animacy, regardless of whether

the target noun is from French or from English, including in cases of "exceptional" animacy assignment. Gender assignment, on the other hand, is observed in this study to treat French- and English-origin nouns differently, aligning most French-origin nouns with the gender of the French equivalent, while the majority of English-origin nouns are assigned default masculine gender. Thus, masculine is the most prevalent gender assigned to English-origin nouns for both animate *and* inanimate nouns. The observation of different gender assignment patterns for animacy and gender based on source language that emerged from quantitative analysis thus suggests that a treatment of Michif as having two separate systems of nominal classification is preferable to a merged analysis.

In addition, possible motivations and patterns applied in instances of mismatches between Michif and Cree animacy values differ from those of mismatches between Michif and French gender values. Most cases of misalignment for animacy between Michif and Cree involve a process of semantic analogy (especially for more recent innovations and English borrowings), whereas this is not attested for masculine/feminine gender at all in the corpus. Meanwhile, the trend for cases of misalignment between French and Michif is for nouns that are arbitrary feminines in French to become default masculines in Michif, which could be seen as extending the same gender assignment pattern as French to novel cases. Thus, the overall dynamics of how animacy and gender are assigned to Michif nouns differ. Speakers appear to be applying other kinds of knowledge of Michif nouns in their generalizations about animacy and gender assignment that extend beyond the natural semantics of referents of those nouns. These observations suggest that the most plausible analysis is that Michif has two co-existing systems of nominal classification, one based on a masculine/feminine distinction, and the

other based on an animate/inanimate distinction. In this light, Michif would present an instance of a combined nominal classification system in which multiple distinct systems of categorization exist (see Chapter Three for discussion), which is cross-linguistically rare (cf. Corbett 1991). With few contact languages having been considered in cross-linguistic studies of gender systems, closer investigation of the dynamics of noun classification in contact languages like Michif may have important contributions to make to current typologies of nominal classification.

8.3 Implications for Michif language documentation and lexicography

The results of this study also have implications for Michif documentation and lexicography, underscoring the importance of documenting both animacy and gender values for all Michif nouns. From a documentary perspective, Michif animacy and gender values have not been systematically noted in prior research. This study offers a source of information on Michif animacy and gender values that is unavailable in other written sources. Moreover, construction of the corpus for this study has served to identify a number of lexical items that, to my knowledge, have not been documented in any other Michif lexical resources to date (e.g., BICH_LAÑP 'bitumen lamp,' BWAA_DI_KOR 'cordwood,' GRAÑDRII 'granary,' JEEÑG 'girl; girlfriend,' KARABIN 'rifle,' PAAÑS 'rumen (cow stomach),' PAEÑGFAEÑ 'lice comb', etc.). This work thus serves an immediate purpose in addressing these gaps in present descriptions of Michif, as well as contributing to a fuller picture of nominal classification in the language.

In terms of lexicography, it is standard practice for the animate/inanimate values of nouns to be provided in Algonquian lexical resources, and likewise, for

masculine/feminine values to be provided in French lexical resources. However, no current Michif lexical resources, either academic or community-based, indicate both animacy and gender values for nominal entries. One source (Rosen 2016b) currently provides masculine/feminine gender values but no animacy values, while no other lexical resources for Michif currently provide either animacy or gender values (e.g., Fleury 2018; Laverdure & Allard 1983). Given how central animacy has been argued to be to the overall grammatical structure of Michif (e.g., being marked in virtually all forms of predication), the absence of animacy values from current lexical resources presents a significant gap in the basic documentation of the language. This information is also critically important for language education and revitalization programs. While Michif animacy and gender values are seen in this study to frequently correspond to those of their French and Cree equivalents, this is not always the case, and both consistent differences between Michif and its source languages and variability in Michif animacy and gender assignment patterns are attested (cf. Chapter Seven). This is especially true of English borrowings, whose animacy and gender assignment patterns are at times relatively less clearly derived from French and Cree. Taken together, these observations highlight the pressing need for further documentation of the Michif lexicon, and for this documentation to be reflected in Michif lexical resources.

8.4 Future directions

These results point to several avenues for future work, both cross-linguistically and for Michif specifically. Having established that Michif has two separate systems of nominal classification and as such presents an example of a language with a combined nominal

classification system (cf. Corbett 1991), a typological survey is in order to determine whether or not the same outcome can be found for other mixed languages, or if this development is unique to Michif. As such systems are uncommon, exploring the hypothesis that the sociolinguistic conditions under which mixed languages emerge may also foster the development of combined nominal classification systems may provide one means of identifying other instances of combined nominal classification systems, while also situating Michif more precisely within a larger typological context.

Along with this typologically oriented research, further language-specific investigation of gender-marking constructions in Michif would also be of value, delving further into aspects of gender agreement where notable departures from the source languages are observed. This study has taken its evidence from several constructions (e.g., articles, demonstratives, possessives) to determine that masculine/feminine gender is a productive category in Michif. Given this conclusion, the apparent lack of consistency in gender agreement in prenominal adjectives is all the more striking, as all other gender-indicating constructions in Michif appear to have retained these distinctions more systematically. More research is needed to determine whether or not gender agreement is a productive process among Michif prenominal adjectives, as has been claimed by some (e.g., Bakker 1997: 102; Bakker & Papen 1997), or if the observed variability can be attributed to fossilization of adjective-noun pairs or free variation in agreement marking in this constructional context (cf. Rosen 2007 and Gillon & Rosen 2018 for further discussion on the lack of gender agreement in adjectival contexts). This would complement the results of this study by providing a fuller picture of the overall retention of exponents of the inherited French-based gender system in Michif.

Preliminary work in this area on the basis of the corpus developed for this study and a limited amount of targeted elicitation with speakers of Michif suggests that there is exceptional variability in the degree of observed adjective-noun agreement between both individual speakers and individual adjective-noun pairs (see Section 4.2.3), although this would require further, dedicated research to describe adequately.

This study has observed several general trends in animacy and gender assignment in Michif, relating them primarily to the semantics of the nouns involved. However, there remains the possibility that the phonological forms of nouns may also contribute to this process. Thus, an understanding of the dynamics of masculine/feminine gender assignment in Michif may benefit from a phonological categorization of gender assignment rules similar to what Tucker, Lambert, and Rigault (1977) have proposed for Standard French. As noted in 8.1 above, it is possible that at least some Michif speakers may be applying phonological generalizations to the noun lexicon to determine what gender is assigned to novel word forms. This line of inquiry would likely require more extensive documentation of Michif gender values for a wider range of lemmas than is presently available, although elicitation of gender values assigned to nonce words might provide another means of approaching speakers' intuitions in this area. In this vein, another potential avenue of research would be to compare Michif's assignment of gender to loanwords (and potentially nonce/non-words) with how gender is assigned to loanwords in French and Cree. Some research in this area is available for French (e.g., Lupu 2005; Nymannson 1995; Őrsi 2012; Saugera 2017, among others), but I am not aware of any similar work concerning animacy and loanwords in Cree or related Algonquian languages.

Finally, perhaps one of the most critical areas on which the results of this study may have some bearing is in Michif language education and revitalization. This study establishes that native speakers of Michif show sensitivity to both animacy and gender, and that this knowledge represents one part of their larger communicative competence. These classifications are not entirely predictable based on the semantics of their referents, and do not consistently map onto natural animacy and gender. At the same time, the use of both animacy and gender directly impacts agreement in many other frequent constructions in the language, such as verbal inflection and demonstratives, making them a particularly important part of the target of acquisition for emergent speakers of Michif. The descriptive linguistic results of the kind presented in this study regarding the distribution and dynamics of animacy and gender in Michif may therefore set the stage for research in applied linguistics that explores how the animacy and gender values and assignment patterns identified can be effectively integrated into language education programs and resources. Particularly given the present state of endangerment and the strong interest expressed by many Métis individuals and communities in reclaiming Michif as one of their heritage languages, it might be hoped that the results of this study would contribute in some way to supporting these efforts in the future.

References

Ahenakew, Freda. 1987. *Cree Language Structures: A Cree Approach*. Winnipeg: Pemmican Publications.

Ahenakew, Vince. 1997. *Michif/Cree dictionary: Nehiyawewin masinahikan*. Saskatoon: Saskatchewan Indian Cultural Centre.

Ahenakew, Vince. 2009. *Nêhiyawêwin Masinahikan: Michif/Cree dictionary*. Saskatoon, Saskatchewan: The Gabriel Dumont Institute of Native Studies and Applied Research.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices* (Oxford Studies in Typology and Linguistic Theory). Oxford/New York: Oxford University Press.

Aikhenvald, Alexandra Y. 2017. A typology of noun categorization devices. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *The Cambridge Handbook of Linguistic Typology* (Cambridge Handbooks in Language and Linguistics), 361–404. Cambridge University Press.

Andrella, Orlando. 1983. *Coexistent systems: The evidence from Mechif*. Grand Forks, North Dakota: University of North Dakota Master's thesis.

Ansaldo, Umberto. 2011. Metatypy in Sri Lanka Malay. In Rajendra Singh & Ghanshyam Sharma (eds.), *Annual Review of South Asian Languages and Linguistics: 2011* (Trends in Linguistics 241), 3–16. Berlin/Boston: Walter de Gruyter.

Antonov, Anton. 2015. The Obviative as a Differential Object Marker in Michif. University of New Mexico, Albuquerque.

Austin, Peter K. 2010. Current issues in language documentation. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 7, 12–33. London: SOAS.

Ayoun, Dalila. 2010. Corpus data: Shedding the light on French grammatical gender … or not. *EUROSLA Yearbook* 10(1). 119–141. doi:10.1075/eurosla.10.08ayo.

Ayoun, Dalila. 2018. Grammatical gender assignment in French: Dispelling the native speaker myth. *Journal of French Language Studies* 28(1). 113–148. doi:10.1017/S095926951700014X.

Baayen, R.H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge UP.

Bakker, Peter. 1989a. Is Michif a creole language? *Amsterdam Creole Studies* 10. 40–47.

Bakker, Peter. 1989b. Relexification in Canada: The case of Métif (French-Cree). *Canadian Journal of Linguistics / Revue canadienne de Linguistique* 34(3). 339–350.

Bakker, Peter. 1990. The genesis of Michif: A first hypothesis. In William Cowan (ed.), *Papers of the 21st Algonquian Conference*, 12–35. Ottawa: Carleton University.

Bakker, Peter. 1991. The Ojibwa Element in Michif. (Ed.) William Cowan. *Papers of the 22nd Algonquian Conference* 22. 11–20.

Bakker, Peter. 1992. *A language of our own: The genesis of Michif – the mixed Cree-French language of the Canadian Métis*. Amsterdam: University of Amsterdam PhD thesis.

Bakker, Peter. 1994. Michif, the Cree-French mixed language of the Metis buffalo hunters in Canada. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15*

case studies in language intertwining (Studies in Language and Language Use 13), 13–33. Dordrecht: ICG Printing. (10 August, 2017).
Bakker, Peter. 1996. "When the Stories Disappear, Our People Will Disappear": Notes on Language and Contemporary Literature of the Saskatchewan Plains Cree and Métis. *Studies in American Indian Literatures* 8(4). 30–45.
Bakker, Peter. 1997. *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Metis*. [Rev. ed.]. Oxford: Oxford University Press.
Bakker, Peter. 2004a. The Michif language of the Metis. In Lawrence Barkwell (ed.), *La lawng: Michif peekishkwewin: The heritage language of the Canadian Metis*, vol. Two: Language Theory, 5–9. Winnipeg, Manitoba: Pemmican Publications.
Bakker, Peter. 2004b. Spelling Systems for Michif: An Overview. In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. Two: Language Theory, 11–28. Winnipeg, Manitoba, Canada: Pemmican Publications.
Bakker, Peter. 2004c. The Verb in Michif. In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. Two: Language Theory, 63–80. Winnipeg, Manitoba, Canada: Pemmican Publications.
Bakker, Peter. 2004d. What is Michif? In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. 1, 5–7. Winnipeg, Manitoba, Canada: Pemmican Publications.
Bakker, Peter. 2011. Emerging publications in the Michif language. In Kerstin Knopf (ed.), *North America in the 21st Century: Tribal, Local, and Global: Festschrift Für Hartmut Lutz*, 223–37. Trier: WVT Wissenschaftlicher Verlag Trier.
Bakker, Peter & Lawrence Barkwell. 2004. Storytelling and mythology. In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. Two: Language Theory, 83–96. Winnipeg, Manitoba, Canada: Pemmican Publications.
Bakker, Peter & Norman Fleury. 2004. *Learn Michif by listening* (Audio CD Produced by Peter Bakker and Norman Fleury). Michif Language Program, Manitoba Métis Federation.
Bakker, Peter, John Gosselin & Ida Rose Allard. 1998. Hoe Brave Hond Bob aan zijn naam kwam (How good dog Bob got his name). In Adrienne Bruyn & Jacques Arends (eds.), *Mengelwerk voor Muysken. Voor Pieter C. Muysken bij zijn afschied van de Univesiteit van Amsterdam*, 262–266. Amsterdam: Publikaties van het Intituut voor Algemene Taalwetenschap.
Bakker, Peter & Robert A. Papen. 1996. Michif language and other languages of the Canadian Metis. In Stephen Adolphe Wurm, Peter Mühlhäusler & Darrell Trevor Tryon (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*, vol. 1: Maps, 1171–1184. Berlin: Mouton de Gruyter.
Bakker, Peter & Robert A. Papen. 1997. Michif: A Mixed Language based on Cree and French. In Sarah Thomason (ed.), *Contact Languages: A Wider Prespective*, 295–363. Amsterdam: Benjamins.
Barkwell, Lawrence (ed.). 2004a. *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*. Vol. 1: Language Practice. Winnipeg, Manitoba, Canada: Pemmican Publications.

Barkwell, Lawrence (ed.). 2004b. *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*. Vol. 2: Language Theory. Winnipeg, Manitoba, Canada: Pemmican Publications.

Barkwell, Lawrence, Norman Fleury & David Morin. 2017. *Michif language research, literature review, teaching resources, and annotated bibliography*. https://www.scribd.com/document/367524485/Michif-Language-Research-Literature-Review-Teaching-Resources-and-annotated-Bibliography (8 January, 2018).

Barkwell, Lawrence J., Leah Dorian, Darren Préfontaine & Nicole Rosen. 2004. Annotated bibliography. In Lawrence J. Barkwell (ed.), *La lawng: Michif peekishkwewin: The heritage language of the Canadian Métis*, vol. 2, 97–128. Winnipeg: Pemmican Publications.

Barkwell, Lawrence J., Leah Dorion & Darren Préfontaine (eds.). 2001. *Metis legacy: A Metis historiography and annotated bibliography*. Winnipeg: Pemmican Publications.

Bidot, Émile. 1925. *La Clef du genre des substantif français (Méthode dispensant d'avoir recours au dictionnaire)*. Poitiers: Imprimerie nouvelle.

Bitterman, Chester, Clarice Crawford, Barbara Glaser, Karen Peterson, Douglas Sabin, Glenn Sage & Lucia Tedesco. 1976. Speech acts and sentence types: Michif field methods. Summer Institute of Linguistics, University of North Dakota Session, ms.

Black, Mary R. 1967. *An Ethnoscience Investigation of Ojibwa Ontology and World View*. Stanford: Stanford University.

Black-Rogers, Mary B. 1982. Algonquian gender revisited: Animate nouns and Ojibwa 'power' — An impasse? *Papers in Linguistics* 15(1). 59–76. doi:10.1080/08351818209370560.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston,.

Bloomfield, Leonard. 1946. Algonquian. In Harry Hoijer (ed.), *Linguistic structures of native America* (Viking Fund Publications in Anthropology 6), 85–129. New York, New York, United States: Viking Fund.

Boerger, Brenda. 2011. To BOLDly Go Where No One Has Gone Before. http://scholarspace.manoa.hawaii.edu/handle/10125/4499 (16 November, 2017).

Boteler, Bette. 1971. *The relationship between the conceptual outlooks and the linguistic description of disease and its treatment among the Chippewa and/or Cree Indians of the Turtle Mountain Reservation*. Grand Forks: University of North Dakota Master's thesis.

Bowerman, Melissa & Eric Pederson. 1993. Topological relations picture series. In Eve Danziger & Deborah Hill (eds.), *Manual for the space stimuli kit 1.2*, Max Planck Institute for Psycholinguistics. Nijmegen: Max Planck Institute for Psycholinguistics.

Browne, Timothy Di Leo. 2005. *Michif: A Study in Language Contact and Relexification*. Cambridge, MA: Harvard University.

Brugman, Hennie & Albert Russel. 2004. Annotating multi-media / multi-modal resources with ELAN. In Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa & Raquel Silva (eds.), *Proceedings of the 4th International*

Conference on Language Resources and Language Evaluation, 2065–2068. Paris: European Language Resources Association.
Burnouf, Laura, Norman Fleury & Guy Lavallée. 2007. *The Michif resource guide: Lii Michif niiyanaan, aan Michif biikishwanaan*. Saskatoon: Gabriel Dumont Institute.
Burton, Wilfred & Anne Patton. 2011. *Call of the fiddle: Li vyayloon ka taypwatikooyen*. (Trans.) Norman Fleury. Saskatoon: Gabriel Dumont Institute.
Chafe, Wallace (ed.). 1980. *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Ablex Publishing Corp.
Chelliah, Shobhana L. & Willem J. de Reuse. 2011. *Handbook of Descriptive Linguistic Fieldwork*. Springer Netherlands. //www.springer.com/la/book/9789048190256 (27 March, 2018).
Coetsem, Frans van. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg: Universitätsverlag Winter.
Comrie, Bernard. 1999. Grammatical Gender Systems: A Linguist's Assessment. *Journal of Psycholinguistic Research* 28(5). 457–66.
Conathan, Lisa. 2011. Archiving and language documentation. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 235–254. Cambridge: Cambridge University Press.
Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2006. *Agreement*. Cambridge, UK ; Cambridge University Press,.
Corbett, Greville G. 2011. The penumbra of morphosyntactic feature systems. *Morphology* 21(2). 445–480. doi:10.1007/s11525-010-9171-4.
Corbett, Greville G. 2013a. Number of Genders. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/30 (5 July, 2017).
Corbett, Greville G. 2013b. Sex-based and Non-sex-based Gender Systems. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/31 (5 July, 2017).
Corbett, Greville G. 2013c. Systems of Gender Assignment. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/32 (5 July, 2017).
Corbett, Greville G. (ed.). 2014. *The Expression of Gender* (The Expression of Cognitive Categories (ECC) 6). Berlin/Boston: Mouton de Gruyter. https://www.degruyter.com/view/product/203689 (22 August, 2017).
Crawford, John C. 1973. Linguistic and sociolinguistic relationships in the Michif language. *Proceedings of the Linguistic Circle of Manitoba and North Dakota* 8. 18–22.
Crawford, John C. 1976. Michif: A new language. *North Dakota English* 4. 3–10.
Crawford, John C. 1979. The standardization and instrumentalization of creole languages: Standardization of orthography in Michif. Presented at the Conference on Theoretical Orientations in Creole Studies, St. Thomas, Virgin Islands.

Crawford, John C. 1985a. Speaking Michif in four Metis communities. *The Canadian Journal of Native Studies* 3(1). 47–55.
Crawford, John C. 1985b. Dialects of Michif: A beginning. *Proceedings of the Linguistic Circle of Manitoba and North Dakota* 25. 14–15.
Crawford, John C. 1985c. What is Michif? Language in the Métis tradition. In Jacqueline Peterson & Jennifer S. H. Brown (eds.), *The New Peoples: Being and Becoming Métis in North America*, 131–141. Winnipeg: University of Manitoba Press.
Cruz, Levi Cirilo. 2016. *Documenting conversations using basic oral language documentation (BOLD) method*. Chiang Mai, Thailand: Payap University Master's thesis.
Dahlstrom, Amy. 1991. *Plains Cree morphosyntax*. New York: Garland.
Dahlstrom, Amy. 1995. Motivation vs. predictability in Algonquian gender. In David Pentland (ed.), *Papers of the 26th Algonquian Conference*, 52–66. Winnipeg: University of Manitoba.
Dahlstrom, Amy Louise. 1986. *Plains Cree morphosyntax*. Berkeley, California, United States: University of California, Berkeley Unpublished doctoral dissertation.
Darnell, Regna. 1991. Thirty-nine postulates of Plains Cree conversation, "power," and interaction: A culture-specific model. (Ed.) William Cowan. *Papers of the 22nd Algonquian Conference* 22. 89–102.
Darnell, Regna & Anthony L. Vanek. 1976. The Semantic Basis of the Animate/Inanimate Distinction in Cree. *Papers in Linguistics* 9(3–4). 159–180.
Donohue, Mark. 2001. Animacy, class and gender in Burmeso. In Andrew Pawley, Tom Dutton & Malcolm Ross (eds.), *The boy from Bundaberg: studies in Melanesian linguistics in honour of Tom Dutton* (Pacific Linguistics 514), 97–115. 1st. publ. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies.
Douaud, Patrick C. 1980. Metis: A Case of Triadic Linguistic Economy. *Anthropological Linguistics* 22(9). 392–414.
Douaud, Patrick C. 1989. Mitchif: an aspect of francophone Alberta. *Journal of Indigenous Studies* 1(2). 80–90.
E-MELD. 2006. E-MELD School of Best Practice. Text. http://emeld.org/school/index.html (17 November, 2017).
Enger, Hans-Olav. 2009. The Role of Core and Non-Core Semantic Rules in Gender Assignment. *Lingua: International Review of General Linguistics* 119(9). 1281–1299.
Evans, Donna. 1982. On coexistence and convergence of two phonological systems in Michif. *Work Papers of the Summer Institute of Linguistics, University of North Dakota* 26. 158–173.
Fitzsimmons, Matthew, Alexah Konnelly, Sarah Provan & Alison Root. 2015. An analysis of the Split-Phonology Hypothesis in Michif: Sociolinguistic and phonological perspectives. *Working Papers of the Linguistics Circle* 25(1). 10–24.
Flamand, Rita. 2002. *Michif conversational lessons for beginners*. Winnipeg: Métis Resource Centre.
Fleury, Norman. 1999a. *Tawnshi en itwayk en Michif?* Winnipeg: Manitoba Métis Federation Michif Languages Project.
Fleury, Norman. 1999b. *Michif vocabulary*. Winnipeg: Manitoba Métis Federation Michif Languages Project.

Fleury, Norman. 2000. *La lawng: Michif peekishkwewin: The Canadian Michif language dictionary*. (Ed.) Lawrence J. Barkwell. Winnipeg: Metis Resource Centre and Manitoba Metis Federation.
Fleury, Norman. 2018. Michif Dictionary. http://www.metismuseum.ca/michif_dictionary.php (15 January, 2018).
Fleury, Norman, Gilbert Pelletier, Jeanne Pelletier, Joe Welsh, Norma Welsh & Janice DePeel. 2008. *Stories of Our People Lii zistwary di la naasyoon di Michif: A Metis Graphic Novel Anthology*. Saskatoon: Gabriel Dumont Institute.
Frantz, Donald G. 2009. *Blackfoot Grammar*. Second. Toronto, Ontario, Canada: University of Toronto Press.
Frawley, William, Kenneth C. Hill & Pamela Munro. 2002. Making a dictionary: Ten issues. In William Frawley, Kenneth C. Hill & Pamela Munro (eds.), *Making dictionaries : Preserving indigenous languages of the Americas*, 1–22. Berkeley: University of California Press.
Gabriel Dumont Institute of Métis Studies and Applied Research. 2018. The Virtual Museum of Métis History and Culture. http://www.metismuseum.ca/index.php (8 January, 2018).
Gillon, Carrie. 2015. Innu-aimun plurality. *Lingua* 162. 128–148. doi:http://dx.doi.org/10.1016/j.lingua.2015.05.006.
Gillon, Carrie. 2016. Plural plurals in Innu-aimun and Michif. Presented at the Workshop on the Structure and Constituency of the Languages of the Americas, University of Arizona.
Gillon, Carrie & Nicole Rosen. 2015a. Reanalyzing Michif determiners. Presented at the Linguistics Department Colloquium, University of Arizona.
Gillon, Carrie & Nicole Rosen. 2015b. The DP domain in Michif: Michif as an Algonquian Language. Presented at the Gender, Class, and Determination: A Conference on the Nominal Spine, University of Ottawa.
Gillon, Carrie & Nicole Rosen. 2016. Critical mass in Michif. *Journal of Pidgin & Creole Languages* 31(1). 113–140. doi:10.1075/jpcl.31.1.05gil.
Gillon, Carrie & Nicole Rosen. 2018. *Nominal Contact in Michif*. Oxford, New York: Oxford University Press.
Goddard, Ives. 1984. The obviative in Fox narrative discourse. In William Cowan (ed.), *Papers of the 15th Algonquian Conference*, vol. 15, 273–286. Ottawa, Ontario, Canada: Carleton University.
Goddard, Ives. 1990. Aspects of the topic structure of Fox narratives: Proximate shifts and the use of overt and inflectional NPs. *International Journal of American Linguistics* 7(1). 317–340.
Goddard, Ives. 2002. Grammatical gender in Algonquian. In H.C. Wolfart (ed.), *Papers of the 33rd Algonquian Conference*, 195–231. Winnipeg: University of Manitoba.
Good, Jeff. 2011. Data and language documentation. In Peter K. Austin & Julia Sallabank (eds.), *Handbook of Endangered Languages*, 212–234. Cambridge: Cambridge University Press.
Good, Jeff. 2012. Typologizing grammatical complexities: or Why creoles may be paradigmatically simple but syntagmatically average. *Journal of Pidgin & Creole Languages* 27(1). 1–47.

Good, Jeff, Tom Myers & Alexander Nakhimovsky. 2010. Interoperability for language documentation: The role of semantic web tools.
Gordey, Louise & Norman Fleury. 2011. *Speaking Michif resources: Lessons in Michif-Cree for the whole family*. Winnipeg: Louis Riel Institute.
Government of Canada, Statistics Canada. 2016. Language Highlight Tables, 2016 Census - Aboriginal mother tongue, Aboriginal language spoken most often at home and Other Aboriginal language(s) spoken regularly at home for the population excluding institutional residents of Canada, provinces and territories, 2016 Census – 100% Data. *Language Highlight Tables, 2016 Census*. http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/lang/Table.cfm?Lang=E&T=41&SP=1&geo=01 (2 January, 2018).
Greenberg, Joseph H. 1954. Concerning Inferences from Linguistic to Nonlinguistic Data. In Harry Hoijer (ed.), *Language in Culture*, 3–19.
Gries, Stefan Th. 2009. *Quantitative Corpus Linguistics with R: A Practical Introduction*. New York / London: Routledge.
Hallowell, A. Irving. 1955. *Culture and experience*. Prospect Heights, IL : Waveland Press, 1955.
Hallowell, A. Irving. 1976. *Contributions to anthropology : selected papers of A. Irving Hallowell*. Chicago: University of Chicago Press.
Hasler, Laurel Anne & H.C. Wolfart. 2002. Sustained and shifting obviation in a Sheshatshiu Innu-aimun story. *Papers of the 33rd Algonquian Conference*, vol. 33, 232–250. Winnipeg, Manitoba, Canada: University of Manitoba Press.
Haugen, Einar. 1969. *The Norwegian language in America; a study in bilingual behavior* (Publications of the American Institute, University of Oslo, in Cooperation with the Department of American Civilization, Graduage School of Arts and Sciences, University of Pennsylvania). Bloomington: Indiana University Press.
Hervé, Maxime. 2018. *RVAideMemoire: Testing and Plotting Procedures for Biostatistics*. https://CRAN.R-project.org/package=RVAideMemoire.
Hockett, Charles F. 1966. What Algonquian Is Really like. *International Journal of American Linguistics* 32(1). 59–73.
Hogmen, Wes. 1981. Agreement for animacy and gender in the Buffalo Narrows dialect of French/Cree. *MASA: Journal of the University of Manitoba Anthropology Students' Association* 7. 81–94.
Ibrahim, Muhammad Hasan. 1973. *Grammatical gender: Its origin and development* (Janua Linguarum.). Vol. 166. The Hague: Mouton.
Kibort, Anna & Greville G. Corbett. 2008. *Grammatical Features Inventory: Gender*. University of Surrey. http://dx.doi.org/10.15126/SMG.18/1.01 (13 September, 2017).
Kilarski, Marcin. 1997. The case of the unmarked gender in loanwords. *Studia Angliga Posnaniensia* 32. 89–99.
Kilarski, Marcin. 2007. Algonquian and Indo-European gender in a historiographic perspective. *Historiographia Linguistica* 34. 333–349.
Kitaoka, Daiho & Kathleen Strader. 2016a. The structure of Michif relative clauses.
Kitaoka, Daiho & Kathleen Strader. 2016b. Relative clauses in Michif. Presented at the Annual Conference of the Canadian Linguistics Association, University of Calgary, Calgary, Alberta.

Kitaoka, Daiho & Kathleen Strader. 2017a. Discontinuous quanitifers in Michif.
Kitaoka, Daiho & Kathleen Strader. 2017b. Discontinuing objects in Michif. Presented at the Annual Conference of the Canadian Linguistics Association, Ryerson University, Toronto, Ontario.
Kitaoka, Daiho & Kathleen Strader. 2017c. Discontinuous DPs in Michif: preverbal focus position. Presented at the 47th Algonquian Conference, University of Manitoba, Winnipeg MB.
Klapka, Lyne. 2002. *Etude comparative: L'accord du genre en français québecois au 19ième et au 20ième siècle*. Ottawa: University of Ottawa.
Labov, William. 1996. When intuitions fail. In Lisa McNair, Kora Singer, Lise M. Dobrin & Michelle M. Aucon (eds.), *Papers from the Parasession on Theory and Data in Linguistics: April 11-13, 1996*, vol. 32. Chicago: Chicago Linguistic Society.
Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press. https://www.press.uchicago.edu/ucp/books/book/chicago/W/bo3632089.html (30 July, 2018).
Laverdure, Patline & Ida Rose Allard. 1983. *The Michif dictionary. Turtle Mountain Chippewa Cree*. (Ed.) John C. Crawford. Winnipeg, MB: Pemmican Publications.
Ledoux-Zoldy, Grace. 2010. *Li liiv oche Michif ayamiiawina: The book of Michif prayers*. 2nd edn. Winnipeg: Louis Riel Institute.
Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the world*. 17th edn. Dallas, Texas: SIL International.
L'Hirondelle, Lorna, Marjorie Memnook, Sally Warr & Donna Paskemin (eds.). 2001. *Plains Cree grammar guide and glossary*. Edmonton, Alberta: Faculty of Native Studies, University of Alberta.
Louis Riel Institute. 2013. Michif Language Examples "The Lost Dog" - YouTube. https://www.youtube.com/watch?v=jvSnGEwKgUk (23 July, 2018).
Lovell, Larry. 1984. *The structure of reflexive clauses in Michif: A relational grammar approach*. Grand Forks: University of North Dakota Master's thesis.
Lupu, Mihaela. 2005. La masculinisation du lexique français: le rôle catalyseur des anglicismes. *Analele Universităţii „Alexandru Ioan Cuza" din Iaşi. Secţiunea IIIe. Lingvistică, LI*, 261–271. Iaşi: Editura Universităţii „Alexandru Ioan Cuza". http://www.diacronia.ro/en/indexing/details/A2697 (12 May, 2018).
Macaulay, Monica. 2009. On prominence hierarchies: Evidence from Algonquian. *Linguistic Typology* 13. 357–389. doi:10.1515/LITY.2009.019.
Martinez, Michal Temkin. 2014. Using the Livescribe Echo Smartpen for Language Documentation. http://scholarspace.manoa.hawaii.edu/handle/10125/24604 (16 November, 2017).
Mathieu, Éric & Kathleen Strader. 2015. On gender in Michif. Presented at the 47th Algonquian Conference, University of Manitoba, Winnipeg MB.
Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
Matras, Yaron & Peter Bakker (eds.). 2003. *The Mixed Language Debate, Theoretical and Empirical Advances*. Berlin/Boston: De Gruyter Mouton. https://www.degruyter.com/view/product/36460 (15 December, 2017).

Matthews, Peter H. 2014. *The Concise Oxford Dictionary of Linguistics*. Oxford University Press. http://www.oxfordreference.com.login.ezproxy.library.ualberta.ca/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128 (1 August, 2018).
Max Planck Institute for Psycholinguistics. 2016. *ELAN*. Nijmegen: The Language Archive. https://tla.mpi.nl/tools/tla-tools/elan/ (22 February, 2016).
Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Books for Young Readers.
Mazzoli, Maria. in press. Challenges and opportunities of collaborative language research: The Michif case study. In Kerstin Knopf, Eeva Sippola, Marivic Lesho & Janelle Rodriques (eds.), *Postcolonial Knowledges, Proceedings of the Third Bremen Conference on Language and Literature in Colonial and Postcolonial Contexts*. Heidelberg: Heidelberg University Press.
McCreery, Dale. 2012. piikishkweetaak: Time to learn! *piikishkweetaak*. http://muskwatch.blogspot.com/2012/03/time-to-learn.html (26 July, 2018).
McCreery, Dale. 2013. *Challenges and Solutions in Adult Acquisition of Cree as a Second Language*. Master's thesis. (26 July, 2018).
McDonough, Joyce & Jordan Lachler. 2012. *The toy game: A technique for collecting natural conversation and prosody in the field*. Rochester: University of Rochester.
Meakins, Felicity. 2013. Mixed languages. In Yaron Matras & Peter Bakker (eds.), *Contact languages: A comprehensive guide*, 159–228. Berlin: Mouton de Gruyter.
Meakins, Felicity. 2017. Mixed Languages. In Mark Aronoff (ed.), *Oxford Research Encyclopedias: Literature*, 1–29. Oxford: Oxford University Press. (11 May, 2017).
Mel'čuk, Igor. 1974. Statistics and the Relationship Between the Gender of French Nouns and Their Endings. In V.Ju. Rozencvejg (ed.), *Machine Translation and Applied Linguistics. Vols. I and II.* (Soviet Papers in Formal Linguistics: 17), 11–42. Frankfurt am Main: Athenaion.
Mel'čuk, Igor. 2008. *Aspects of the Theory of Morphology*. (Ed.) David Beck. Berlin/Boston: De Gruyter Mouton. https://www.degruyter.com/view/product/36398 (11 May, 2018).
Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). 2013. *Atlas of Pidgin and Creole Language Structures Online*. Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology. http://apics-online.info/.
Millar, Patsy, Doris Mikolayenko, Agathe Chartrand, June Bruce & Lorraine Coutu. 2016. *Michif French: as spoken by most Michif people of St. Laurent*.
Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.
Mithun, Marianne. 2007. Linguistics in the face of language endangerment. In W. Leo Wetzels (ed.), *Language endangerment and endangered languages: Linguistic and anthropological studies with special emphasis on the languages and cultures of the Andean-Amazonian border area*, 15–35. Leiden: Publications of the Research School of Asian, African, and Amerindian Studies (CNWS), Leiden University.
Mithun, Marianne. 2014. The value of good conversation. Presented at the Annual meeting of the Linguistic Society of America, Minneapolis, MN.

http://www.linguisticsociety.org/sites/default/files/Mithun-Value_of_Good_Conversation.pdf (2 January, 2014).
Mühlbauer, Jeffrey. 2007. The speaker's knowledge and obviation in Plains Cree. In H.C. Wolfart (ed.), *Proceedings of the 38th Algonquian Conference*, vol. 38, 343–389. Winnipeg, Manitoba, Canada: University of Manitoba Press.
Mühlbauer, Jeffrey Thomas. 2008. *kâ-yôskâtahk ôma nêhiyawêwin: The representation of intentionality in Plains Cree*. Vancouver: University of British Columbia Unpublished doctoral dissertation.
Muysken, Pieter C. 1981. Half-way between Quechua and Spanish: The case for relexification. In Arnold R. Highfield & Albert Valdman (eds.), *Historicity and Variation in Creole Studies*, 52–79. Ann Arbor, Michigan: Karoma.
Nymannson, Karin. 1995. Le genre grammatical des anglicismes contemporains en français. *Cahiers de Lexicologie* 66. 95–113.
Okimāsis, Jean L. 2004. *Cree: Language of the Plains. nēhiyawēwin: paskwāwi-pīkiskwēwin*. Regina, Saskatchewan: Canadian Research Center, University of Regina.
OLAC. 2018. Open Language Archives Community. http://www.language-archives.org/ (15 May, 2018).
Orser, Lori L. 1984. *Michif: A problem in classification*. Lawrence: University of Kansas Master's thesis.
Őrsi, Tibor. 2012. Gender assignment of anglicisms in French. *Eger Journal of English Studies* XII. 77–83.
Oxford, William Robert. 2014. *Microparameters of agreement: A diachronic perspective on Algonquian verb inflection*. Toronto, Ontario, Canada: University of Toronto PhD thesis.
Panas, J.D. & Olive Whitford. 2004. *The beavers' big house: Lii kastorr leu groos maenzoon*. (Trans.) Norman Fleury. Saskatoon: Gabriel Dumont Institute.
Papen, Robert. 2004. Michif Spelling Conventions: Proposal for a Unified Michif Writing Systems. In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. Two: Language Theory, 29–53. Pemmican Publications.
Papen, Robert A. 1984a. Quelques remarques sur un parler français méconnu de l'Ouest canadien : le métis. *Revue québecoise de linguistique* 14(1). 113–39. doi:10.7202/602530ar.
Papen, Robert A. 1984b. Un parler français méconnu de l'Ouest Canadien: le métis.'Quand même qu'on parle français, ça veut pas dire qu'on est des Canayens!'. In Pierre-Yves Mocquais, André Lalonde & Bernard Wilhelm (eds.), *La langue, la culture, et la société des francophones de l'Ouest*, 121–136. Regina: Institut de recherche du centre d'études bilingues.
Papen, Robert A. 1986. Can two distinct grammars coexist in a single language? The case of Métif. In A. Murray Kinloch (ed.), *Papers from the tenth annual meeting of the Atlantic Provinces Linguistic Association / Actes du dixième colloque annuel de l'Association de Linguistique des Provinces Atlantiques*, 144–157. Fredericton, NB: University of New Brunswick Press.

Papen, Robert A. 1987a. Linguistic variation in the French component of Métif grammar. In William Cowan (ed.), *Papers of the 18th Algonquian Conference*, 247–259. Ottawa: Carleton University.
Papen, Robert A. 1987b. Sur quelques processus phonologiques, morphologiques et lexicaux du Métif. In Rose Mary Babitch (ed.), *Papers from the eleventh annual meeting of the Atlantic Provinces Linguistic Association / Actes du onzième colloque annuel de l'Association de Linguistique des Provinces Atlantiques*, 107–114. Shippagan, NB: University of Moncton.
Papen, Robert A. 1993. La Variation dialectale dans le parler français des Métis de l'Ouest canadien. *Francophonies d'Amérique* (3). 25–38. doi:10.7202/1004439ar.
Papen, Robert A. 2003a. Le mitchif: un problème de genre. In Robert A. Stebbins, Claude Romney & Micheline Ouellet (eds.), *Francophonie et langue dans un monde divers en évolution: contacts interlinguistiques et socioculturels. Actes du dix-neuvième colloque du CEFCO (18–20 octobre 2001)*, 119–141. Winnipeg, MB: Presses universitaires de Saint-Boniface.
Papen, Robert A. 2003b. Michif: One phonology or two? In Yunhee Chung, Carrie Gillon & Rachel Wojdak (eds.), *Papers from the workshop on the structure and constituency in the languages of the Americas*, 47–58. Vancouver: University of British Columbia Working Papers in Linguistics.
Papen, Robert A. 2005a. On developing a writing system for Michif. *Linguistica Atlantica* 26. 75–97.
Papen, Robert A. 2005b. Le Mitchif: langue franco-crie des Plaines. In Albert Valdman, Julie Auger & Deborah Piston-Hatlen (eds.), *Le Français en Amérique du Nord*, 327–347. Saint-François, QC: Presses de l'Université Laval.
Papen, Robert A. 2011. Un nours, un zours, un lours? la question de la liaison en mitchif. In France Martineau & Terry Nadasdi (eds.), *Le français en contact: Hommages à Raymond Mougeon*, 217–245. Saint-François, QC: Presses de l'Université Laval.
Papen, Robert A. 2017. On confirming the split phonology hypothesis for Michif. Presented at the 49th Algonquian Conference, Université du Québec à Montréal.
Papen, Robert A. & Davy Bigot. 2010. Sontaient, ontvaient, et fontsaient en Français Mitchif: Variation et systematicité. In Carmen Leblanc, France Martineau & Yves Frenette (eds.), *Vues sur les français d'ici*, 201–225. Québec: Presses de l'Université Laval.
Pasch, Timothy. 2013. DJing language: "Elastic" audio for clip-based language documentation. Poster presented at the 3rd International Conference on Language Documentation and Conservation, Honolulu, HI.
Patton, Anne & Wilfred Burton. 2007. *Fiddle dancer: Li danseur di vyaeloon*. (Trans.) Norman Fleury. Saskatoon: Gabriel Dumont Institute.
Pelletier, Jeanne. 2007. *The story of the rabbit dance: Li nistwaar di la daañs di liyév*. (Trans.) Rita Flamand. Saskatoon: Gabriel Dumont Institute.
Peske, Mary. 1981. *The French of the French-Cree (Michif) language*. Grand Forks, North Dakota: University of North Dakota Master's thesis.
Peterson, Jacqueline. 1978. Prelude to Red River: A social portrait of the Great Lakes Métis. *Ethnohistory* 25. 41–67.

Poplack, Shana, Alicia Pousada & David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 57(1). 1–28. doi:10.1016/0024-3841(82)90068-7.
Prichard, Hilary & Kobey Shwayder. 2014. Against a Split Phonology of Michif. *University of Pennsylvania Working Papers in Linguistics* 20(1, Article 29). 270–280.
Pyle, Kai. 2018. The Michif Internet Resource Center. *The Michif Internet Resource Center*. https://michif.wordpress.com/ (8 January, 2018).
Quinn, Conor. 2001. A preliminary survey of animacy categories in Penobscot. In John D. Nichols (ed.), *Papers of the 32nd Algonquian Conference*, 395–426. Winnipeg: University of Manitoba.
Ratt, Solomon. 2016. *mâci-nêhiyawêwin: Beginning Cree*. Regina, Saskatchewan: University of Regina Press.
Rhodes, Richard. 1982. Algonquian trade languages. (Ed.) William Cowan. *Papers of the 13th Algonquian Conference* 13. 1–10.
Rhodes, Richard A. 1977. French Cree: A case of borrowing. In William Cowan (ed.), *Actes du Huitième Congrès des Algonquinistes*, 6–25. Ottawa: Carleton University.
Rhodes, Richard A. 1986. Métchif: A second look. In William Cowan (ed.), *Actes du Dix-Septième Congrès des Algonquinistes*, 287–296. Ottawa: Carleton University.
Rhodes, Richard A. 1987. Les conte Métif – Métif myths. *Papers of the 19th Algonquian Conference* 19. 297–301.
Rhodes, Richard A. 2001. Text strategies in Métchif. In H.C. Wolfart (ed.), *Papers of the Thirty-second Algonquian Conference*, 455–469. Winnipeg: University of Manitoba.
Rhodes, Richard A. 2008. Ojibwe in the Cree of Métchif. (Ed.) Karl S. Hele & Regna Darnell. *Papers of the 39th Algonquian Conference* 39. 569–580.
Rhodes, Richard A. 2009. The phonological history of Métchif. In France Martineau & Luc Baronian (eds.), *Le français d'un continent à l'autre: Mélanges offerts à Yves Charles Morin* (Les Voies Du Français 1), 423–442. Québec: Presses de l'Université Laval.
Rhodes, Richard A. 2013. An overview of Métchif adjectives. *Canadian Journal of Native Studies* 33(2). 101–115.
Robinson, Sinclair & Donald Smith. 1990. *Dictionary of Canadian French = Dictionnaire du français canadien*. Toronto: Stoddart.
Rosen, Nicole. 2003. Demonstrative Position in Michif. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 48(1–2). 39–69. doi:10.1353/cjl.2004.0011.
Rosen, Nicole. 2004a. Recommendations for best practices for the development of Michif language instruction for adults. In Lawrence J. Barkwell (ed.), *La lawng: Michif peekishkwewin: The heritage language of the Canadian Metis*, vol. 1, 67–81. Winnipeg: Pemmican Publications.
Rosen, Nicole. 2004b. Stress Assignment in Michif. In Lawrence Barkwell (ed.), *La Lawng: Michif Peekishkwewin: The Heritage Language of the Canadian Metis*, vol. Two: Language Theory, 55–61. Pemmican Publications.

Rosen, Nicole. 2006. Language contact and Michif stress assignment. *STUF - Sprachtypologie und Universalienforschung* 59(2). 170–190.
Rosen, Nicole. 2007. *Domains in Michif phonology*. Toronto: University of Toronto PhD thesis.
Rosen, Nicole. 2008. French-Algonquian interaction in Canada: A Michif case study. *Clinical Linguistics & Phonetics* 22(8). 610–624. doi:10.1080/02699200802221570.
Rosen, Nicole. 2016a. Multilingual language revitalization: The case of the Red River Métis. Washington, D.C.
Rosen, Nicole. 2016b. Michif Online Dictionary. *Michif Online Dictionary*. http://michif.atlas-ling.ca/ (10 February, 2016).
Rosen, Nicole & Janelle Brodner. 2012. Michif vowel space. Presented at the Society for the Study of the Languages of the Americas Annual Meeting, Portland, OR.
Rosen, Nicole & Carrie Gillon. 2015. Reanalyzing Michif determiners. Presented at the Canadian Linguistics Association, University of Ottawa.
Rosen, Nicole & Carrie Gillon. 2017. An examination of Michif gender. Presented at the 49th Algonquian Conference, Université du Québec à Montréal.
Rosen, Nicole & Élyane Lacasse. 2014. Une comparaison des voyelles postérieures du français mitchif et du franco-manitobain. In Robert A. Papen & Sandrine Hallion (eds.), *À l'ouest des Grands Lacs : communautés francophones et variétés de français dans les Prairies et en Colombie-Britannique*, 109–153. Québec: Presses de l'Université Laval.
Rosen, Nicole, Jeff Muehlbauer & Élyane Lacasse. 2010a. L'espace des voyelles postérieures en michif, français et cri des Plaines. Presented at the Les français d'ici, Université de Montréal.
Rosen, Nicole, Jeff Muehlbauer & Élyane Lacasse. 2010b. Back vowel space in Michif and its lexifying languages. Presented at the 14th Alberta Conference on Linguistics, Calgary, Alberta.
Rosen, Nicole & Heather Souter. 2009a. *Piikishkweetak añ Michif!* Winnipeg, Manitoba: Louis Riel Institute.
Rosen, Nicole & Heather Souter. 2009b. Language revitalization in a multilingual community: The case of Michif. Honolulu. http://hdl.handle.net/10125/5090.
Rosen, Nicole & Heather Souter. 2015. *Piikishkweetak aa'n Michif!* Winnipeg, Manitoba: Louis Riel Institute.
Rosen, Nicole, Jesse Stewart & Olivia N. Sammons. in prep. Diachronic effects of language contact in the synchronic Michif vowel system.
Rosen, Nicole, Jesse Stewart & Olivia N. Sammons. 2015. Phonetics of the synchronic Michif vowel system. Presented at the 47th Algonquian Conference, University of Manitoba, Winnipeg MB.
Rosen, Nicole, Jesse Stewart & Olivia N. Sammons. 2016a. Diachronic effects of language contact in the synchronic Michif vowel system. Presented at the 90th Annual Meeting of the Linguistic Society of America, Washington, D.C.
Rosen, Nicole, Jesse Stewart & Olivia N. Sammons. 2016b. Diachronic effects of language contact in the synchronic Michif vowel system. Presented at the New Ways of Analyzing Variation 45, Vancouver, BC.

Rosenblum, Daisy & Olivia N. Sammons. 2014. Documenting multimodal interaction: Workflows, data management, and archiving. Presented at the Annual meeting of the Linguistic Society of America, Minneapolis, MN. http://www.linguisticsociety.org/sites/default/files/Mithun-Value_of_Good_Conversation.pdf (2 January, 2014).
Ross, Malcolm. 2007. Calquing and Metatypy. *Journal of Language Contact* 1(1). 116–143. doi:10.1163/000000007792548341.
Russell, Kevin. 1991. Obviation as discourse structure in a Swampy Cree âcimowin. In William Cowan (ed.), *Papers of the 22nd Algonquian Conference*, vol. 22, 320–335. Ottawa, Ontario, Canada: Carleton University.
Sakel, Jeanette & Daniel L. Everett. 2012. *Linguistic Fieldwork*. Cambridge: Cambridge University Press.
Sammons, Olivia N. in revision. Applicative constructions in Northern Michif.
Sammons, Olivia N. in revision. Enriching the record: The role of conversation in language documentation.
Sammons, Olivia N. 2013a. Leaving Ste. Madeleine: A Michif account. *Canadian Journal of Native Studies* 33(2). 149–164.
Sammons, Olivia N. 2013b. Collaboration, communities, and distance. Poster presented at the Annual meeting of the Linguistic Society of America, Boston.
Sammons, Olivia N. 2013c. Linguistic variation and language revitalization: Perspectives from Michif. Poster presented at the Annual meeting of the Canadian Linguistic Association, Victoria, BC.
Sammons, Olivia N. 2013d. Inflectional morphology in contact: Michif possessive constructions. Presented at the Language Documentation and Linguistic Theory, London.
Sammons, Olivia N. 2015. Enriching the record: The role of conversation in language documentation. Presented at the Language in the Present, Victoria, BC.
Sammons, Olivia N. 2017. Nominal classification in Michif. Presented at the 49th Algonquian Conference, Université du Québec à Montréal.
Saugera, Valérie. 2017. *Remade in France: Anglicisms in the Lexicon and Morphology of French*. Oxford University Press.
Sealey, D. Bruce & Antoine S. Lussier. 1975. *The Métis: Canada's forgotten people*. Winnipeg: Pemmican Publications.
Séguin, Hubert. 1969. *Les marques du genre dans le lexique du français écrit contemporain compilation des cas et essai de classement*. Montréal: Université de Montréal.
Siewierska, Anna. 2013. Alignment of Verbal Person Marking. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/100 (14 July, 2018).
SIL International. *SayMore: Language documentation productivity*. http://saymore.palaso.org/ (22 February, 2016).
Simons, Gary F. & Charles D. Fennig. 2018. Language Status. In Gary F. Simons & Charles D. Fennig (eds.), *Ethnologue: Languages of the World*. 21st edn. Dallas, Texas: SIL International. Online version: https://www.ethnologue.com/about/language-status (23 March, 2018).

Souter, Heather. 2008. Michif Verb Rummy. http://www.metismuseum.ca/browse/index.php/987 (26 July, 2018).
Souter, Heather. 2018a. *Ti parii chiiñ? (Are you ready?): Preparing adult learners and proficient speakers for the challenge of Michif reclamation*. Victoria: University of Victoria Master's thesis.
Souter, Heather. 2018b. *A Guide to Ti Parii Chiiñ: Michif Master-Apprentice Phrases and Micro Grammar Lessons*. https://www.youtube.com/watch?v=-qwd1QjcLiY&t=0s&list=PLI2E8ojtLxiZnsi1-V67w_NJcoJDLGS7c&index=2 (26 July, 2018).
Speers, Breck. 1983. *Discourse analysis of a Michif narrative*. Grand Forks: University of North Dakota Master's thesis.
St. Pierre, Edwin. 2012. *Remembering my Métis past: Reminiscences of Edwin St. Pierre*. (Trans.) Harriet Oaks St. Pierre. Saskatoon: Gabriel Dumont Institute.
Stewart, Jesse. 2015. Intonation Patterns in Pijal Media Lengua. *Journal of Language Contact* 8(2). 223–262. doi:10.1163/19552629-00802003.
Stoltzfus, Daniel-Paul & Aubrée Boissard. 2016. Lexical patterning in Michif: reconsidering the role of French. Presented at the 48th Algonquian Conference, University of Wisconsin--Milwaukee.
Strader, Kathleen. 2013. Michif Discontinuous DP elements. Presented at the 45th Algonquian Conference, University of Ottawa.
Strader, Kathleen. 2014. *Michif Determiner Phrases*. Winnipeg, Manitoba, Canada: University of Manitoba Master's thesis.
Strader, Kathleen. 2016. Agreement in Michif adjectives. Presented at the The 21st Workshop on Structure and Constituency in Languages of the Americas, Montreal: Université du Québec à Montréal.
Strader, Kathleen. 2017a. Michif Adjectives. Ottawa, Ontario, ms.
Strader, Kathleen. 2017b. Michif: The Unique Syntax Hypothesis. Ottawa, Ontario, ms.
Strader, Kathleen. 2017c. The unique nominal syntax of Michif. Presented at the Formal Approaches to Creole Studies, University of Kentucky, Lexington.
Straus, Anne Terry & Robert A. Brightman. 1982. The implacable raspberry. By Anne Terry Straus and Robert Brightman. *Papers in Linguistics* 15(1–4). 97–137.
Thieberger, Nicholas & Andrea L. Berez. 2012. Linguistic data management. In Nicholas Thieberger (ed.), *Oxford handbook of linguistic fieldwork*, 90–120. New York: Oxford University Press.
Thomason, Lucy Grey. 2003. *The Proximate and Obviative Contrast in Meskwaki*. Austin, Texas, United States: The University of Texas at Austin PhD thesis.
Thomason, Sarah G. 2001. *Language Contact: An Introduction*. Washington, D.C., United States: Georgetown University Press.
Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Totem Field Storyboards. 2013. Storyboards for Language Documentation. http://totemfieldstoryboards.org/ (20 November, 2017).
Tucker, G. Richard, Wallace E. Lambert & André Rigault. 1977. *The French speaker's skill with grammatical gender: an example of rule-governed behavior*. The Hague: Mouton.

Weaver, Deborah. 1982. *Obviation in Michif*. Grand Forks, North Dakota, United States: University of North Dakota Master's thesis.
Weaver, Deborah. 1983. The effects of language change and death on obviation in Michif. In William Cowan (ed.), *Actes du Quatorzième Congrès des Algonquinistes*, vol. 14, 261–268. Ottawa, Ontario, Canada: Carleton University.
Weston, Lori Orser. 1982. Alternative structures in a mixed language: Benefactives and dubitatives in Michif. *Papers of the Mid-America Linguistics Conference* 427–432.
Wildeman, Carol. 1989. *The Michif technique: Code-switch cue*. Grand Forks: University of North Dakota Master's thesis.
Winford, Donald. 2008. Processes of Creole Formation and Related Contact-Induced Language Change. *Journal of Language Contact* 2(1). 124–145. doi:10.1163/000000008792525228.
Wolfart, H. C. 2010. Choice and balance in Michif negation. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 55(1). 115–129. doi:10.1353/cjl.0.0068.
Wolfart, H. Christoph. 1973. *Plains Cree: A grammatical study* (Transactions of the American Philosophical Society, New Series). Vol. 63, part 2. Philadelphia, Pennsylvania: American Philosophical Society.
Wolfart, H.C. 1978. How many obviatives: Sense and reference in a Cree verb paradigm. In Eung-Do Cook & Jonathan D. Kaye (eds.), *Linguistic studies of Native Canada*, 255–272. Lisse, Netherlands / Vancouver, British Columbia, Canada: Peter de Ridder Press / University of British Columbia Press.
Wolfart, H.C. 1996. Sketch of Cree, an Algonquian language. In Ives Goddard (ed.), *Handbook of North American Languages*, vol. 17: Languages, 390–439. Washington, D.C., United States: Smithsonian Institution.
Wolvengrey, Arok. 2001. *nēhiýawēwin: itwēwina. Cree: Words*. 2 vols. Regina, Saskatchewan: Canadian Research Center, University of Regina.
Wolvengrey, Arok. 2011a. *Semantic and Pragmatic Functions in Plains Cree Syntax* (LOT Publications 268). Utrech, Netherlands: LOT.
Wolvengrey, Arok. 2011b. *Cree: Words*. Regina: University of Regina Press. (27 July, 2018).
Wolvengrey, Arok (ed.). 2015. Plains Cree Dictionary. http://dictionary.plainscree.atlas-ling.ca/#/help (18 May, 2018).
Woodbury, Tony. 2003. Defining documentary linguistics. (Ed.) Peter K. Austin. *Language documentation and description* 1. 35–51.
2018a. Algonquian Linguistic Atlas. http://atlas-ling.ca/ (27 March, 2018).
2018b. English to French, Italian, German & Spanish Dictionary - WordReference.com. http://www.wordreference.com/ (18 May, 2018).

Appendix 1: Michif lemma frequencies in the dataset

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
4X4$_{MASC:AN}$ '4 X 4'	ENGLISH	1	1	0	1	0	camion$_{MASC}$	âwatâswâkan$_{AN}$
AMBULANCE$_{MASC:AN}$ 'ambulance'	ENGLISH	2	2	0	2	0	ambulance$_{FEM}$	ahkosiwtapan$_{IN}$
ARAÑZH$_{FEM:AN}$ 'orange'	FRENCH	1	0	1	1	0	orange$_{FEM}$	osâwâs$_{AN}$
BAA$_{MASC:IN}$ 'sock'	FRENCH	1	1	0	0	1	bas$_{MASC}$	asikan$_{AN}$
BALL$_{MASC:AN}$ 'ball'	ENGLISH	1	1	0	1	0	balle$_{FEM}$	pâkahatowân$_{AN}$
BAÑ$_{MASC:IN}$ 'bench'	FRENCH	1	1	0	0	1	banc$_{MASC}$	tehtapowinis$_{IN}$
BARLII$_{MASC:IN}$ 'barrel'	ENGLISH	1	1	0	0	1	tonneau$_{MASC}$	mahkahk$_{IN}$
BARYER$_{VAR:VAR}$ 'barrier; gate'	FRENCH	4	2	2	1	3	barrière$_{FEM}$	nakânikan$_{IN}$
BASKET$_{MASC:IN}$ 'basket'	ENGLISH	9	9	0	1	8	panier$_{MASC}$	mawiswâkan$_{IN}$; watapîwat$_{IN}$
BEEBII$_{MASC:AN}$ 'baby'	FRENCH	7	7	0	7	0	bébé$_{MASC}$	pêpîsis$_{AN}$
BELMER$_{FEM:AN}$ 'mother-in-law'	FRENCH	1	0	1	1	0	belle-mère$_{FEM}$	-sikos-$_{AN}$
BICH_LAÑP$_{MASC:IN}$ 'bitumen lamp'	ENGLISH	1	1	0	0	1	lampe$_{FEM}$ 'lamp'	wâsaskotênikan$_{IN}$ 'lamp'
BICYCLE$_{MASC:AN}$ 'bicycle'	ENGLISH	14	14	0	12	2	bicycle$_{MASC}$,[44] vélo$_{MASC}$	nîsokâcis$_{IN}$
BIKE$_{MASC:AN}$ 'bicycle'	ENGLISH	1	1	0	1	0	bicycle$_{MASC}$, vélo$_{MASC}$	nîsokâcis$_{IN}$
BILEE$_{MASC:AN}$ 'billy goat'	ENGLISH	7	7	0	7	0	bouc$_{MASC}$	wâpatihk$_{AN}$
BILLY_GOAT$_{MASC:AN}$ 'billy goat'	ENGLISH	4	4	0	4	0	bouc$_{MASC}$	wâpatihk$_{AN}$
BISH$_{FEM:AN}$ 'elk'	FRENCH	1	0	1	1	0	biche$_{FEM}$	wâwâskêsiw$_{AN}$
BITAEÑ$_{MASC:IN}$ 'clothes'	FRENCH	18	18	0	1	17	butin$_{MASC}$	wiyâhcikana$_{IN}$

44 This form is noted to be frequently used in colloquial Québecois French (Robinson & Smith 1990: 160).

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
BLAAÑ$_{MASC:AN}$ ‘white (man)’	FRENCH	1	1	0	1	0	blanc$_{MASC}$	moniyâw$_{AN}$
BOCHIN$_{FEM:IN}$ ‘boot’	FRENCH	1	0	1	0	1	bottine$_{FEM}$	mistikwaskisin$_{IN}$
BODAEÑ$_{MASC:IN}$ ‘blood sausage’	FRENCH	2	2	0	0	2	boudin noir$_{MASC}$	otakisîhkân$_{IN}$ ‘sausage’
BOER$_{MASC:IN}$ ‘butter’	FRENCH	2	2	0	0	2	beurre$_{MASC}$	tohtôsâpôwipimiy$_{IN}$
BONEE$_{MASC:IN}$ ‘hat’	FRENCH	11	11	0	0	11	bonnet$_{MASC}$	astotin$_{IN}$
BOÑFRER$_{MASC:AN}$ ‘brother-in-law’	FRENCH	2	2	0	2	0	beau-frère$_{MASC}$	-îstâw$_{AN}$
BOÑPER$_{MASC:AN}$ ‘father-in-law’	FRENCH	1	1	0	1	0	beau-père$_{MASC}$	-sis$_{AN}$
BOTEE$_{FEM:IN}$ ‘bottle’	FRENCH	5	0	5	0	5	bouteille$_{FEM}$	môtêyâpisk$_{IN}$
BOWL$_{MASC:IN}$ ‘bowl’	ENGLISH	1	1	0	0	1	bol$_{MASC}$	wiyâkan$_{IN}$
BOYFRIEND$_{MASC:AN}$ ‘boyfriend’	ENGLISH	1	1	0	1	0	copain$_{MASC}$	-îcimos$_{AN}$
BRAA$_{MASC:IN}$ ‘arm’	FRENCH	4	4	0	0	4	bras$_{MASC}$	mispiton$_{IN}$
BRAÑSH$_{FEM:IN}$ ‘branch’	FRENCH	1	0	1	0	1	branche$_{FEM}$	watihkwan$_{IN}$
BREZ$_{FEM:IN}$ ‘ember(s)’	FRENCH	1	0	1	0	1	braise$_{FEM}$	kaskaskisiw$_{IN}$
BROSH$_{FEM:IN}$ ‘brush’	FRENCH	1	0	1	0	1	brosse$_{FEM}$	sinikohtakahikan$_{IN}$
BWAA_DI_KOR$_{MASC:IN}$ ‘cordwood’	FRENCH	2	2	0	0	2	bois de corde$_{MASC}$	asastân$_{IN}$
BWAA-1$_{MASC:IN}$ ‘stick’	FRENCH	1	1	0	0	1	bout de bois$_{MASC}$	mistik$_{IN}$
BWAA-2$_{MASC:IN}$ ‘wood’	FRENCH	5	5	0	0	5	bois$_{MASC}$	mihti$_{IN}$
BWASOÑ$_{MASC:IN}$ ‘alcohol’	FRENCH	1	1	0	0	1	boisson$_{FEM}$	iskotêwâpoy$_{IN}$
BWATOÑ$_{MASC:AN}$ ‘button’	FRENCH	1	1	0	1	0	bouton$_{MASC}$	sakwâskwahon$_{AN}$
BWET$_{FEM:IN}$ ‘box’	FRENCH	7	0	7	0	7	boîte$_{FEM}$	mistikowat$_{IN}$
BYER$_{FEM:IN}$ ‘beer’	FRENCH	3	0	3	0	3	bière$_{FEM}$	iskwêsisâpoy$_{IN}$
CAKE$_{MASC:AN}$ ‘cake’	ENGLISH	1	1	0	1	0	gateau$_{MASC}$	wîhkihkasikan$_{AN}$
CHIM_DI_ZHVOO$_{MASC:AN}$ ‘team of horses’	FRENCH	3	3	0	3	0	time$_{FEM}$	peyakwahpitew misitatimwak$_{AN}$
COOKSTOVE$_{MASC:AN}$ ‘cookstove’	ENGLISH	1	1	0	1	0	poêle$_{MASC}$	kotawânâpisk$_{IN}$
COWBOY_HAT$_{MASC:IN}$ ‘cowboy hat’	ENGLISH	3	3	0	0	3	chapeau de cow-boy$_{MASC}$	astotin$_{IN}$ ‘hat’
DAÑS$_{FEM:IN}$ ‘dance’	FRENCH	1	0	1	0	1	danse$_{FEM}$	nîmihitowin$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
DEEZHEUNII$_{MASC:IN}$ ‘breakfast’	FRENCH	1	1	0	0	1	déjeuner$_{MASC}$	kîkisêpâwahcikêwin$_{IN}$
DIBRII$_{MASC:IN}$ ‘tripe’	FRENCH	10	10	0	0	10	débris$_{MASC}$	wînâstakay-$_{IN}$
DRIVERS_LICENSE$_{MASC:IN}$ ‘driver’s license’	ENGLISH	1	1	0	0	1	permis de conduire$_{MASC}$	pamihchikew masinahikan$_{IN}$
DWE$_{MASC:IN}$ ‘finger’	FRENCH	2	2	0	0	2	doigt$_{MASC}$	yîkicihcân$_{IN}$
FAAM$_{FEM:AN}$ ‘woman’	FRENCH	17	0	17	17	0	femme$_{FEM}$	iskwêw$_{AN}$
FARIIN$_{FEM:AN}$ ‘flour’	FRENCH	4	0	4	4	0	farine$_{FEM}$	pahkwêsikan$_{AN}$
FARM$_{FEM:IN}$ ‘farm’	ENGLISH	2	0	2	0	2	ferme$_{FEM}$	kistikân$_{IN}$
FESTOÑ$_{MASC:IN}$ ‘feast (moose)’	FRENCH	2	2	0	0	2	festin$_{MASC}$	wîhkohtowin$_{IN}$
FEU$_{MASC:IN}$ ‘fire’	FRENCH	13	13	0	2	11	feu$_{MASC}$	iskotêw$_{IN}$
FII$_{FEM:AN}$ ‘girl’	FRENCH	45	2	43	45	0	fille$_{FEM}$	iskwêsis$_{AN}$
FIMEL$_{FEM:AN}$ ‘female’	FRENCH	1	0	1	1	0	femelle$_{FEM}$	nôsê-aya$_{AN}$
FOOR$_{MASC:IN}$ ‘oven’	FRENCH	2	2	0	0	2	four$_{MASC}$	sêkowêpinâpisk$_{IN}$
FRER$_{MASC:AN}$ ‘brother’	FRENCH	11	11	0	11	0	frère$_{MASC}$	-îtisân$_{AN}$
FUN$_{FEM:IN}$ ‘fun’	ENGLISH	1	0	1	0	1	fonne$_{MASC}$ / fun$_{MASC}$	miyawâtamowin$_{IN}$
FWAEÑ$_{MASC:IN}$ ‘hay; grass’	FRENCH	6	6	0	0	6	foin$_{MASC}$ ‘hay’	maskosiy$_{IN}$
GAA$_{MASC:AN}$ ‘guy’	FRENCH	34	34	0	34	0	gars$_{MASC}$	nâpêw$_{AN}$ ‘man’
GALET$_{FEM:AN}$ ‘bannock’	FRENCH	8	0	8	8	0	galette$_{FEM}$	pahkwêsikan$_{AN}$
GAME$_{FEM:IN}$ ‘game’	ENGLISH	1	0	1	0	1	jeu$_{MASC}$	mêtawêwin$_{IN}$
GARSOÑ$_{MASC:AN}$ ‘boy’	FRENCH	70	70	0	70	0	garcon$_{MASC}$	nâpês$_{AN}$
GLAS$_{FEM:AN}$ ‘ice’	FRENCH	2	0	2	2	0	glace$_{FEM}$	maskwamiy$_{AN}$
GLASS$_{MASC:IN}$ ‘glass’	ENGLISH	1	1	0	0	1	verre$_{MASC}$	minihkwêwiyâkanis$_{IN}$
GOURNOY$_{VAR:ANIM}$ ‘frog’	FRENCH	33	17	16	33	0	grenouille$_{FEM}$	ayîk$_{AN}$
GRAÑDRII$_{FEM:IN}$ ‘granary’	FRENCH	2	0	2	0	2	grenier$_{MASC}$	kistikânikamik$_{IN}$
GRAVEYARD$_{MASC:IN}$ ‘graveyard’	ENGLISH	1	1	0	0	1	cimetière$_{MASC}$	kihkwahaskân$_{IN}$
GRES$_{FEM:IN}$ ‘grease’	FRENCH	3	0	3	0	3	graisse$_{FEM}$	pimîs$_{IN}$
HAÑSH$_{FEM:IN}$ ‘hip’	FRENCH	1	0	1	0	1	hanche$_{FEM}$	mitokan$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
IDEE$_{FEM:IN}$ ‘idea’	FRENCH	1	0	1	0	1	idée$_{FEM}$	itêyihtâkan$_{IN}$
IIBUU$_{MASC:AN}$ ‘owl’	FRENCH	3	3	0	3	0	hibou$_{MASC}$	ôhow$_{AN}$
ISHEL$_{MASC:IN}$ ‘ladder’	FRENCH	3	3	0	0	3	echelle$_{FEM}$	kîhcêkosîwinâhtik$_{IN}$
JEANS$_{MASC:AN}$ ‘jeans’	ENGLISH	1	1	0	1	0	jean$_{MASC}$	iskwêwitâs$_{AN}$
JEEÑG$_{FEM:AN}$ ‘girl; girlfriend’	FRENCH	1	0	1	1	0	copine $_{FEM}$	-îcimos$_{AN}$
JINII$_{MASC:IN}$ ‘dinner; lunch’	FRENCH	1	1	0	0	1	déjeuner$_{MASC}$	nîmâwin$_{IN}$
KAAB$_{MASC:IN}$ ‘rope’	FRENCH	10	9	1	0	10	câble$_{MASC}$	pîminahkwân$_{IN}$
KAAN$_{FEM:IN}$ ‘can’	FRENCH	2	0	2	0	2	canne$_{FEM}$	kinwâskicês$_{IN}$
KACHIM$_{MASC:IN}$ ‘catechism’	FRENCH	1	1	0	0	1	catéchisme$_{MASC}$	ayamihâw kiskinohamâsiwin$_{IN}$
KANAAR$_{MASC:AN}$ ‘duck’	FRENCH	8	8	0	8	0	canard$_{MASC}$	sîsîp$_{AN}$
KAPOO$_{MASC:IN}$ ‘overcoat’	FRENCH	1	1	0	0	1	capot$_{MASC}$	miskotâkay$_{IN}$, waskitasâkay$_{IN}$
KARABIN$_{FEM:IN}$ ‘rifle’	FRENCH	1	0	1	0	1	carabine$_{FEM}$	sâpohtak$_{IN}$
KAREM$_{MASC:IN}$ ‘Lent’	FRENCH	2	2	0	0	2	carême$_{MASC}$	awihew$_{IN}$
KASKET$_{VAR:INAN}$ ‘cap’	FRENCH	5	2	3	0	5	casquette$_{FEM}$	astotin$_{IN}$
KATAEÑ$_{FEM:AN}$ ‘doll’	FRENCH	1	0	1	1	0	catin$_{FEM}$	awâsisîhkân$_{AN}$
KEU$_{VAR:INAN}$ ‘tail’	FRENCH	31	6	25	0	31	queue$_{FEM}$	-soy$_{IN}$
KEUV$_{FEM:IN}$ ‘tub’	FRENCH	1	0	1	0	1	cuve$_{FEM}$	mahkahk$_{IN}$
KLOSH$_{FEM:AN}$ ‘bell’	FRENCH	8	0	8	7	1	cloche$_{FEM}$	sêwêyâkan$_{AN}$
KOK$_{MASC:AN}$ ‘rooster’	FRENCH	22	21	1	22	0	coq$_{MASC}$	nâpê-pâhkahahkwân$_{AN}$
KOÑFICHEUR$_{MASC:IN}$ ‘preserves’	FRENCH	1	1	0	0	1	confiture$_{FEM}$	mînisâpôhkân$_{IN}$
KOP$_{FEM:AN}$ ‘penny’	FRENCH	3	0	3	3	0	cope$_{FEM}$	pîwâpiskos$_{AN}$
KOR$_{MASC:IN}$ ‘body’	FRENCH	2	2	0	0	2	corps$_{MASC}$	-iyaw$_{IN}$
KORAL$_{MASC:IN}$ ‘corral’	FRENCH	8	8	0	0	8	corral$_{MASC}$	wasakanikan$_{IN}$ (?)
KOSH$_{FEM:IN}$ ‘diaper’	FRENCH	1	0	1	0	1	couche$_{FEM}$	âsiyân$_{AN}$
KREM$_{FEM:IN}$ ‘cream’	FRENCH	1	0	1	0	1	crème$_{FEM}$	manahikan$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
KRIIYOÑ$_{MASC:IN}$ ‘pencil’	FRENCH	2	2	0	0	2	crayon$_{MASC}$	masinahikanâhcikos$_{AN}$
KROSH$_{MASC:AN}$ ‘crook’	FRENCH	1	1	0	1	0	croche$_{MASC}$	wawakisiw$_{AN}$
KROSHEE$_{MASC:IN}$ ‘hook’	FRENCH	3	3	0	0	3	crochet$_{MASC}$	sakâskwahonis$_{IN}$
KRWEE$_{FEM:IN}$ ‘cross’	FRENCH	1	0	1	0	1	croix$_{FEM}$	pimitâskwahikan$_{IN}$
KUULII$_{FEM:IN}$ ‘creek’	FRENCH	1	0	1	0	1	coulee$_{FEM}$	sîpîsis$_{IN}$
KUULOT$_{FEM:AN}$ ‘pants’	FRENCH	2	0	2	2	0	culotte$_{FEM}$	-tâs$_{AN}$
KUUVECHOER$_{MASC:IN}$ ‘roof; cover’	FRENCH	1	1	0	0	1	couverture$_{FEM}$	apahkwân$_{IN}$
KWASHOÑ$_{MASC:AN}$ ‘pig’	FRENCH	4	4	0	4	0	cochon$_{MASC}$	kohkôs$_{AN}$
L/N/ARZHAÑ$_{MASC:AN}$ ‘money’	FRENCH	2	2	0	2	0	argent$_{MASC}$	sôniyâw$_{AN}$
L/N/IGLIIZ$_{FEM:IN}$ ‘church’	FRENCH	1	0	1	0	1	église$_{FEM}$	ayamihêwikamik$_{IN}$
L/N/ZAÑFAÑ$_{MASC:AN}$ ‘child’	FRENCH	3	3	0	3	0	enfant$_{MASC}$	awâsis$_{AN}$
L/N/ZIKOL$_{FEM:IN}$ ‘school’	FRENCH	4	0	4	0	4	école$_{FEM}$	kiskinwahamâtowikamik$_{IN}$
L/N/ZISTWAR$_{FEM:IN}$ ‘story’	FRENCH	3	0	3	0	3	histoire$_{FEM}$	âcimowinis$_{IN}$
L/N/ZITWEL$_{FEM:AN}$ ‘star’	FRENCH	1	0	1	1	0	étoile$_{FEM}$	atâhk$_{AN}$
L/N/ZOM$_{MASC:AN}$ ‘man’	FRENCH	21	21	0	21	0	homme$_{MASC}$	nâpêw$_{AN}$
L/N/ZOÑK$_{MASC:AN}$ ‘uncle’	FRENCH	11	11	0	11	0	oncle$_{MASC}$	ohkomisimâw$_{AN}$
L/N/ZOOR$_{MASC:AN}$ ‘bear’	FRENCH	20	20	0	20	0	ours$_{MASC}$	maskwa$_{AN}$
L/N/ZOTOMOBIL$_{FEM:AN}$ ‘car’	FRENCH	1	0	1	1	0	auto(mobile)$_{FEM}$	sêhkêw$_{AN}$
L/NUUVRAAZH$_{MASC:IN}$ ‘work’	FRENCH	2	2	0	0	2	ouvrage$_{MASC}$	atoskewin$_{IN}$
LADDER$_{MASC:IN}$ ‘ladder’	ENGLISH	1	1	0	0	1	echelle$_{FEM}$	kîhcêkosîwinâhtik$_{IN}$
LAK$_{MASC:IN}$ ‘lake’	FRENCH	2	2	0	0	2	lac$_{MASC}$	sâkahikan$_{IN}$
LAÑG$_{FEM:IN}$ ‘language’	FRENCH	3	0	3	0	3	langue$_{FEM}$	pîkiskwêwin$_{IN}$
LAÑP$_{FEM:IN}$ ‘lamp’	FRENCH	3	0	3	0	3	lampe$_{FEM}$	wâsaskotênikan$_{IN}$
LAVEUZ$_{MASC:IN}$ ‘tub; washer’	FRENCH	1	1	0	0	1	laveuse$_{FEM}$	kisîpêkini-mahkahk$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
LEN$_{FEM:IN}$ ‘string’	FRENCH	1	0	1	0	1	laine$_{FEM}$ (‘wool’)	sakâpîhkanis$_{IN}$ ‘string’; pîswêkin$_{IN}$ ‘wool’
LII$_{MASC:IN}$ ‘bed’	FRENCH	2	2	0	0	2	lit$_{MASC}$	nipêwin$_{IN}$
LIIÑG-1$_{MASC:IN}$ ‘line’	FRENCH	1	1	0	0	1	ligne$_{FEM}$	miniskwahpicikan$_{IN}$
LIIÑG-2$_{FEM:IN}$ ‘fishing line’	FRENCH	6	0	6	0	6	ligne de pêche$_{FEM}$	okwâskwêpicikanê yâpiy$_{IN}$
LIIV$_{MASC:IN}$ ‘book’	FRENCH	1	1	0	0	1	livre$_{MASC}$	masinahikan$_{IN}$
LIWEE$_{FEM:IN}$ ‘law’	FRENCH	1	0	1	0	1	loi$_{FEM}$	wiyasiwêwin$_{IN}$
LUU$_{MASC:AN}$ ‘wolf’	FRENCH	17	17	0	17	0	loup$_{MASC}$	mahihkan$_{AN}$
LUUMYER$_{MASC:IN}$ ‘light’	FRENCH	1	1	0	0	1	lumière$_{FEM}$	wâsaskotênikan$_{IN}$
LUUN$_{FEM:AN}$ ‘moon’	FRENCH	3	1	2	3	0	lune$_{FEM}$	tipiskâwi-pîsim$_{AN}$
LYEV$_{MASC:AN}$ ‘rabbit’	FRENCH	2	2	0	2	0	lièvre$_{MASC}$ ‘hare’	wâpos$_{AN}$
MAAMA$_{MASC:AN}$ ‘mom’	FRENCH	2	2	0	2	0	maman$_{FEM}$	nimâmâ$_{AN}$
MAARD$_{FEM:IN}$ ‘shit’	FRENCH	2	0	2	0	2	merde$_{FEM}$ ‘shit’	mêyi$_{IN}$ ‘excrement’
MAEÑ$_{FEM:IN}$ ‘hand’	FRENCH	2	0	2	0	2	main$_{FEM}$	micihciy$_{IN}$
MAGAZAEÑ$_{MASC:IN}$ ‘store’	FRENCH	1	1	0	0	1	magasin$_{MASC}$	atâwêwikamik$_{IN}$
MAÑZHII$_{MASC:IN}$ ‘food’	FRENCH	3	3	0	0	3	manger$_{MASC}$	mîciwin$_{IN}$
MASHIN$_{VAR:INAN}$ ‘machine’	FRENCH	4	1	3	0	4	machine$_{FEM}$	âpacihcikan$_{IN}$
MATAEÑ$_{MASC:IN}$ ‘morning’	FRENCH	2	2	0	0	2	matin$_{MASC}$	kîkisepâ$_{IN}$
MEEZOÑ$_{FEM:IN}$ ‘house’	FRENCH	40	0	40	0	40	maison$_{FEM}$	wâskahikan$_{IN}$
MICHIF$_{MASC:AN}$ ‘Métis (person)’	FRENCH	3	3	0	3	0	métis$_{MASC}$	âpihtawikosisân$_{AN}$
MICHIN$_{FEM:IN}$ ‘medicine’	FRENCH	2	0	2	0	2	médecine$_{FEM}$	maskihkîwin$_{IN}$
MISTAKE$_{MASC:IN}$ ‘mistake’	ENGLISH	1	1	0	0	1	erreur$_{FEM}$	patâpahtamowin$_{IN}$
MOLAEÑ$_{MASC:IN}$ ‘machine’	FRENCH	1	1	0	0	1	moulin$_{MASC}$	âpacihcikan$_{IN}$
MOÑD$_{MASC:AN}$ ‘people’	FRENCH	39	39	0	39	0	monde$_{MASC}$	ayisiniwak$_{AN}$
MORSOO$_{MASC:IN}$ ‘piece’	FRENCH	1	1	0	0	1	morceau$_{MASC}$	pîwipicikan$_{IN}$
MOSHWEE$_{MASC:VAR}$ ‘handkerchief’	FRENCH	5	5	0	3	2	mouchoir$_{MASC}$	tâpiskâkan$_{AN}$
MOVII-IHKAAN$_{MASC:IN}$ ‘movie’	ENGLISH	1	1	0	0	1	film$_{MASC}$	âcikâstêpicikan$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
NII$_{MASC:IN}$ ‘nose’	FRENCH	2	2	0	0	2	nez$_{MASC}$	mikot$_{IN}$
NIIZH$_{FEM:AN}$ ‘snow’	FRENCH	1	0	1	1	0	neige$_{FEM}$	kôna$_{AN}$
NIK$_{VAR:INAN}$ ‘beehive’	FRENCH	5	2	3	0	5	nic$_{MASC}$ / nique$_{MASC}$	âmowaciston$_{IN}$
NIVEU$_{MASC:AN}$ ‘nephew’	FRENCH	3	3	0	3	0	neveu$_{MASC}$	nikosim$_{AN}$
NOÑ$_{MASC:IN}$ ‘name’	FRENCH	5	5	0	0	5	nom$_{MASC}$	wîhowin$_{IN}$
PAA$_{MASC:IN}$ ‘feast’	FRENCH	6	6	0	0	6	pas$_{MASC}$	wîhkohtowin$_{IN}$
PAAÑS$_{FEM:IN}$ ‘rumen (cow stomach)’	FRENCH	3	0	3	0	3	panse$_{FEM}$ ‘rumen’	omâw$_{IN}$
PAAPA$_{MASC:AN}$ ‘dad’	FRENCH	1	1	0	1	0	papa$_{MASC}$	nipâpâ$_{AN}$
PAAR$_{MASC:IN}$ ‘fence’	FRENCH	3	3	0	0	3	parc$_{MASC}$ ‘pen (of animals)’	mênikan$_{IN}$
PAAT$_{VAR:INAN}$ ‘leg’	FRENCH	3	1	2	0	3	patte$_{FEM}$	miskât$_{IN}$
PAEÑGFAEÑ$_{MASC:IN}$ ‘lice comb’	FRENCH	1	1	0	0	1	peigne fin$_{MASC}$ ‘fine tooth comb’	sîkahon$_{IN}$ ‘comb’
PAPII$_{MASC:IN}$ ‘paper’	FRENCH	3	3	0	0	3	papier$_{MASC}$	masinahikan$_{IN}$
PARAÑTII$_{FEM:AN}$ ‘relatives’	FRENCH	3	0	3	3	0	parenté$_{FEM}$	wâhkômâkan$_{AN}$
PARSON$_{VAR:ANIM}$ ‘person’	FRENCH	3	1	2	3	0	personne$_{FEM}$	ayisiyiniw$_{AN}$
PAYIIÑ$_{MASC:IN}$ ‘basket’	FRENCH	29	29	0	1	28	panier$_{MASC}$	mawiswâkan$_{IN}$; watapîwat$_{IN}$
PAZH$_{FEM:IN}$ ‘page’	FRENCH	3	0	3	0	3	page$_{FEM}$	pâskekinikan$_{IN}$
PEN$_{MASC:IN}$ ‘pen’	ENGLISH	1	1	0	0	1	plume$_{FEM}$	masinahikanâpisko s$_{AN}$
PER$_{MASC:AN}$ ‘priest’	FRENCH	3	3	0	3	0	père$_{MASC}$	ayamihêwiyiniw$_{AN}$
PHONE$_{MASC:IN}$ ‘telephone’	ENGLISH	1	1	0	0	1	téléphone$_{MASC}$	ayamâkan$_{IN}$
PII$_{MASC:IN}$ ‘foot’	FRENCH	1	1	0	0	1	pied$_{MASC}$	misit$_{IN}$
PLAÑSH$_{FEM:IN}$ ‘board’	FRENCH	11	0	11	0	11	planche$_{FEM}$	napakihtak$_{IN}$
PLAÑSHII$_{MASC:IN}$ ‘floor’	FRENCH	1	1	0	0	1	plancher$_{MASC}$	anâskânâhtik$_{IN}$ ‘floorboard’
PLEM$_{FEM:IN}$ ‘pen’	FRENCH	1	0	1	0	1	plume$_{FEM}$	masinahikanâpisko s$_{AN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
PLOT$_{FEM:IN}$ ‘ball’	FRENCH	1	0	1	0	1	pelote$_{FEM}$	pâkahatowân$_{AN}$
PLUU_VYEU$_{MASC:AN}$ ‘oldest (one)’	FRENCH	7	7	0	7	0	plus vieux$_{MASC}$	ostêsimâw$_{AN}$ ‘oldest male’
PLYBOARD$_{MASC:IN}$ ‘plyboard’	ENGLISH	1	1	0	0	1	planche$_{FEM}$ ‘board’	napakihtak$_{IN}$ ‘board’
POM$_{FEM:AN}$ ‘apple’	FRENCH	1	0	1	1	0	pomme$_{FEM}$	picikwâs$_{AN}$
POOSHIISH$_{MASC:AN}$ ‘cat’	CREE	4	4	0	4	0	minou$_{MASC}$	pôsîs-$_{AN}$ (Eng. borrowing)
PORT$_{FEM:IN}$ ‘door’	FRENCH	8	0	8	0	8	porte$_{FEM}$	iskwâhtêm$_{IN}$
PORTREE$_{MASC:IN}$ ‘picture’	FRENCH	12	12	0	0	12	portrait$_{MASC}$	masinipayiwin$_{IN}$
POSYER$_{FEM:IN}$ ‘dust’	FRENCH	1	0	1	0	1	poussière$_{FEM}$	pihko$_{IN}$
PROZHEE$_{MASC:IN}$ ‘project’	FRENCH	1	1	0	0	1	projet$_{MASC}$	oyesehcikewin$_{IN}$
PUUL-1$_{FEM:AN}$ ‘chicken’	FRENCH	12	0	12	12	0	poule$_{FEM}$	pâhkahâhkwân$_{AN}$
PUUL-2$_{FEM:IN}$ ‘chicken meat’	FRENCH	1	0	1	0	1	poule$_{FEM}$	pâhkahâhkwâniwiyâs$_{IN}$
PWASOÑ$_{MASC:AN}$ ‘fish’	FRENCH	13	13	0	13	0	poisson$_{MASC}$	kinosêw$_{AN}$
PWEL$_{FEM:AN}$ ‘stove’	FRENCH	1	0	1	1	0	poêle$_{MASC}$	kotawânâpisk$_{IN}$
RABABUU$_{FEM:IN}$ ‘stew; rabbit soup’	FRENCH	1	0	1	0	1	ragoût$_{MASC}$	mîcimâpôhkân$_{IN}$
RIVYER$_{FEM:IN}$ ‘river’	FRENCH	1	0	1	0	1	rivière$_{FEM}$	sîpiy$_{IN}$
ROB$_{FEM:IN}$ ‘dress’	FRENCH	3	1	2	0	3	robe$_{FEM}$	miskotâkay$_{IN}$
ROCHII$_{MASC:IN}$ ‘roast’	FRENCH	1	1	0	0	1	rôti$_{MASC}$	nawacîwin$_{IN}$
ROOF$_{MASC:IN}$ ‘roof’	ENGLISH	1	1	0	0	1	toit$_{MASC}$	apahkwân$_{IN}$
ROOM$_{MASC:IN}$ ‘room’	ENGLISH	1	1	0	0	1	pièce$_{FEM}$	apiwikamik$_{IN}$ ‘living room’
ROSH$_{FEM:AN}$ ‘rock’	FRENCH	15	0	15	14	1	roche$_{FEM}$	asiniy$_{AN}$
RUUBARB$_{FEM:AN}$ ‘rhubarb’	FRENCH	1	0	1	1	0	rhubarbe$_{FEM}$	pikwânâhtik$_{AN}$
SAK$_{MASC:IN}$ ‘bag, sack’	FRENCH	1	1	0	0	1	sac$_{MASC}$	maskimot$_{IN}$
SALOP$_{FEM:IN}$ ‘slop’	ENGLISH	1	0	1	0	1	bouette$_{FEM}$	wîyipapoy$_{IN}$
SAÑ$_{MASC:IN}$ ‘blood’	FRENCH	3	3	0	0	3	sang$_{MASC}$	mihko$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
SAÑD$_{FEM:IN}$ ‘ash’	FRENCH	3	0	3	0	3	cendre$_{FEM}$	pihko$_{IN}$
SAVOÑ$_{MASC:AN}$ ‘soap’	FRENCH	1	1	0	1	0	savon$_{MASC}$	kisîpêkinikan$_{IN}$
SAWMILL$_{MASC:IN}$ ‘sawmill’	ENGLISH	1	1	0	0	1	scierie$_{FEM}$	tâskipocikêwikamik$_{IN}$
SCHOOLBUS$_{MASC:AN}$ ‘schoolbus’	ENGLISH	1	1	0	1	0	autobus d’écoliers$_{MASC}$	âwatawâsiswâkan$_{AN}$
SEL$_{MASC:IN}$ ‘salt’	FRENCH	1	1	0	0	1	sel$_{MASC}$	sîwîhtâkan$_{IN}$
SEUK$_{MASC:IN}$ ‘sugar’	FRENCH	2	2	0	0	2	sucre$_{MASC}$	sôkâw$_{AN}$
SHAÑB$_{FEM:IN}$ ‘room’	FRENCH	1	0	1	0	1	chambre$_{FEM}$	apiwikamik$_{IN}$ ‘living room’
SHAPOO$_{MASC:IN}$ ‘hat’	FRENCH	14	14	0	0	14	chapeau$_{MASC}$	astotin$_{IN}$
SHAR-1$_{MASC:AN}$ ‘car’	FRENCH	1	1	0	1	0	char$_{MASC}$	sêhkêw$_{AN}$
SHAR-2$_{MASC:IN}$ ‘train’	FRENCH	1	1	0	0	1	char$_{MASC}$ ‘car, chariot’	iskotêwitâpân$_{AN}$
SHEEZH$_{FEM:IN}$ ‘chair’	FRENCH	4	1	3	0	4	chaise$_{FEM}$	têhtapiwin$_{IN}$
SHEV$_{FEM:AN}$ ‘goat’	FRENCH	6	0	6	6	0	chèvre$_{FEM}$	wâpatihk$_{AN}$
SHMIIZH$_{FEM:IN}$ ‘shirt’	FRENCH	2	0	2	0	2	chemise$_{FEM}$	pakowayân$_{IN}$
SHORT$_{MASC:AN}$ ‘shorts’	FRENCH	1	1	0	1	0	short$_{MASC}$	kîskicâsis$_{AN}$
SHOVREU$_{MASC:AN}$ ‘deer’	FRENCH	8	8	0	8	0	chevreuil$_{MASC}$ / chevreux$_{MASC}$	wêpâyôs$_{AN}$
SHYAEÑ$_{MASC:AN}$ ‘dog’	FRENCH	44	44	0	44	0	chien$_{MASC}$	atim$_{AN}$
SHYEN$_{FEM:AN}$ ‘dog’	FRENCH	1	0	1	1	0	chienne$_{FEM}$	atim$_{AN}$
SIMAÑ$_{MASC:AN}$ ‘cement’	FRENCH	1	1	0	1	0	ciment$_{MASC}$	asinîwipayihcikan$_{AN}$
SIMICHYER$_{MASC:IN}$ ‘cemetery’	FRENCH	1	1	0	0	1	cimetière$_{MASC}$	kihkwahaskân$_{IN}$
SLACKS$_{MASC:AN}$ ‘slacks’	ENGLISH	1	1	0	1	0	pantalon$_{MASC}$	nitâs$_{AN}$
SOER$_{FEM:AN}$ ‘sister’	FRENCH	15	0	15	15	0	soeur$_{FEM}$	-îtisân$_{AN}$
SOPII$_{MASC:IN}$ ‘supper’	FRENCH	1	1	0	0	1	souper$_{MASC}$	otâkwani-mîcisowin$_{IN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
STOOL$_{MASC:IN}$ ‘stool’	ENGLISH	2	2	0	0	2	tabouret$_{MASC}$	têhcikâpawin$_{IN}$
STORE$_{MASC:IN}$ ‘store’	ENGLISH	6	6	0	0	6	magasin$_{MASC}$	atâwêwikamik$_{IN}$
STORY$_{MASC:IN}$ ‘story’	ENGLISH	3	3	0	0	3	histoire$_{FEM}$	âcimowinis$_{IN}$
SUUYII$_{MASC:IN}$ ‘shoe’	FRENCH	3	3	0	0	3	soulier$_{MASC}$	maskisin$_{IN}$
SWIS$_{VAR:ANIM}$ ‘squirrel; gopher’	FRENCH	3	1	2	3	0	suisse$_{MASC}$ ‘chipmunk’	anikwacâs$_{AN}$
TAAB$_{MASC:IN}$ ‘barn’	FRENCH	1	1	0	0	1	étable$_{FEM}$	mistatimokamik$_{IN}$
TABAA$_{MASC:AN}$ ‘tobacco’	FRENCH	1	1	0	1	0	tabac$_{MASC}$	ciscêmâs$_{AN}$
TAÑ$_{MASC:IN}$ ‘time’	FRENCH	3	3	0	0	3	temps$_{MASC}$	tipahikan$_{IN}$
TAÑT$_{FEM:AN}$ ‘aunt’	FRENCH	14	0	14	14	0	tante$_{FEM}$	nitôsis$_{AN}$
TER$_{FEM:IN}$ ‘earth; ground; land’	FRENCH	10	0	10	0	10	terre$_{FEM}$	askiy$_{IN}$
TERAEÑ$_{MASC:IN}$ ‘ground, land’	FRENCH	2	2	0	0	2	terrain$_{MASC}$	-skamikâw$_{IN}$
TET$_{FEM:IN}$ ‘head’	FRENCH	7	0	7	0	7	tête$_{FEM}$	mistikwân$_{IN}$
TII$_{MASC:IN}$ ‘tea’	FRENCH	1	1	0	0	1	thé$_{MASC}$	maskihkîwâpoy$_{IN}$
TOROO$_{MASC:AN}$ ‘bull’	FRENCH	4	4	0	4	0	taureau$_{MASC}$	pônîw$_{AN}$
TRAEÑ$_{MASC:IN}$ ‘noise’	FRENCH	2	2	0	0	2	traînement$_{MASC}$	kisewehtâwin$_{IN}$
TRAMB$_{MASC:AN}$ ‘tree’	FRENCH	12	12	0	11	1	tremble$_{MASC}$ ‘aspen’	mistik$_{AN}$
TREATY_CARD$_{MASC:AN}$ ‘treaty card’	ENGLISH	1	1	0	1	0	carte de traité$_{MASC}$	iskonikanîwasinahikan$_{IN}$
TRIANGLE$_{MASC:IN}$ ‘triangle’	ENGLISH	3	3	0	0	3	triangle$_{MASC}$	matwêyâpiskahikan$_{IN}$ ‘sounding triangle’
TRIPALET$_{MASC:IN}$ ‘small intestine’	FRENCH	1	1	0	0	1	intestin grêle$_{MASC}$	nitakisiya$_{IN}$
TRUCK$_{MASC:AN}$ ‘truck’	ENGLISH	2	2	0	2	0	camion$_{MASC}$	âwatâswâkan$_{AN}$
TRUU$_{MASC:IN}$ ‘hole’	FRENCH	10	10	0	0	10	trou$_{MASC}$	wâci$_{IN}$
TV$_{MASC:IN}$ ‘TV, television’	ENGLISH	1	1	0	0	1	télévision$_{FEM}$	cikâstêpayihcikan$_{IN}$
UNDERWEAR$_{MASC:IN}$ ‘underwear’	ENGLISH	1	1	0	0	1	sous-vêtements$_{MASC}$	atâmayiwinis$_{AN}$
VAEÑ$_{MASC:IN}$ ‘wine’	FRENCH	1	1	0	0	1	vin$_{MASC}$	sôminâpoy$_{IN}$
VASH$_{FEM:AN}$ ‘cow’	FRENCH	8	0	8	8	0	vache$_{FEM}$	mostos$_{AN}$

Lemma	Source	Tokens	Gender/animacy by token				French	Cree
			Masc	Fem	Anim	Inan		
VISEL$_{FEM:IN}$ ‘dishes’	FRENCH	2	0	2	0	2	vaisselle$_{FEM}$	oyâkan$_{IN}$
VYAÑD$_{FEM:IN}$ ‘meat’	FRENCH	5	0	5	0	5	viande$_{FEM}$	wiyâs$_{IN}$
VYEE-1$_{FEM:AN}$ ‘old lady’	FRENCH	7	0	7	7	0	vieille$_{FEM}$	nôtokwêsiw$_{AN}$
VYEE-2$_{FEM:AN}$ ‘wife’	FRENCH	1	0	1	1	0	vieille$_{FEM}$	wîwimâw$_{AN}$
VYEU-1$_{MASC:AN}$ ‘old man’	FRENCH	13	13	0	13	0	vieux$_{MASC}$	kisêyiniw$_{AN}$
VYEU-2$_{MASC:AN}$ ‘husband’	FRENCH	9	9	0	9	0	vieux$_{MASC}$	nikisêyinîm$_{AN}$
WART$_{MASC:AN}$ ‘wart’	ENGLISH	1	1	0	1	0	verrue$_{FEM}$	micîhcîkom$_{AN}$
WASHING_MACHINE$_{MASC:AN}$ ‘washing machine’	ENGLISH	2	2	0	2	0	machine à laver$_{FEM}$	kisîpêkinikâkan$_{IN}$
WAZAEÑ$_{MASC:AN}$ ‘neighbor’	FRENCH	2	2	0	2	0	voisin$_{MASC}$	wîtapimâkan$_{AN}$
WEEZOÑ$_{MASC:AN}$ ‘bird’	FRENCH	1	1	0	1	0	oiseau$_{MASC}$	piyêsîs$_{AN}$
WEST$_{MASC:IN}$ ‘west’	ENGLISH	1	1	0	0	1	ouest$_{MASC}$	pahkisimôtâhk$_{IN}$
WIL$_{MASC:IN}$ ‘oil’	FRENCH	2	2	0	0	2	huile$_{FEM}$	pimiy$_{IN}$
YAEÑS$_{FEM:AN}$ ‘niece’	FRENCH	1	0	1	1	0	nièce$_{FEM}$	nistim$_{AN}$
ZANIMOO$_{MASC:AN}$ ‘cattle’	FRENCH	1	1	0	1	0	animaux$_{MASC}$	mostos$_{AN}$
ZHAAM$_{FEM:IN}$ ‘leg’	FRENCH	1	0	1	0	1	jambe$_{FEM}$	miskât$_{IN}$
ZHARDAEÑ$_{MASC:IN}$ ‘garden’	FRENCH	2	2	0	0	2	jardin$_{MASC}$	kiscikânis$_{IN}$
ZHILII$_{FEM:IN}$ ‘jelly’	FRENCH	1	0	1	0	1	gelée$_{FEM}$	nanamipayîs$_{IN}$
ZHORNII$_{FEM:IN}$ ‘day’	FRENCH	1	0	1	0	1	journée$_{FEM}$	kîsikâw$_{IN}$
ZHWAL$_{MASC:AN}$ ‘horse’	FRENCH	24	24	0	24	0	cheval$_{MASC}$	misatim$_{AN}$
ZWEE-1$_{MASC:AN}$ ‘goose’	FRENCH	5	5	0	5	0	oie$_{FEM}$	niska$_{AN}$
ZWEE-2$_{MASC:IN}$ ‘goose meat’	FRENCH	1	1	0	0	1	oie$_{FEM}$	niska-wiyâs$_{IN}$
ZYEU$_{MASC:IN}$ ‘eye’	FRENCH	1	1	0	0	1	yeux$_{MASC}$ ‘eyes’	miskîsik$_{IN}$

Appendix 2: Michif animates in the dataset by semantic category

Semantic category	Example(s)	
HUMAN BEINGS	*beebii*	'baby'
	belmer	'mother-in-law'
	blaañ	'white (man)'
	boñfrer	'brother-in-law'
	boñper	'father-in-law'
	boyfriend	'boyfriend'
	faam	'woman'
	fii	'girl'
	fimel	'female'
	frer	'brother'
	gaa	'guy'
	garsoñ	'boy'
	jeeñg	'girl; girlfriend'
	krosh	'crook'
	l/n/zañfañ	'child'
	l/n/zom	'man'
	l/n/zoñk	'uncle'
	maama	'mom'
	Michif	'Métis (person)'
	moñd	'people'
	niveu	'nephew'
	paapa	'dad'
	parañtii	'relatives'
	parson	'person'
	per	'priest'
	pluu vyeu	'oldest (one)'
	soer	'sister'
	tañt	'aunt'
	vyee	'old lady; wife'
	vyeu	'old man; husband'
	wazaeñ	'neighbor'
	yaeñs	'niece'
ANIMALS	*bilee*	'billy goat'
	billy goat	'billy goat'
	bish	'elk'
	chim di zhvoo	'team of horses'
	gournoy	'frog'
	iibuu	'owl'
	kanaar	'duck'
	kok	'rooster'
	kwashoñ	'pig'

Semantic category	Example(s)	
	l/n/zoor	‘bear’
	luu	‘wolf’
	lyev	‘rabbit’
	pooshiish	‘cat’
	puul	‘chicken’
	pwasoñ	‘fish’
	shev	‘goat’
	shovreu	‘deer’
	shyaeñ	‘dog (male)’
	shyen	‘dog (female)’
	swis	‘squirrel; gopher’
	toroo	‘bull’
	vash	‘cow’
	weezoñ	‘bird’
	zanimoo	‘cattle’
	zhwal	‘horse’
	zwee	‘goose’
MOST TREES	*tramb*	‘tree’
CERTAIN PLANTS AND THEIR PRODUCTS	*arañzh*	‘orange’
	cake	‘cake’
	fariin	‘flour’
	galet	‘bannock’
	pom	‘apple’
	ruubarb	‘rhubarb’
SOME HUMAN BODY PARTS AND PHYSICAL CONDITIONS	*wart*	‘wart’
TOBACCO AND RELATED ITEMS	*tabaa*	‘tobacco’
SOME NATURAL OBJECTS	*rosh*	‘rock’
SOME PERSONAL/HOUSEHOLD ITEMS	*ball*	‘ball’
	bwatoñ	‘button’
	cookstove	‘cookstove’
	kataeñ	‘doll’
	klosh	‘bell’
	kop	‘penny’
	l/narzhañ	‘money’
	pwel	‘stove’
	savoñ	‘soap’
	simañ	‘cement’
	treaty card	‘treaty card’
PHENOMENA OF THE NATURAL ENVIRONMENT / CELESTIAL BODIES	*glas*	‘ice’
	l/n/zitwel	‘star’
	niizh	‘snow’
SOME ITEMS OF CLOTHING	*jeans*	‘jeans’
	kuulot	‘pants’
	short	‘shorts’

Semantic category	Example(s)	
	slacks	‘slacks’
SOME MACHINES	*washing machine*	‘washing machine’
MEANS OF TRANSPORTATION	*4 X 4*	‘4 X 4’
	ambulance	‘ambulance’
	bicycle	‘bicycle’
	bike	‘bicycle’
	l/n/zotomobil	‘car’
	schoolbus	‘schoolbus’
	shar	‘car’
	truck	‘truck’