An Analysis of the Split-Phonology Hypothesis in Michif: Sociolinguistic and Phonological Perspectives

Matthew Fitzsimmons, Alexah Konnelly, Sarah Provan, & Alison Root
University of Victoria
mfitzsim@uvic.ca; a.konnelly@mail.utoronto.ca; provan.sarah@gmail.com; alisonlroot@gmail.com

Michif, a mixed language spoken by the Métis people, is highly unique. Unlike other mixed languages that take their lexicon from one base language and their grammar from another, Michif is divided etymologically within the lexicon: nouns and noun phrases come from French, and verbs and verb phrases come from Cree. This ‘split’ in the lexicon has led many researchers to question whether or not the phonology of Michif is similarly divided. This paper seeks to contribute to this Michif split-phonology debate. In so doing, we discuss the context that fostered the unique characteristics of Michif as a mixed language, summarize previous claims made in the split-phonology debate, and present a case study grounded in a phonological process from one of Michif’s base languages—the French-origin liaison. While results show that liaison occurs only between etymologically French words, we argue that this does not necessarily mean that the entire phonology is split. There are unique challenges and considerations involved in analyzing liaison in Michif which make any conclusion tentative at best. As such, it may be that the two base phonologies are intertwined in complex ways that our current linguistic understanding of phonological systems does not yet know how to account for.

Keywords: Michif; mixed languages; split-phonology hypothesis; liaison

1 Introduction

Spoken in Western Canada, North Dakota, and Montana by the descendants of French fur traders and Indigenous Cree speakers, Michif1 is one of the world’s few mixed languages, a language resulting from contact between multiple varieties (Bakker, 2003). Unlike pidgins and creoles, mixed languages do not have a superstrate and a substrate; typically, they instead take their lexicon from one language and their grammar from another (Bakker & Muysken, 1995).

1 Michif is also sometimes spelled Mitchif, Métchif or Méchif.
However, even within mixed languages Michif is unique: instead of taking its lexicon from one base language and its grammar from the other, Michif is divided within the lexicon. It takes its nouns from French and its verbs from Cree, with Cree phrases following Cree syntax and French phrases following French syntax. Therefore, its lexicon is widely accepted as being ‘split’. This has raised questions among linguists as to whether or not the Michif phonological inventory is similarly divided along etymological lines: can Michif be said to have one unified phonological system, or are there two separate, coexistent ones? While some researchers (e.g., Evans, 1982; Bakker, 1997) claim that there are in fact two, described by Bakker (1997) as “one for the Cree part and one for the French part, each with its own rules” (p. 7), others argue that there is only one system (e.g., Papen, 2011; Prichard & Shwayder, 2014). As such, the nature of phonological patterning in Michif is not yet clear.

This project seeks to build on this previous scholarship and provide evidence for a split phonology in Michif. In so doing, we discuss the context which fostered the unique characteristics of Michif as a mixed language, summarize previous claims made in the split-phonology debate, and conclude with a case study grounded in a phonological process from one of Michif’s base languages: the French-origin liaison. Liaison is a French phonological process where underlyingly silent, word-final consonants are pronounced when followed by a vowel-initial word (e.g., [patikɔpɛ] ‘petit copain’ vs. [patítɔmi] ‘petit ami’). Given that French makes up a large portion of Michif’s lexicon, whether or not liaison is a productive process with both etymologically Cree and French words will be significant for the Split Phonology Hypothesis.

Our research questions are therefore twofold: (1) Is the French liaison rule productive in Michif and, if so, (2) is it restricted to only the etymologically French portion of the language? If liaison does apply to only the French portion of the lexicon, this could be taken as evidence that a split phonology does exist, with French phonological processes applying only to French words. If liaison is fully productive (that is, applies to both etymologically French and Cree words), this could be taken as evidence that a split phonology does not exist, and phonological processes—at least those inherited from French—apply to words whose origins are in either language. Our results show that liaison occurs only between etymologically French words, and even following the fossilized [t] in tout. However, this does not necessarily mean that the entire phonology is split. There are unique challenges and considerations involved in analyzing liaison in Michif which make any conclusion tentative at best. It may be that the two base phonologies are intertwined in complex ways that our current linguistic understanding of phonological systems does not yet know how to account for. We will return to these issues below.

2 Here, the term phrase is used in the theoretically syntactic sense (e.g., noun phrase, verb phrase, etc.). Generally, other lexical categories such as adjective, determiner, adverb, etc. within a phrase are all from the same source language, although items from another source language can be added (Bakker, 1997, p. 102).
2 Sociolinguistic perspectives

Michif is the language of the Métis people in Canada and the United States. There are many languages spoken by the Métis, including English, Métis French, Métis Cree, and Métis Saulteaux, but “only Michif is truly theirs” (Martin & Bonneville, 2001, n.p.). Michif is unique; it is one of the few mixed languages of the world—that is, a language formed when “two halves are combined into one organic whole” (Bakker, 1997, p. 210). Mixed languages (such as Michif, Mednyj Aleut, Ma’a, Media Lenga, etc.) are often likened to pidgins and creoles, though there are important differences. A pidgin is formed when two groups come into contact and form a grammatically reduced language for the purpose of communication; it is not spoken as a native language (Bakker, 1997). Creoles are contact languages with complete grammatical systems, and a creole is thought to be derived from a pidgin once a new generation of speakers uses it as a first language (Bakker, 1997). Creoles typically have a single dominant (i.e., superstrate) language as the main influence, providing the majority of the grammatical rules (Bakker & Muysken, 1995). This is contrasted with a mixed language, which is a language that “shows positive genetic similarities, in significant numbers, with two different languages” (Bakker, 1997, p. 195). Thomason (2001) thus characterizes mixed languages as ones “whose grammatical and lexical subsystems cannot all be traced back primarily to a single source language,” contrary to all other forms of language, as the grammars of mixed languages can be directly traced to two languages (p. 21).

A mixed language is formed in a dual language situation. However, this is not simply a language contact occurrence, but actually involves an intertwining. For this to occur, bilingual speakers with an understanding of both parent languages are necessary. The result of this intertwining is generally that one language supplies the grammar for the new mixed language, while the other provides that lexicon. Bakker (1997) has proposed that, if this highly bilingual environment is present, there are three factors that play a role the formation of a mixed language: (1) a high number of speakers of both languages, (2) no outside linguistic pressure, and (3) a distinct group of people. For the first factor, Bakker argues that the most important aspect is that the parent languages be uniformly divided within the community. The second factor is that there must be no outside linguistic pressure, such as a dominant language group asserting pressure on the speakers of the two parent languages. Finally, the speakers of the two parent languages must make up a distinct group of people that are separate from the cultures of the origin languages. All three of these factors appear in the history of the Métis people.

The Métis were the descendants of French fur traders and Cree women, and began to self-identify as such in the early 1800s. The first records of the Michif language are from approximately 1820. Margaret Desjarlais, a native speaker of Michif, describes the origins of the language in an interview with Peter Bakker (1996):
When the French Canadians came from across the ocean, they started to marry Indian women and then they had children. The Indian woman couldn't speak French to her children. The Frenchman couldn't speak Cree to his children, so he spoke to them in French. Therefore some of them learned to speak French and Cree. Therefore he speaks only French and Cree mixed. (p. 8)

This description of the birth of Michif is analogous to the birth of the Métis culture, as a group of people who had ties to both the French and Cree peoples but did not necessarily belong to either group. This was fuelled by a desire to be both—and neither—Cree and French. The Métis thought of themselves as their own people, and the Michif language helped to define Métis culture. The development of the language was further helped along by Bakker’s (1997) three factors discussed above: almost all speakers were found in an identical environment, as the men were French speakers, and the women were Cree speakers. There was also no outside linguistic influence. Due to struggles in the Red River area and other conflicts with the government, the Métis withdrew from many aspects of mainstream Canadian society (Rosen, 2015). By doing so, they avoided linguistic and social pressures to conform to English or French, which helped preserve Michif for some time. This leads to the third factor, that the speakers of the parent languages must be removed from their original cultures. Since the Métis were not accepted by either the colonial settlers or First Nations people, the mixed language was able to continue to form in this regard.

However, Michif is distinct among mixed languages. Bakker (1997) argues that the native language of the males (i.e., the fathers) usually provides the lexicon, while the language of the females (i.e., the mothers) provides the grammar (p. 207). Michif appears to defy this, as the language is not ‘split’ between the lexicon and the grammar, but within the grammar itself: the nouns and noun phrase syntax come from French, while the verbs and verb phrase syntax come from Cree. This irregular split in the lexicon has led to much debate about the phonology of this mixed language—specifically, whether or not the phonology of the language is similarly divided along etymological lines.

---

3 While this at first seems to go against the usual trends of mixed languages, Bakker (1997) suggests that this difference is due to the polysynthetic nature of Cree. Cree verbs exhibit high degrees of affixation, and because of this complexity, a Cree verb can also be said to be a representation of the grammar.
3 Split-phonology debate

Michif being composed of fully realized portions of French and Cree raises a question regarding its phonology: are there separate phonologies that deal with each historically distinct portion of the language, or is the phonology unified? This question has seen a great deal of debate among researchers of the language.

The first area of debate is the phonemic inventory of Michif. Some researchers have posited two wholly distinct phonemic inventories for the French and Cree portions (Rhodes, 1986; Bakker, 1997). This is due partly because they observe words from each origin using a different set of phonemes than the other, some being completely restricted to one etymological origin. Recently, Rosen (2007) has argued that it is not necessary to stipulate two distinct phonologies because there is an uneven distribution of phonemes between the French and Cree components, and she provides acoustic evidence that similar phonemes considered to be distinct by other researchers have in fact mapped onto each other. Additionally, Rosen (2007) believes there to be no evidence that both the long-short contrast of Cree phonology and the tense-lax contrast of French are both independently operational; she argues that in Michif, there is a single collapsed two-way contrast system based on vowel quality rather than vowel length.

In addition to distributional differences in the phonemic inventories, there are also cases where single phonemes are said to function differently between the two source components. For example, Rhodes (1986) found /a/ to be realized as [a], [ʌ], or [ɛ] in the Cree component of the grammar, while in the French component it surfaces as only [a] or [æ]. Conversely, Rosen (2007) found that for the speakers in her study, /a/ surfaced as [a] or [ʌ] in both French and Cree contexts.

One area that does seem to support a unified phonology is stress assignment. The difference in stress assignment between Cree and French is only apparent in words over two syllables long, and in Michif it appears that French words have been converted to the Cree system of stress assignment (Bakker, 1997). The length of stressed syllables in three-and four-syllable words reflects Cree patterning, with the antepenultimate syllable receiving main stress, rather than stress being lexically determined as in French.

A significant area of contention is that of phonological processes, which in some cases seem to be limited to French or Cree contexts. Papen (2003) argues that certain vowel deletion processes are limited to French contexts: for example, he claims that the French phonological rule that some grammatical words with a CV shape elide their vowels before vowel initial words is still functional in Michif. This can be seen with the word *di* (historically French *de*). 
For example, the [d] is fully realized in (1), but before a vowel-initial word such as in (2), it is elided. Papen (ibid) additionally argues that a similar French-origin process, wherein /i/ is deleted following a single pronounced consonant, is also operational in Michif. He claims that these rules are only applied to French origin words, and therefore provide evidence for a split phonology.

Finally, the question of whether the French phonological process of liaison remains operational in Michif has received a great deal of attention. Liaison describes a process wherein a final, unpronounced consonant becomes pronounced as the onset of a following vowel-initial word. There are a number of contexts where liaison takes place, but the most common context is within a noun phrase, where the final consonant of determiner or adjective is pronounced as the onset of a following noun. This can be seen in example (3).

(3)  les amis

DET friends

‘the friends’

The /s/ in ‘les’ would not be pronounced in isolation or before a consonant-initial noun, but here it is syllabified as the onset of the vowel-initial ‘amis’ and pronounced.

Liaison within entirely French-origin noun phrases has drawn the most attention. Papen (2003) argues that liaison in Michif functions in much the same way as it does in French: in a study of dictionary material, he found that vowel-initial nouns in Michif either remain vowel-initial or have the expected liaison consonant (e.g., from an an in/definite article) as their onset the majority of the time. There are, however, cases where historically vowel-initial words have unexpected initial consonants—these appear to correspond to liaison consonants that have become fossilized as onsets of the noun. Rosen (2007), however, takes a much different view, arguing that all vowel-initial French-origin nouns have been re-analyzed as consonant-initial, with liaison consonants fossilizing as onsets. She believes that cases where the initial consonant is variable can be analyzed as morphological rather than as instances of liaison.

Other contexts of liaison have been largely ignored until a recent paper by Prichard & Shwayder (2014), which examines the extent to which liaison is limited to etymologically French words. They analyze two instances of liaison. The first instance consists of a French-origin pronoun followed by a Cree-origin.
verb. The second instance of liaison consists of a French-origin adjective being triggered by a vowel initial Cree noun (a few do exist as exceptions to the general tendency). Liaison occurs in both instances, and they take this as evidence of a unified phonology; though liaison consonants may be limited to French-origin words, the fact that non-French-origin words can trigger them suggests that liaison itself is operational in all components of the grammar.

4 Case study

It appears as if there is as much evidence for a split phonology as there is against it. With this in mind, we turn to our case study.

4.1 Data and methodology

The data for our analysis have been drawn from online Michif educational materials, specifically the narrated children’s book section of the Gabriel Dumont Institute’s Michif Museum website. All books were translated from English and narrated in Michif by Flamont and Pelletier (2006a, 2006b, 2006c, 2006d). The children’s books were the only Michif-language materials available on the Michif Museum’s website that were both narrated and written in Michif, as well as having an accompanying English translation. As such, the data set is limited to tokens found within these materials only, and only six were found.

<table>
<thead>
<tr>
<th>Context</th>
<th>Alfred’s First Day at School</th>
<th>Lisa and Sam</th>
<th>The Big Storm</th>
<th>The Pow Wow</th>
</tr>
</thead>
<tbody>
<tr>
<td>French-French (F-F)</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>French-Cree (F-C)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total N</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1: Number of potential liaison contexts by children’s book.

Audio files from each book and their accompanying text were downloaded and searched for potential liaison contexts—that is, French-origin words ending in consonants and preceding a vowel-initial word. The etymology of words in potential liaison context was highly important. Since whether or not liaison can apply to non-French origin words is one of the main research questions of this project, it was necessary to determine, with as much accuracy as possible, whether the words in the potential liaison contexts came from French or Cree. However, there are no Michif-English, Michif-French, or Michif-Cree dictionaries available, and there is no conventional writing system for the language; consequently, while some items are more transparent than others (e.g., Michif ‘ikol from French l’école, l’itee from French l’été), the origin of all words was not entirely clear. As such, based on the fact that Michif takes its nouns from
French and its verbs from Cree, we have assumed that all nominal items (as evidenced in the translations) in liaison contexts came from French and verbal items in liaison contexts came from Cree⁴.

A summary of the potential liaison contexts identified in the texts can be found in Table 1. ‘French-French’ contexts refer to those where a consonant-final word of French origin precedes a vowel-initial word of French origin. ‘French-Cree’ contexts refer to those where a consonant-final word of French origin precedes a vowel-initial word of Cree origin. Since there is no evidence for liaison occurring in Cree, Cree words ending in a consonant preceding a vowel-initial word of any origin were not considered to be potential liaison contexts. Following data collection, tokens with the requisite phonological environment for liaison were analyzed in Praat to determine whether or not the word-final consonant was pronounced and elided.

4.2 Results

All of the French-French contexts analyzed contained liaison, while liaison in French-Cree contexts was not found. Fossilization and orthography added additional complications, to be discussed below.

The analysis was made challenging due to the lack of a standardized orthography for Michif, because orthographic representations of speech may obscure processes of liaison. For example, in the phrase tout dahn la ville, (cf. Figure 1), [d] could represent either a fossilized consonant or an instance of liaison. In this case, we analyzed the [d] as a liaised consonant due to the lack of evidence for dahn (but ample evidence for ah) as a lexical item in Michif. Throughout the texts, there are many instances of ah used as a preposition in a parallel distribution to the dahn in the above phrase (such as in the phrase sa premier jour ah l’ikol), suggesting that this word comes from the French à, meaning to.⁵

---

⁴ Based on the English translations available in the books, we were able to look up words in the Gabriel Dumont Institute’s online English-Michif dictionary (http://www.Metismuseum.ca/michif_dictionary.php), which features over 11,500 translations and audio pronunciations by Michif-language expert Norman Fleury.

⁵ Additionally, as a grammatical function word, ah occurs in too many variable contexts for any consonant to proceed regularly enough to fossilize as an onset. Also, de would never precede à in French, making the grammaticality of the source construction questionable. It is possible that such a construction is not ungrammatical in Michif, but without access to native speakers, this cannot be certain.
The word *tout* in Michif itself appears to contain an instance of fossilization: while the final [t] is only pronounced in French in liaised or grammatically feminine contexts, it is always pronounced in Michif as [tut] irrespective of the phonological environment or grammatical gender. This is shown in Figure 2, where *tout* precedes the consonant-initial grammatically masculine (evidenced by the masculine singular definite article *li*) *swyr*.

---

6 It is also possible that this is not in fact fossilization, but the retention of a Québécois French dialect feature. According to Laurence Godard (personal communication), in colloquial Québécois French, the final [t] in *tout* is sometimes pronounced irrespective of grammatical environment. However, more speakers are needed to determine the validity of this possibility.

7 The [t] is audible, but is not obvious on the spectrogram. This is hardly surprising assuming that this /t/ is unaspirated, which /t/ often is in Michif. Optional aspiration of /t/ is a result of influence from English (Bakker, 1997).
Figure 2: Spectrogram for phrase tout li swayr with final [t] of tout pronounced.

Taking this fossilization into account, this would appear to conform to French native speaker intuitions. Even if there is a pause, if there is a word that ends in a consonant followed by a word beginning in a vowel, speakers cannot drop the liaison (Goddard, personal communication). In Figure 1, the speaker applies this rule productively. He pronounces the fossilized [tut], pauses, and then applies the liaison rule to add [d] to ah.

While liaison appears to be triggered despite the fossilized [t] consonant in a French-French context, it does not do so with French-Cree contexts. This is shown in the spectrogram of the French-Cree phrase tout-oh-wuk in Figure 3, where the fossilized [t] in tout is pronounced, but there is no [d] as in Figure 1.
Figure 3: Spectrogram for (F-C) phrase tout-oh-wuk with pronounced fossilized [t], but no liaison

The other remaining French-Cree token, *apres aqua*, does not appear to exhibit liaison either, as shown in Figure 4. However, there is some variation in whether or not the final consonant in *apres* can undergo liaison in French; it is described as an "optional liaison context" (Lawless, n.d.). Consequently, it is not entirely unsurprising that liaison is not present in this instance.

Figure 4: Spectrogram for (F-C) phrase apres aqua, with no liaison

Fossilization notwithstanding, there are French-French instances where liaison appears to be more straightforward; for example, Figures 5 and 6.
However, orthography is again problematic, particularly for *zeclarr* in Figure 6. As with *dah*, based on dictionary consultation, there is no evidence for *zeclarr* as a lexical item in Michif—yet there is plenty of evidence for *eclarr*, not just as an independent item but also in compounds such as *fayr d’iklayr* (‘lightning rod’). It does seem telling that the orthography in the text shows a *<z>* as being part of the onset of a vowel-initial word following a plural determiner where liaison would be present in French, but the same noun in non-liaison contexts (e.g., in a compound following an always realized consonant, such as *d’*) does not have a *z*, either in the orthography or in the pronunciation. This strongly suggests that this is an instance of liaison, but that it is not orthographically realized as such. This is unsurprising, considering the lack of a standardized orthography. This may indicate that speakers are unaware of the process of liaison in Michif. Since most Michif speakers aren’t bilingual in French (Martin & Bonneville, 2001), it is unsurprising that French orthography would not influence Michif orthography.
5 Discussion and Implications

Rosen’s (2007) analysis of vowel-initial French-origin nouns in Michif as categorically possessing fossilized onset consonant is not supported by our data. We have found a number of instances where historically vowel-initial nouns had different onset consonants in various contexts that one would expect to see liaison consonants, such as with the word *ikol*. This word was realized as *d’ikol* when preceded by the preposition *di*, and as *l’ikol* when preceded by the definite article. It seems clear from this example and the many others like it that liaison is functioning at some level within Michif.

Our results also counter those of Prichard and Shwayder (2014), who concluded that vowel-initial Cree-origin words can trigger liaison in French-origin words. That does not appear to be true for our data; we found liaison only within French-origin noun phrases. However, Prichard and Shwayder’s (2014) results were based upon just one instance of liaison with a Cree-origin word, and it happens that the French word supposedly undergoing liaison is *tout*. In this study, we identified *tout* as having a fossilized final /t/, and the pronunciation of this /t/ therefore does not represent liaison. It may be that the *tout* token that they analyzed also had a fossilized /t/, in which case liaison was not, in fact, triggered by the following Cree-origin word. This would be consistent with our findings. Alternatively, if we take their conclusion at face value, there would seem to be some degree of variability in the application of liaison for speakers of Michif. The existence of variability would be evidence against a split phonology; it hardly seems likely that Michif speakers who do use liaison have a split phonology, while those who do not have a unified one. If the application of liaison rules is variable for individual speakers (which it likely is, assuming variability between speakers), then the presence or absence of liaison cannot be accounted for with wholly different models of the phonology. This would require that the phonology be unified at some times and split at others.

This variability, as well as the fossilization that has occurred on some nouns, presents a problem for the development of a standard form of the language. Should some forms, perhaps more common ones, be accepted as correct, or should variability be permitted? If the latter route is chosen, how will this be represented in a standard orthography? Existing orthographies seem to favour writing liaison consonants as the onset of the following word, but this makes the difference between liaison consonants and fossilized consonants opaque, which presents a challenge to learners of Michif. Additionally, variable spellings may present a complication in the creation of dictionaries. It can also make looking up words more difficult for learners and create problems for language software that could be used to support the language. However, these are issues that we respectfully leave to the Métis community to decide.

In addition to the possibility of variability, it may also be that these results simply point to Michif as participating in a wider, cross-linguistic process of phonological intertwining that depends largely on the contact a given language has with another (Hyekyeong Ceong, personal communication). In other words,
French-origin processes may be productive in Michif in a similar way that Latin-origin processes are productive in Latin loan words in English, or Chinese-origin words in Korean and Japanese. This is an issue that we leave open to further research, as a larger body of data—and perhaps access to native speakers—would be necessary.

6 Conclusion

Determining the productivity of liaison in Michif is complicated because the language does not have a standardized orthography, and as such, it is difficult to determine whether certain consonants are instances of liaison or are fossilized onto lexemes.

Although our data suggest that liaison occurs only in the French-French contexts, particularly following the fossilized [t] in tout, it does not necessarily mean that the entire phonology is split. The literature surrounding the Michif split-phonology debate shows conflicting evidence for and against the hypothesis, but all studies on the topic suffer from low token counts for analysis, and this study is no exception. This, combined with the difficulties posed by the lack of Michif spelling conventions, makes any conclusion tentative at best. It may be that the two base phonologies are intertwined in complex ways that our current linguistic understanding of phonological systems does not yet know how to account for. Future studies would benefit from access to either a larger body of data from which to draw tokens, access to native speakers for elicitations, or both. While we are reluctant to form generalizations based on the small amount of data, it is our hope that we have made some small contribution into growing scholarship on the unique language that is Michif.

Acknowledgements

The authors would like to acknowledge Dr. Nicole Rosen for her helpful suggestions.

References


