The philological interpretation of small corpora in extinct languages, illustrated by relative clause data from Lule (Argentine Chaco)

Willem J. de Reuse
University of North Texas

The philological interpretation of small corpora in extinct languages, illustrated by relative clause data from Lule (Argentine Chaco)[∗]

Willem J. de Reuse
University of North Texas
willemdereuse@my.unt.edu

1. Introduction

Lule was spoken by a semi-nomadic people who lived in the Gran Chaco between the Pilcomayo river and the Andean foothills of northwestern Argentina. It is an extinct language, which was also spoken by other groups of the present-day Argentine provinces of Tucumán and Santiago del Estero. Virtually all of our documentation of Lule comes from the Sardinian Jesuit missionary Antonio Maccioni (1688-1753) in his *Arte y vocabulario de la lengua lule y tonocoté,* published under the name Machoni in Madrid in 1732. Mentions of Lule disappear from the historical record after the expulsion of the Jesuits in 1776. We may surmise that the Lule language lost its last speakers sometime in the course of the early 19th century.

2. Genetic affiliation and philological challenges

The Lule language is generally considered to be genetically related to the Vilela language, forming a Lule-Vilela family, but it remains possible that the lexical similarities are due to contact rather than to shared ancestry. The language described in Machoni (1732) provides interesting challenges for the study of linguistic diversity in the Chaco area, because it is typologically rather distinct from the other known Chaco languages as well as from neighboring Andean languages, and because we have to rely almost entirely on Maccioni's examples and linguistic abilities in trying to reconstruct its grammar in modern terms. The total corpus is estimated at 602 lines of examples or text in Machoni's (1732) grammar and text sections; the dictionary section of Machoni (1732) contains about 2,550 lexemes. The retranscription and morphological analysis in this paper is based on the comprehensive grammar of Lule which is being written by Raoul Zamponi and myself.

3. Relative clause formation in Machoni's (1732) grammar section
3.1. Introduction

There are 18 examples that can be interpreted as relative clauses in Machoni's (1732) grammar section, so it is possible to present the whole corpus of examples here. Each example is presented in five lines, so no information is lost: the first line is the sentence in the original spelling, the second line is Maccioni's translation or gloss in Spanish (with archaic spellings normalized to modern Spanish practice), the third line is a phonemic spelling with morpheme and clitic boundaries indicated, the fourth line is the word by word and morphological analysis, and the fifth line is a fairly literal translation into English. On that line, there is also a number in parentheses, which is the page number in Machoni's (1732) grammar section. One can distinguish four types of relative clause formation strategies, discussed in sections 3.2 through 3.5.

3.2. The agent nominalizer *-ton*

One type of construction involves the agent nominalizer *-ton,* as in examples (1) through (8):

(1) *pelè in lè otiqueyuton, vecinèp*
     el indio que trabajó ayer murió
     pele   inle   wotikeyu-ton   wetsi[-i]   ne-p
     man   yesterday work-NMLZ   die[-THEM]   P.REC.VIS-3sSBJ[1]
     'the man who worked yesterday died' (56)

(2) *pelè in lè otiqueyuton, vyetip*
     el indio que trabajó ayer murió
     pele   inle   wotikeyu-ton   uye-ti-p
     man   yesterday work-NMLZ   NEG-do-3sSBJ
     'the man who worked yesterday died (lit. was no more)' (56)

(3)   *pelè in lè otiqueyuton, vecip*
      el indio que trabajó ayer murió
      pele   inle   wotikeyu-ton   wetsi-p
     man   yesterday work-NMLZ   die-3sSBJ
     'the man who worked yesterday is dead' (56)

(4) *ecitòn zotà Dios meticant*
     Dios llevará al Cielo al que fuere bueno
     etsi-ton   tso=ta   Dios   metika-n-t
     be.good-NMLZ   heaven=LOC   God   take-FUT-3sSBJ

---

[1] Abbreviations used follow the Leipzig Glossing Rules, with the addition of: ANA 'anaphoric'; COLL 'collective'; P.REC.VIS 'recent visual past'; P.REM 'remote past'; PRO, 'independent pronoun'; •RED 'reduplicative process, meaning uncertain', and THEM 'thematic' (suffixed to imperatives and to the V1 of a serial verb construction). Material in square brackets are my philological interpretations or additions.

'God will take the one who is good to heaven' (56)

(5) *yuisinis yetòn Petro cumuepinèp*
     he oído al que dice, o dijo que Pedro se casó
     iwis-i   ni-s   ye-ton   Petro   kumue-p-i   ne-p
     hear-THEM   P.REC.VIS-1sSBJ   say-NMLZ   Peter   get.married.3SBJ-?   P.REC.VIS-3sSBJ
     'I heard the one who says/said that Peter got married' (56)

(6) *yuisinis yetòn Petro cumuèp*
     he oído al que dice, o dijo que Pedro se casó
     iwis-i   ni-s   ye-ton   Petro   kumue-p
     hear-THEM   P.REC.VIS-1sSBJ   say-NMLZ   Peter   get.married-3sSBJ
     'I heard the one who says/said that Peter is getting married' (56)

(7) *quys yepnicsintòn*
     yo tengo la obligación de enseñar
     kis   yepniks-in-ton
     1sPRO   teach-FUT-NMLZ
     'I am one who will teach' (56)

(8) *pelè quis tacevinstòn tià vyè*
     el indio que tengo de castigar no está aquí
     pele   kis   tak-tseu-in-s-ton   tia   uye
     man   1sPRO   INS-punish-FUT-1sSBJ-NMLZ   here   NEG
     'the man I will punish is not here' (64)

3.3. The agent nominalizer *-ton* and the anaphoric and resumptive pronoun *mima*
      Another type of construction involves the agent nominalizer *-ton,* as well as the anaphoric and resumptive pronoun *mima*, as in example (9). There is only one example of this construction in the corpus. I will deal with the clause translated as 'what God commands' in section 4.3., so let us ignore it for now.

(9) *Dios amaicitòn mimà tit vsa Dios nequequèsp*
     el que ama a Dios guarda sus mandamientos o lo que Dios manda
     Dios   amaitsi-ton   mima   ti-t   usa   Dios   ne-kekes-p
     God   love-NMLZ   ANA.DIST   do-3sSBJ   what God   INS-command-3sSBJ
     'the one who loves God, that one observes what God commands' (64)

3.4. The anaphoric and resumptive pronoun *mima* only
      Another type involves no nominalization, but only the anaphoric and resumptive pronoun *mima*, as in examples (10) through (15). The pronoun *mima* can be optional, as in examples (10) and (11).

(10) *pelè Dios alapip, zotà cant, o mimà zotà caànt*
     el hombre que Dios escogió, irá al Cielo

3

pele	Dios	alap-ip	(mima)	tso=ta	ka-n-t
man	God	choose-3sSBJ	(ANA.DIST)	heaven=LOC	go-FUT-3sSBJ

'the man God chooses, (that one) will go to heaven' (64)

(11) *pelè Dios alapinèp, zotà cant*, o *mimà zotà caànt*
el hombre que Dios escogió, irá al Cielo

pele	Dios	alap-i	ne-p	(mima)	tso=ta
man	God	choose-THEM	P.REC.VIS-3sSBJ	(ANA.DIST)	heaven=LOC

ka-n-t
go-FUT-3sSBJ

'the man God chose, (that one) will go to heaven' (64)

(12) *tanta quis cayç mimà poòp*
el pan que estoy comiendo o que como, este es blanco

tanta	kis	kai-ts	mima	p'o-p
bread	1sPRO	eat-1sSBJ	ANA.DIST	be.white-3sSBJ

'the bread I am eating, that one is white' (65)

(13) *talà quis ceinicè, mimà palap*
el vestido que me diste es nuevo

tala	kis	tse-i	ni-tse	mima	palʔa-p
dress	1sPRO	give-THEM	P.REC.VIS-2sSBJ	ANA.DIST	be.new-3sSBJ

'the dress you gave me, that one is new' (65)

(14) *talà quis ceinicè, mimà palatòn*
el vestido que me diste es nuevo

tala	kis	tse-i	ni-tse	mima	palʔa-ton
dress	1sPRO	give-THEM	P.REC.VIS-2sSBJ	ANA.DIST	be.new-NMLZ

'the dress you gave me, that one is a new one' (65)

(15) *saycç meticapssè, mimà saycsnis vyè*
la carta que has de llevar o llevarás aun no está escrita

s-aiks-∅	metika-p-tse	mima	s-aiks[-i]
INS-write-NMLZ	carry-FUT-2sSBJ	ANA.DIST	INS-write[-THEM]

ni-s	uye
P.REC.VIS-1sSBJ	NEG

'the letter you will carry, that one I have not written' (65)

3.5. An unclear case: no relativization marking at all, or possession

Finally, the grammar section contains three potential examples (16-18) of null head relativization without either the agent nominalizer -*ton*, nor the resumptive pronoun *mima*. But the verbs following *Dios* could conceivably be analyzed as possessed nouns, as in the

Latin model, as shown in (17). However, only *olom-* 'word' is attested elsewhere as a noun.

(16) *Dios yelem ti*
    fac secundum legem Dei (Latin), *haz lo que Dios dice o manda*
    Dios   yelem[-p]   ti[-i]
    God   command[-3sSBJ]   do[-THEM]
    'do what God commands!' (78)

(17) *Dios olómp ti*
    fac secundum legem Dei (Latin), *haz lo que Dios dice o manda*
    Dios   olom-p   ti[-i]   (alternatively: *olom-p)*
    God   speak-3sSBJ do[-THEM] (alternatively: word-3sPOSS)
    'do what God says!' (78)

(18) *Dios nequequesp ti*
    fac secundum legem Dei (Latin), *haz lo que Dios dice o manda*
    Dios   ne-kekes-p   ti[-i]
    God   INS-command-3sSBJ   do[-THEM]
    'do what God commands!' (78)

3.6. Discussion

The following is a table of morphosyntactic features of the 18 examples presented so far:

Example(s):	Overt head sentence-initially:	Head is subject of the main clause:	NP position relativized:	Strategy:
(1-3, 7)	yes	yes	Subject	NMLZ on verb
(4-6)	no	no	Subject	NMLZ on verb
(8)	yes	yes	Object	NMLZ on verb
(9)	no	yes	Subject	NMLZ on verb AND *mima*
(10-11)	yes	yes	Object	*mima* OR no marking
(12-14)	yes	yes	Object	*mima*
(15)	yes	no	Object	*mima*
(16-18)	no	no	Object	no marking

Tentative conclusions regarding these 18 examples are the following. I suggest that the nominalizer is the preferred strategy when the NP position relativized is Subject. Example (8) is the only counterexample to this, but then example (8) is suspicious in any case, because it is an example with the nominalizer following a verb inflected for subject person, so it is hard to know what to make of this. If sentence (8) actually contained a passive, and could be translated as 'the man who will be punished by me…', it would no

longer be a counterexample, but then of course I would have to argue that the 1st person subject inflection on 'punish' is actually something else.

I also suggest that *mima* is the preferred strategy when the NP position relativized is Object, unless, as shown in example (9), there is no overt head sentence-initially and the head is the subject of the main clause, in which case the nominalizer is also present. It is possible that the resumptive pronoun strategy is due to Spanish influence, as is the case with the resumptive demonstrative pronoun *chay* in Quechua relativization (R. Cerrón-Palomino, p.c.).

I will not further comment on the 'no marking' strategy, illustrated by (16-18) alone, because these three sentences are close variants of each other, clearly considered just one example by Maccioni, and they might not be cases of relativization at all. Let us note for now that sentence (18) bears a similarity to part of sentence (9)..

4. Relative clause formation in Machoni's (1732) texts section
4.1. Introduction
I counted 33 sentences that can be interpreted as containing relative clauses in Machoni's (1732) texts section. Only a few examples will be provided here, as the texts are quite repetitive. The texts, which are separately paginated, are all parts of 18th century Catholic Church doctrine, including prayers, and two catechisms. Except for the second catechism, no Spanish translation is given by Maccioni, which shows that he expected the Jesuits who were to use his book to be familiar with the texts. So in the examples below, I do not provide a Spanish gloss.

4.2. Examples of relative clause formation in Machoni's (1732) texts section
The texts section shows, overwhelmingly, the relative clause strategy with the agent nominalizer *-to(n)*, (examples 19-21) and also the strategy with no marking at all (examples 22-23), but significantly, I think, no example of the resumptive pronoun *mima* at all.

(19) *Dios ué ayopép Reyna Umuè auei titó,*

Dios	*we*	*ayo-pe-p*	*Reyna*	*Umue*	*ʔawei*	*ti-to*
God	2sPRO	greet-FUT-3sSBJ	Queen	Mother	compassion	have-NMLZ

'May God greet you, Queen (and) Mother who has compassion'
(Salve Regina prayer)

(20) *Yaupecelé yauptó uyáp cepssé.*

yaʔupe-tse=le	*yaʔu-p-to*	*uya-p*	*tse-p-tse*
village-2sPOSS=LOC	walk-3sSBJ-NMLZ	house-3sPOSS	give-FUT-2sSBJ

'you give his house (to) the traveller (lit. one who walks to your village)'
(Works of Mercy)

(21) *Uaton, ueciton tayule Christianospan Dios uasipssé.*

wato-[to]n	*wetsi-ton*	*taʔyule*	*Christianos-pan*	*Dios*
live- NMLZ	die- NMLZ	BEN	Christians-COLL	God

waʔas-ip-tse
pray-FUT-2sSBJ
'you will pray to God for Christians who are living (and) who are dead'
(Works of Mercy)

(22) *lopsauy eycupticén, mequequetó uá ticazpan lopsáucén:*
lopsawi[-i]   *eikupti-tsen*   *meke[•ke]-to*   *wa*
forgive[-THEM]   sin-1pPOSS   be.like[•RED]-NMLZ   1pPRO

tik-as-pan   *lopsawi-tsen*
INS-get.angry-3pSBJ   forgive-1pSBJ
'forgive our sins as we forgive those who offend us (lit. make us angry)' (Lord's
Prayer)

(23) *Tamócx Dios Pé, zó, há tiatép;*
takmoks-ts   *Dios*   *Pe*   *tso*   *ʔa*   *ti[-i]*   *ate-p*
believe-1sSBJ   God   Father   heaven   earth   make[-THEM]   P.REM-3sSBJ
'I believe in God the Father, who made heaven (and) earth' (Credo)

But in addition, there is quite a common construction not discussed in the grammar, with the pronoun *usa* 'what' i.e. 'that which', which in the grammar occurs only once, in sentence (9) above. Examples in the texts are quite common, and include (24-25):

(24) *Usa Christiano tamocxynt ysyauomp locuép*
usa   *Christiano*   *takmoks-in-t*   *is*   *yawon-p*   *lokwe-p*
what   Christian   believe-FUT-3sSBJ   hand   be.all-3sSBJ be.four-3sSBJ
'the things that the Christian will believe (are) fourteen (lit. all hands and four)'
(Articles of the Faith)

(25) *Pelé eciptó zó cent ussiqué, usá Dios nequequesp tyatepán;*
pele   *etsi-p-to*   *tso*   *tse-n-t*
man   be.good-3sSBJ-NMLZ   heaven   give-FUT-3sSBJ

usike   *usa*   *Dios*
because   what   God

ne-kekes-p   *ti[-i]*   *ate-pan*
INS-command-3sSBJ   do[-THEM]   P.REM-3pSBJ
'He will give heaven (to) the man who is good, because they have done what
God commands' (Articles of the Faith)

4.3. Discussion

What one notices in the texts section, is more unusual syntax, possibly calqued on Spanish or maybe on Latin. For example, in (21), the benefactive postposition *taʔyule* 'for', precedes *Christianos*, although everywhere else, Lule postpositions follow their object. But it is also possible that *taʔyule* heads the preceding nominalized verbs *wato[to]n* and *wetsiton*, thus meaning 'for the living and the dead', in which case the placement of *Christianospan* is odd. Also, since the language tends to be SOV in basic constituent order, the initial 'I believe' in example (23), must follow the word order of the Latin 'credo'.

Another unusual thing, occurring often in the texts, is the nominalizer following an inflected verb, as in *yaʔu-p-to* in example (20) and *etsi-p-to* in example (25). This also occurs in the grammar, as shown in example (8), but example (8) is the only occurrence in the grammar.

It is also suspicious that there are no examples of the resumptive pronoun *mima* in the texts. This is presumably because a resumptive pronoun sounds clumsy or colloquial in Spanish (and in Latin). Look again at examples (10-15). It is possible that Maccioni did not use *mima* in the texts to improve on their style. After all, these are religious texts, and should not sound colloquial.

Finally, it is also suspicious that there is only one example of *usa* 'what, that which' in the grammar (example 9), although this *usa* is very common in the texts; two examples are (24) and (25). It is also instructive to compare *usa Dios nekekesp ti* 'do what God commands' in example (25), with *Dios nekekesp ti* 'do what God commands', which is example (18) from the grammar section. So one wonders if this *usa* is really necessary. I suggest that this *usa* is a calque on Spanish 'lo que'. In the grammar, *usa* is simply the interrogative pronoun 'what?'. I suggest, then, that the portion *usa Dios nekekesp* in example (9), is taken out of the texts, although it is preceded by a portion containing *mima*, which sounds more colloquial. Example (9) is interesting then, as it shows characteristics of a more natural syntax evidenced in the grammar, as well as a characteristic of the more stilted syntax found in the texts.

In general, the syntax (and maybe the morphology) of Machoni's (1732) religious texts is more stilted and somewhat less reliable than the examples in the grammar, presumably because of a requirement that religious texts must follow the formal Spanish (or Latin) models closely.

5. Conclusions

I suggest that it is possible to draw tentative conclusions about Lule relative clauses from this very small corpus, following the terminology of WALS (Comrie and Kuteva 2013a, b), and ordering them from the most genuine strategy (i.e. likely to be original Lule) to the least genuine strategy (i.e. likely to be entirely calqued on Spanish or Latin). The most genuine strategy for relative clause formation in Lule is likely to be the gap strategy, normally with a nominalizer at the end of the clause. Maybe less genuine is the non-reduction strategy, in the guise of the paratactic relative clause, with a resumptive pronoun, which might well be from Spanish colloquial influence. Least genuine are the relative clauses with the head *usa*, since these only occur in the texts, and are quite likely the result of a calque on Spanish.

Let me conclude further with some observations regarding the validity of philological interpretations of early missionary writings of extinct languages. Needless to say, the model of Latin grammar is always prevalent among the early missionaries, but beyond that, one can detect innovative ideas, indigenous voices, and, crucially for this paper, places where the missionary accurately observed actual language, and places where he thought it was better to calque the language on Spanish (or Latin) models. I suggested that this happened in the Lule texts, but less so in the Lule grammar section. The moral of all this is, of course, that when studying missionary grammars, we need to make careful comparisons between example forms in the grammatical discussions and forms from texts. We may not assume, as we may safely in later documentary traditions, that the texts represent a more natural language than examples in the grammatical discussions.

6. Bibliographical references

Comrie, Bernard, and Tania Kuteva. (2013a.) Relativization on Subjects. In Matthew S. Dryer, & Martin Haspelmath, (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/122, Accessed on 2015-04-21.)

Comrie, Bernard, and Tania Kuteva. (2013b.) Relativization on Obliques. In Matthew S. Dryer, & Martin Haspelmath, (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/123, Accessed on 2015-04-21.)

Machoni, Antonio. 1732. *Arte y vocabulario de la lengua lule y tonocoté*. Madrid: Herederos de Juan García Infanzón.

Linguistics Program
University of North Texas
College of Information
3940 North Elm St., C232
Denton, TX 76207-7102
USA