Conor Stuart Roe Brook Lillehaugen Structure of Colonial Valley Zapotec Final Project 22 Dec 2017

DISTRIBUTION OF THE ZAPOTEC COVERT SUBJECT CONSTRUCTION

Abstract

This paper investigates the distribution of the covert subject construction, an unusual syntactic construction appearing in several Zapotec languages including Colonial Valley Zapotec (CVZ). A covert subject is the subject of a sentence which is coreferential with the possessor of the object and is not overtly stated; where the subject is coreferential with the possessor of the object and is still present it is termed a redundant subject. Given its presence throughout diverse branches of the Zapotec family, suggesting that it originates with Proto-Zapotec, and its conflict with supposed linguistic universals, the covert subject construction is of high interest to syntactic study.

Across a dataset of seventeen sentences from seven documents, spanning roughly 200 years of written Zapotec history, the covert subject construction displays an even distribution with regard to clause types and historical periods. The data are slightly suggestive of a relationship between covert subject and possession strategy, and/or of a language contact effect in documents translated from Spanish into Zapotec, but also appear consistent with the covert and redundant subject constructions existing in free variation. Further research should investigate covert or redundant subject sentences with free possessors or body part objects.

Introduction and literature review

A number of languages in the Zapotec family, including Colonial Valley Zapotec (CVZ), of the Central Valley branch, and multiple modern varieties of other branches, make use of a typologically unusual syntactic construction known as the "covert subject" (Avelino et. al. 2004). The distribution of the construction within the family strongly suggests that it is reconstructable to Proto-Zapotec, and that it has been lost from most modern varieties, including all modern Central Valley branch varieties (Avelino et. al. 2004 1). Closer study of the syntactic and diachronic distribution of the covert subject construction may reveal interesting insights about the typology of the family.

In its essence, the covert subject construction describes clauses in which the subject is coreferential with the possessor of the object, and the subject has been omitted and appears only as a possessor modifying the object.

```
VERB (SUBJ) OBJ=POSS

1. ...baa pe-zooba - s-tolla=lo
already PERF.CAUS-be.located POSS-sin=2S
'Have you already confessed your sins?'
(Levanto 1732 p. 102:4, analysis Lillehaugen et. al. 2017b)
```

Note that the covert subject construction does not appear to be mandatory when the subject and the possessor of the object are coreferential.

```
VERB=SUBJ OBJ=POSS

14. ...na-pa=ya yoholichi=a

STA-have=1S house=1S

'I have my houses'

(AGN 1614 p. 1:28, analysis Munro 2011:44)
```

I refer to constructions such as 14., in which a coreferential subject and possessor of the object both overtly appear, "redundant subjects," which is not to suggest that the overt use of the subject does not convey anything meaningful. The syntactic distribution, and potential implication in meaning, of the choice between covert and redundant subject constructions has not

yet been determined. Previous literature (Avelino et. al. 2004, Munro 2011) has mainly been concerned with describing the covert subject construction and defending the notion that it truly lacks a subject.

Accepting that conclusion and moving forward, I attempt here to investigate the question of what syntactic, language contact, or temporal factors determine the choice between covert and redundant subject. To this end, I utilize a dataset of covert and redundant subject constructions appearing across seven documents, ranging from a 1565 will to Leonardo Levanto's 1766 *Cathecismo*.

Data Analysis

Possession Marking

The first criteria of interest is possession marking strategy. CVZ utilizes three different strategies to indicate possessive and genitive relationships, which may differently affect the grammaticality of the covert subject.

Possession in 1. is marked with a "direct alienable" (Galant 2011b) construction. The enclitic pronoun **-lo** '2s' is syntactically independent, and identical in form to the second person pronoun that would appear as an enclitic subject. The alienable possession relationship is indicated with the prefix typically spelled **x-** or **xi-** (spelled **s-** in 1.), which marks that the following noun is a possessum and the noun or pronoun after is its possessor.

Possession is marked in 14. via a "direct inalienable" construction. In CVZ, body parts and kinship terms, as well as a handful of other nouns like **yoholichi** 'house,' are obligatorily possessed, and the possessor is simply indicated by juxtaposition after the possessum. If the possessor is pronominal it takes the same clitic pronoun form as in direct alienable possession or other grammatical uses, like –(y)a '1s', which appears as both a subject and a possessor in 14.

The last possession marking strategy is the "periphrastic" construction (Galant 2011b), which uses the preposition **xitene** 'of' between the possessum and possessor to indicate a genitive relationship, as in 11.

Xitene (and its orthographic variants, including **xteni** and **xitini**), like a verb or directly possessed possessum, may be followed by either an unbound noun or pronoun or a phonologically bound enclitic pronoun.

Plotting the distribution of covert and redundant subject constructions according to possession strategy reveals a dataset without clear tendencies but some suggestive patterns. In the tables below, the numbers correspond to numbered sentences given in Appendix B:

	covert subject	redundant subject		
direct inalienable	2, 4	14, 17		
direct alienable	1, 3, 6, 7, 8, 9			
periphrastic	5, 10, 11, 12	13, 15, 16		

	covert subject	redundant subject		
bound possessor	1, 2, 3, 5, 6, 7, 8, 9, 11, 12	13, 14, 15, 16, 17		
free possessor	3, 4, 10			

One pattern of note is that redundant subject constructions do not appear in my dataset with direct alienably possessed objects, and that covert subject constructions are relatively rarer with direct inalienably possessed objects. Perhaps this can be semantically justified – a body part or member of kin, closely associated with the possessor, may be less likely to be acted on by a third party, in which case gapping the subject may introduce less ambiguity about the agent. However, the data don't seem strongly skewed enough to be certain that the pattern isn't spurious.

Additionally, redundant subjects do not appear in my dataset with free possessors. This again cannot be ruled out as a random artifact given the small number of free possessors in the dataset. However, it seems justifiable phonologically – since free possessors are usually polysyllabic, repeating them in subject position might sound especially repetitive and thus be dispreferred. One variant of the redundant subject which has not to my knowledge been seen, but would have interesting implications for this hypothesis were it to be found, is a construction in which a coreferent subject and possessor are both overtly stated, but one appears as a bound pronoun while the other appears as an unbound pronoun or full noun phrase.

Clause Type

Another variable of potential relevance to the distribution of the covert subject is clause type. I distinguish four types: indicative clauses, which may be statements with any of several aspectual/modal prefixes on the verb; interrogative clauses, which are questions with a similar variety of possible aspectual and modal marking; exhortative, which is always marked for irrealis mood in an independent clause; and "dependent," which includes adverbial and embedded clauses marked for irrealis mood. The overwhelming majority of my data were classified as indicative or interrogative.

Given the data so far, there is no clearly emergent relationship between clause type and construction choice:

	covert subject	redundant subject		
indicative	4, 5, 7, 9, 10, 11, 12	14, 15, 16		
dependent	3	13		
exhortative	2			
interrogative	1, 6, 8	17		

The amount of data for dependent and exhortative clauses is certainly too small to rule out exhortative redundant subjects. The contrastive distribution of covert and redundant subject constructions in all other clause types, with no noticeable skew, suggests that clause type is not a determining factor in construction choice.

Original Language

Another factor to consider is whether the source of a sentence was written originally in Spanish and then translated or written originally in Zapotec, to account for possible second-language effects. Many of the existing documents in CVZ were written by Spaniards translating European documents or concepts, usually religious, into the indigenous language, so there is a significant quantity of data to inspect for language contact effects.

There is a noticeable absence in my data of redundant subjects in documents translated from Spanish:

	covert subject	redundant subject
originally Zapotec	4, 5, 10, 11, 12	13, 14, 15, 16
originally Spanish	1, 2, 3, 6, 7, 8, 9	17

An explanation of this as a legitimate language-contact effect is difficult. A priori, I might expect the opposite – an increased frequency of redundant subjects in documents translated from Spanish if the author were more directly calquing – since Spanish obligatorily marks subject.

A possible interactive effect might appear when translating from Spanish sentences in which the possessum is a body part, since Spanish does not typically modify object body parts with possessors, but rather puts the possessor as a dative experiencer. Perhaps in such cases, an author translating into Zapotec especially literally would avoid a coreferential possessor on the body part, qualifying the sentence as neither a covert nor redundant subject construction. Such an effect would likely contribute to gaps in my data.

To further investigate this question, a dataset of sentences translated from Spanish, with body part objects, would be needed. Strategies for avoiding a coreferential possessor might be various, since CVZ body parts generally require some possessor.

Time

A last criteria to consider is historical time. I have data from documents written in the 16th, 17th, and 18th centuries, giving a reasonably broad span of time to examine for diachronic effects. Since the covert subject construction has disappeared in all modern Central Valley Zapotec languages, it presumably began to fade from spoken language some time during or shortly after the period from which we have CVZ documents, so it might be expected to become less frequent in written CVZ as well over time.

Unfortunately, for my dataset, time is difficult to extricate from original language – all of my 17^{th} century documents are originally Zapotec, and all of my 18^{th} century documents are originally Spanish.

	covert subject	redundant subject
16 th century	2, 10, 11	13
17 th century	4, 5, 12	14, 15, 16
18 th century	1, 3, 6, 7, 8, 9	17

Overall, no significant trend appears. There are examples of each construction from each century, with somewhat even distribution. Certainly, with six examples of covert subjects in the

18th century, the distribution of the data suggests that the covert subject remained firmly established in written CVZ through the end of the written record. If it was already disappearing from spoken Valley Zapotec at the time, diglossia seems to have kept it present in writing.

Conclusion

No unmistakable distributional trends appear in my data. To the contrary, they seem to provide reasonably strong evidence that clause type and historical period have little to no effect on the choice between the two constructions. My data offer slightly suggestive evidence that possession strategy and original language may affect the choice of construction, but either claim would require an augmented dataset to convincingly claim. In particular, covert subjects seem to possibly be preferred with free possessors or direct inalienable possession, or when translating from Spanish. To further investigate these possibilities, more data with free possessors or body part objects would be useful.

Across all criteria, both covert and redundant subject constructions appear in diverse enough conditions that any factor affecting the choice between covert and redundant subjects seems likely to be statistical at best, rather than ruling one or the other absolutely ungrammatical. Given this finding, the possibility should be taken seriously that covert and redundant subjects are in totally free variation, with any tendencies in my dataset due simply to statistical chance.

Scholarly References

- Avelino, Heriberto, John Foreman, Pamelo Munro, and Aaron Huey Sonnenschein. 2004. "Covert Subjects in Zapotecan." SSILA annual meeting, Boston.
- Lillehaugen, Brook. 2004. "Is Valley Zapotec *lohoh* a dative marker?" First Conference on Otomanguean and Oaxacan Languages, Berkeley.
- Lillehaugen, Brook Danielle, Claire Benham, Janet Chavéz Santiago, Emily Drummond, James Arthur Faville, Avery A. King, Bridget Murray, Tristan Jacobo Pepin, May Helena Plumb, Mindy Renee Reutter, James E. Truitt, Christina Nicole Ulowetz, Mike Zarafonetis & Ian Fisher. 2017. Digital edition of Fray Leonardo Levanto's 1766 Cathecismo de la lengua Zaapoteca, beta version. Ticha Project: https://ticha.haverford.edu/en/texts/levanto/
- Lillehaugen, Brook Danielle, Sabea K. Evans, Kathryn Goldberg, Julie Gonzales, Jaime Metzger, Lyra Piscitelli, Diamond C. J. Ray, and Conor Stuart Roe. 2017. Digital Edition of pp. 97-117 of Leonardo Levanto's 1732 *Arte de la lengua zapoteca*. Online: https://ticha.haverford.edu/en/texts/levanto-arte/98/original.
- Munro, Pamela. 2011. Evidence about Proto-Zapotec from a Colonial Document. Munich: Lincom Europa.
- Munro, Pamela, Kevin Terraciano, Xóchitl Flores-Marcial, Michael Galant, Brook Danielle Lillehaugen, Maria Ornelas, Aaron Huey Sonnenschein, and Lisa Sousa. 2018; to appear. The Zapotec-language testament of Sebastiana de Mendoza, c. 1675. *Tlalocan XXIV*.

Primary Sources

- 1565. Archivo General de la Nación (AGN), Tierras, vol. 241, exp. 7.
- 1614. Archivo General de la Nación (AGN), Tierras, vol. 256, exp. 2, ff. 1-165v.
- 1626. Archivo General de la Nación (AGN), Tierras, vol. 256, exp. 2, p. 4:1-14.
- 1675. Archivo General de la Nación (AGN), legajo 42, exp. 10, ff. 7r-7v.
- Feria, Pedro de. 1567. *Doctrina christiana en lengua castellana y çapoteca*. Mexico: En casa de Pedro Ocharte.
- Levanto, Leonardo. 1732. Arte de la lengua zapoteca. Oaxaca.
- Levanto, Leonardo. 1766. *Cathecismo de la lengua zaapoteca* [sic]. Reimpresso en la Imprenta del Colegio Real de San Ignacio de la Puebla de Los Angeles.

Appendix A - Abbreviations used in glossing

PERF perfective aspect

HAB habitual aspect

STA stative aspect

CAUS causative voice

IRR irrealis mood

DFUT definite future tense

NOM nominalizer

POSS possession marking

1s first person singular pronoun

2s second person singular pronoun

2P second person plural pronoun

3 third person pronoun

Appendix B – Data

Covert Subjects

•	1	1. 1.1	1 1			11	a · 1
interrogative,	direct	alienable	haund	nronolln	nocceccor	Originally	/ Snanich
michoganve,	uncci	andiaoic,	oouna	pronoun	possessor,	originan	opamsn .

VERB (SUBJ) OBJ=POSS

1. ...baa pe-zooba - s-tolla=lo
already PERF.CAUS-be.located POSS-sin=2s
'Have you already confessed your sins?'

(Levanto 1732 p. 102:4, analysis Lillehaugen et. al. 2017b)

exhortative, direct inalienable, bound pronoun possessor, originally Spanish

VERB (SUBJ) OBJ=POSS

2. **chela ya qui-tona - lechela=to**and not IRR-leave spouse=2P
'And you are not to leave your spouses...'
(Feria 1567 p. 5:1-2, analysis Avelino et. al. 2004:10)

dependent, direct alienable, both types of possessor, originally Spanish

VERB (SUBJ) OBJ=POSS

3. ...nijani c-o-xooba - (o)x-tolla=ni benni quicharij in.order IRR-CAUS-be.located POSS-sin=3 person sick this '...so that this sick person may confess (place) their sins.'

(Levanto 1732 p. 102:7, analysis Lillehaugen et. al. 2017b)

indicative, direct inalienable, noun phrase possessor, originally Zapotec

VERB (SUBJ) OBJ POSS

4. ...r-o-caa - laa testigo lao quichi rijni anachij ju[eves].

HAB-CAUS-be.stuck name witness face/on paper this today Thursday

'...the witnesses write their names on this paper, today, Thursday.'

(AGN 1626 p. 4:14, analysis Munro 2011:46)

indicative, periphrastic, bound pronoun possessor, originally Zapotec

VERB (SUBJ) OBJ POSS

5. **r-o-gago - anima xitini=a dios xi-b[ejuana]=ya**HAB-CAUS-eat soul POSS=1S God POSS-lord-1S

'I feed my soul to God my lord...'

(AGN 1614 p. 1:13, analysis Munro 2011:40)

interrogative, direct alienable, bound pronoun possessor, originally Spanish (SUBJ) OBJ=POSS VERB 6. coca-xaa x-tolla=lo? pe-zooba when-EMPH PERF.CAUS-be.located POSS-sin=2s 'Have you confessed your sins?' (Levanto 1732 p. 115:12, analysis Lillehaugen et. al. 2017b) indicative, direct alienable, bound pronoun possessor, originally Spanish (SUBJ) OBJ=POSS **VERB** 7. Chij pe-zooba x-tolla=lo day/when PERF.CAUS-be.located POSS-sin=2s 'When you confessed your sins...' (Levanto 1732 p. 117:4, analysis Lillehaugen et. al. 2017b) interrogative, direct alienable, bound pronoun possessor, originally Spanish VERB OBJ=POSS (SUBJ) 8. c-o-tijxe-lee-chahui **Bixooze Coopa Bitoo?** quiraa x-tolla=lo loo PERF-CAUS-exposed-well all POSS-sin=2s face/to father guard divine "...did you declare all your sins to the priest?" (Levanto 1732 p. 117:4-5, analysis Lillehaugen et. al. 2017b) indicative, direct alienable, bound pronoun possessor, originally Spanish VERB SUBJ VERB (SUBJ) OBJ=POSS 9. ri-cabi-lachi=a x-tola=ya gue-xooba HAB-accept-heart=1s IRR-be.located POSS-sin=1s 'I agree to confess my sins' (Levanto 1732 p. 110:8, analysis Lillehaugen et. al. 2017b) indicative, periphrastic, free pronoun possessor, originally Zapotec (SUBJ) OBJ VERB POSS 10. t-e-neche firma quitaa tono xitene HAB-1P-give signature of all 1 p 'We all give our signatures.' (AGN 1565 lines 15-16, analysis Oudijk 2008:234) indicative, periphrastic, bound pronoun possessor, originally Zapotec VERB (SUBJ) OBJ POSS 11. niacanj c-o-lo firma xitene=a of=1sIRR-CAUS-put signature SO

'And so I write my signature.'

(AGN 1565 lines 23-24, analysis Oudijk 2008 p. 235)

indicative, periphrastic, bound pronoun possessor, originally Zapotec

VERB (SUBJ) OBJ POSS

12. **ti-baqui - āīā xitini=a neza quela=na-bani çelij**HAB-put soul of=1s road NOM=HAB-live eternal

'I put my soul on the road of eternal life.' (AGN 1614 line 10, analysis Munro et. al.)

Redundant Subjects

dependent, periphrastic, bound pronoun possessor, originally Zapotec

VERB OBJ VERB=SUBJ OBJ POSS

13. **çij=nj yoo-cani cu-aqui=nj enstaçia xitene=nj**IRR.receive=3_i land-this IRR-put=3_i estancia of=3_i

'[they want to] ...take the land to put their estancia on it.'

(AGN 1565 lines 4-5, analysis Oudijk 2008:232)

indicative, direct inalienable, bound pronoun possessor, originally Zapotec

VERB=SUBJ OBJ=POSS

14. ...na-pa=ya yoholichi=a

STA-have=1S house=1S

'I have my houses'

(AGN 1614 p. 1:28, analysis Munro 2011:44)

indicative, periphrastic, bound pronoun possessor, originally Zapotec

VERB=SUBJ OBJ POSS

15. r-apa=ya quela=ri-jeni quela=r-acapea xteni=ya

HAB-have=1s NOM=HAB-understand NOM=HAB-know of=1s

'I have my understanding and knowledge.'

(AGN 1675 lines 2-3, analysis Munro et. al. 2018)

indicative, periphrastic, bound pronoun possessor, originally Zapotec

OBJ POSS VERB=SUBJ

16. anima xteni=a ri-go=ya lachi ña Bexuana=na Dios soul of=1s HAB-put=1s heart/in hand Lord=1P God

'I put my soul in the hands of our lord God.'

(AGN 1675 line 11, analysis Munro et. al. 2018)

interrogative, direct inalienable, bound pronoun possessor, originally Spanish

SUBJ VERB=SUBJ OBJ=POSS OBJ OBJ=POSS 17. Yoobi Dios hua=r-apa=ni quique=ni, niia, ñaa=ni... INT=HAB-have=3_i same Godi $head=3_i$ foot $hand=3_i$ 'Does the same God have a head, feet, and hands?' (Levanto 1766 p. 21:3-4, analysis Lillehaugen et. al. 2017a)