

Pre-publication version, to appear in *Estudios de lenguas amerindias 3*, Zarina Estrada Fernández, ed. Hermosillo: Editorial Unison. 2015.

## A SOUTHERN ZAPOTEC CONTRIBUTION TOWARDS THE TYPOLOGY OF INCLUSORY CONSTRUCTIONS\*

Rosemary G. Beam de Azcona  
Escuela Nacional de Antropología e Historia

Resumen: La literatura sobre las “construcciones de conjunto” (inclusory constructions) en lenguas australianas y austronesianas describe el uso obligatorio de un pronombre inclusorio. Lenguas zapotecas de la Sierra Sur carecen de número gramatical, aún en el sistema pronominal, y por esto las construcciones de conjunto tienen otra estructura en estas lenguas. Los pronombres pueden ser opcional o ausentes y los cuantificadores se vuelven elementos obligatorios para contar el conjunto de referentes al cual se refiere la construcción. Este trabajo propone un desarrollo diacrónico para este tipo de construcción en el zapoteco de la Sierra Sur, y la gramaticalización del término de conjunto hacia un comitativo. Otro elemento poco común de la versión de la construcción que se ve en la Sierra Sur es el uso obligatorio de una frase nominal poseída para referir al subconjunto. Las variaciones de la construcción en lenguas zapotecas de la Sierra Sur señalan la necesidad de considerar cómo se indica el conjunto y los subconjuntos en las construcciones de conjunto en otras lenguas del mundo.

Palabras clave: construcciones de conjunto, zapoteco, número, cuantificadores, numerales, gramaticalización, tipología

Abstract: Literature on inclusory constructions in Austronesian and Australian languages describes an inclusory pronominal as an essential element of this construction type. Southern Zapotec languages lack grammatical number, including in the pronominal system, and because of this inclusory constructions in Southern

---

\* Data gathered by the author for this study were collected during fieldtrips sponsored by the Endangered Language Fund, UC Berkeley’s Survey of California and Other Indian Languages, the California Native Language Center, and La Trobe University. The first draft of this paper was written during a fellowship at the Research Centre for Linguistic Typology, later significantly revised during a fellowship at the UC Davis department of Native American Studies funded by the Endangered Languages Documentation Program. Students and colleagues at these various institutions have contributed helpful suggestions. In particular, Alexandra Aikhenvald, Cheryl Black, Stefan Dienst, Nick Evans, Alex François, Alice Gaby, Brian Joseph, Frank Lichtenberk, Ruth Singer, and anonymous reviewers provided helpful insights, data, and references. Any errors are the author’s sole responsibility.

Zapotec languages have quite different structures. Pronouns can be optional or even absent from these constructions and instead quantifiers become an essential element, indicating the superset that the whole construction refers to. This paper proposes a diachronic development for this construction type in Southern Zapotec languages, including grammaticalization of the superset word to a comitative marker. Another unusual characteristic of Southern Zapotec inclusory constructions highlighted here is the obligatory use of a possessed noun phrase to indicate the included subset. The Southern Zapotec variations on the IC highlight the need to consider how both supersets and subsets are indicated in IC's worldwide.

Keywords: Inclusory construction, Zapotec, numerals, quantifiers, number, grammaticalization, typology

## 1. Introduction

Inclusory constructions (IC's), e.g. strings like *we (with) Henry* meaning 'Henry and I,' are common in the world's languages, though the existing typological literature (Lichtenberk 2000, Singer 2001a & b, Bril 2004, Haspelmath 2007) mostly focuses on Australian and Austronesian<sup>1</sup> versions of the construction. In these languages a non-singular pronoun such as *we* is one of the required elements of the construction. Southern Zapotec (SZ) languages generally lack number marking even on pronouns and as a result quantifiers take on an increased functional load, including in one SZ version of IC in which a quantifier has completely replaced the non-singular pronoun that plays a key role in the better documented IC's of the world. This paper will contribute to the typology of this construction type by showing another possible structure for IC's in languages that lack number distinctions on pronouns.

A fuller explanation of the IC is provided in §1.1. The typological literature is reviewed in §1.2. §1.3 introduces the SZ languages and varieties used in this paper. §2 provides an overview of the IC in SZ varieties. Related quantifier constructions in SZ languages are considered in §3, and these relate to the grammaticalization paths for SZ IC's (§4).

### 1.1 *Inclusory constructions*

Many grammars and general works contain descriptions of IC's under terms such as "inclusory conjunction" (Haspelmath 2007), "syntactic associatives" (Corbett 2000), a special type of "part-whole NP" (Evans 1995), "non-singular associative" (François 2000) and "associative conjunction" (Dunn 1999). In Beam de Azcona (2014a) I use *construcciones de conjunto* as a Spanish equivalent.

IC's, shown in italics in examples throughout, are here defined as constructions containing at least one morpheme, word, or phrase referring to a larger group or "superset"<sup>2</sup> (cf. Singer 2001a&b; Gaby 2006), and additionally an NP referring only to a subset<sup>3</sup> of that group.

In (1) the first person dual exclusive is the superset and 'my brother' is the subset. The subset is *included* in the set of referents denoted by the superset.

---

<sup>1</sup> Aissen (1989) provides an in-depth formal study of Tzotzil IC's

<sup>2</sup> In Spanish I call this a *término de conjunto* (Beam de Azcona 2014a).

<sup>3</sup> In Beam de Azcona (2014a) *subconjunto*.

(1) Toqabaqita (Lichtenberk 2000)

*Kamareqa doqora-ku meki lae ma-i qusungadi.*  
1DU.EXCL brother-1S 1DU.EXCL.FUT go VENIT-at tomorrow<sup>4</sup>  
'I and my brother will come tomorrow.'  
(Lit. 'We two my brother we will come tomorrow.')

In (2) the first person non-singular (indicated as dual by marking on the verb) is the superset and 'my older sister' is the subset. The sister is one of but not the only referent of the first person non-singular.

(2) Kulina (Dienst 2006)

[[*o-kha asi*]<sub>NP-POS</sub> *tazaha-ni*]<sub>PP</sub> *i-kada-na.*  
1SG-ASSOC older.sister ECOM-F 1NSG-move.DU-FUT  
'I'm going to go together with my older sister.'  
(Lit. 'Together with my older sister we two will go.')

In (3) the third person plural object marker is co-referent with the third plural absolutive form of 'parents'. These make reference to a superset, 'the parents'. A subset of 'parents' is the singular absolutive noun 'mother'. The 'mother' is included in the superset but is not equal to the whole set of referents indicated by 'parents'. Note that it would presumably suffice to just mention 'the parents' and be done with it. In these constructions the subset is redundant and its separate mention gives it prominence.

(3) Chukchi (Dunn 1999)

*ʔeqe-lʔ-e [...]* *ya-nm-ə-lenat* *ətləy-ə-t* *əmmemə.*  
bad-NOM-ERG PERF-kill-ə-3PO parent-ə-3PABS mother.3SABS  
'Evil-doers killed the father and mother.'  
(Lit. 'The bad killed them, the parents (including the) mother.')

---

<sup>4</sup>Abbreviations used are as follows: 1=first person, 3=third person, A=animal, ABS=absolutive, ANC=animacy classifier, ASSOC=associative, C=known to speaker, CL=clause, COM = comitative, COMPL=completive, D=unknown or stranger, DU=dual, E=exclusive, ECOM=even comitative, ERG=ergative, F=feminine, FOC=focus particle, FUT=future, H=human, HAB=habitual aspect, I=inclusive, LOC=locative, NOM=nominalizer, NSG=non-singular, O=object, POT=potential mood, P=plural, PERF=perfective, POS=possessive, PR=progressive, R=respectful, S (in gloss line)=singular, S (following brackets)=subject of an intransitive clause, TEMP=temporal noun phrase, V=higher animate, VCS=subject of a verbless clause, VENIT=venitive, X=morpheme of unknown gloss.

In (4) the number ‘two’ enumerates a set of two people. The head of the following possessed noun phrase, ‘offspring’ is a subset of the group counted by the number ‘two’. As indicated in the translation, only one son accompanied the father to work, not two. The number ‘two’ refers to a pair of people that includes the father, here represented by a third person pronoun which ostensibly functions only as a possessor of ‘offspring’.

(4) Coatec Zapotec (Beam de Azcona 2004)

Mb-i'd          toóp    [xin          meé]<sub>NP-POS</sub>    ti'n<sup>5</sup>.  
 COMPL-come POT.two offspring    3HR          job  
 ‘He came with his son to work.’ (Lit. ‘The two his son came to work’)

The subset is often a possessed noun phrase, as in (1), (2), and (4), but not always, as ‘mother’ in (3). The term used to refer to the superset can vary. Just in the examples shown above, we find the superset indicated with a dual pronoun in (1), a bound non-singular subject marker on a verb in (2), a plural marked noun (and additionally a plural object marker on the verb) in (3) and a numeral in (4)<sup>6</sup>.

### 1.2. Previous work on inclusory constructions

IC’s have previously been analyzed as subtypes of comitatives and coordination (Schwartz 1998a & 1998b), such that one element was included in the other (see Ladusaw 1989 and Aissen 1989).

Lichtenberk (2000) introduced the term “inclusory construction,” arguing that IC’s are neither comitatives nor coordination, though they may be related to either and sometimes make use of the same markers. Using the Oceanic language of Toqabaqita, Lichtenberk established two typological parameters for categorizing IC’s. Firstly, in some IC’s the superset and subset together form a phrase, while in others the indicator of the superset may be some bound marker in a different phrase, such as an agreement marker on a verb. He called these “phrasal” IC’s and “split” IC’s respectively. Secondly, Lichtenberk contrasts “implicit” IC’s in which there is no overt marker of relation between the superset and the subset, as in (1), (3), and (4), with “explicit” IC’s in which such a marker does exist. Different types of overt markers identified by Lichtenberk (2000) and Singer (2001a) in explicit IC’s are comitative markers, as in (2), the coordinate conjunction ‘and’, ‘one of a group’ markers commonly found in Australian languages, and even special IC case marking.

Singer (2001a) is a typological study of IC’s in Australian languages. Towards the semantic analysis of IC’s she introduces the term “central member of a superset” to refer to

<sup>5</sup> The orthographic representation of tone has been updated here, using a different convention than the one being used at the time of the dissertation being cited. Now, glottalized vowels are written with an apostrophe *a'*, low tone with an unmarked vowel *a*, high tone with an accented vowel *á*, falling tone with an accented followed by an unaccented vowel *áa*, and rising tone with an unaccented followed by an accented vowel *aá*. The low-mid front lax vowel, formerly written as *ẽ* is here rendered as *eh*.

<sup>6</sup> Similar to SZ IC’s like that in (4), but not clearly the same phenomenon, is the use of family group classifiers in Yi languages (Bradley 2001). In these Tibeto-Burman languages family group classifiers (or in the case of Lahu, compound nominal) are usually formed by naming a dyad such as ‘mother-child’ or ‘father-child’ but in some cases can be truncated to simply ‘mother’ or ‘father’. These are used with numerals such that a string like 3 mother(-child) refers to a mother and two children.

“that referent which is already specified by the superset.” For example, the first person singular is automatically specified by all first person pronouns and the second person singular is always a referent of a second person pronoun. Third person pronouns do not have a central member since the reference of a third person singular is not fixed to a speech participant.

Bril (2004) focuses on Oceanic languages and suggests grammaticalization paths affecting IC’s over time. What she calls “appositive” (Lichtenberk’s “implicit”) IC’s may over time take on a comitative or coordination marker, rendering explicit IC’s. Bril also found that pronouns used to denote supersets could become reanalyzed as either coordination markers or comitative prepositions, as shown below for Mwotlap.

(5) Mwotlap (François 2000:388)

IC

Mayanag *kōyō* mo-gom.  
 chief 3DU PERF-ill

‘The chief (and his wife) are ill.’ (Lit. ‘The chief they-two are ill.’)

(6) Mwotlap (François 2000:262)

Coordination

imam *kōyō* tita mino  
 father 3DU mother my  
 ‘my father and mother’

### 1.3. Southern Zapotec languages

SZ is an areal-genetic grouping (Beam de Azcona 2014a, 2014b, in preparation) belonging to the Zapotecan branch of Otomanguean. SZ comprises an estimated 8 mutually unintelligible languages. For the present study I have considered IC’s from eight varieties of four SZ languages in a contiguous area we can think of as the core of the SZ region. The relative location of each variety considered is shown in Figure 1, which also introduces the abbreviations used for source varieties of examples throughout this article.

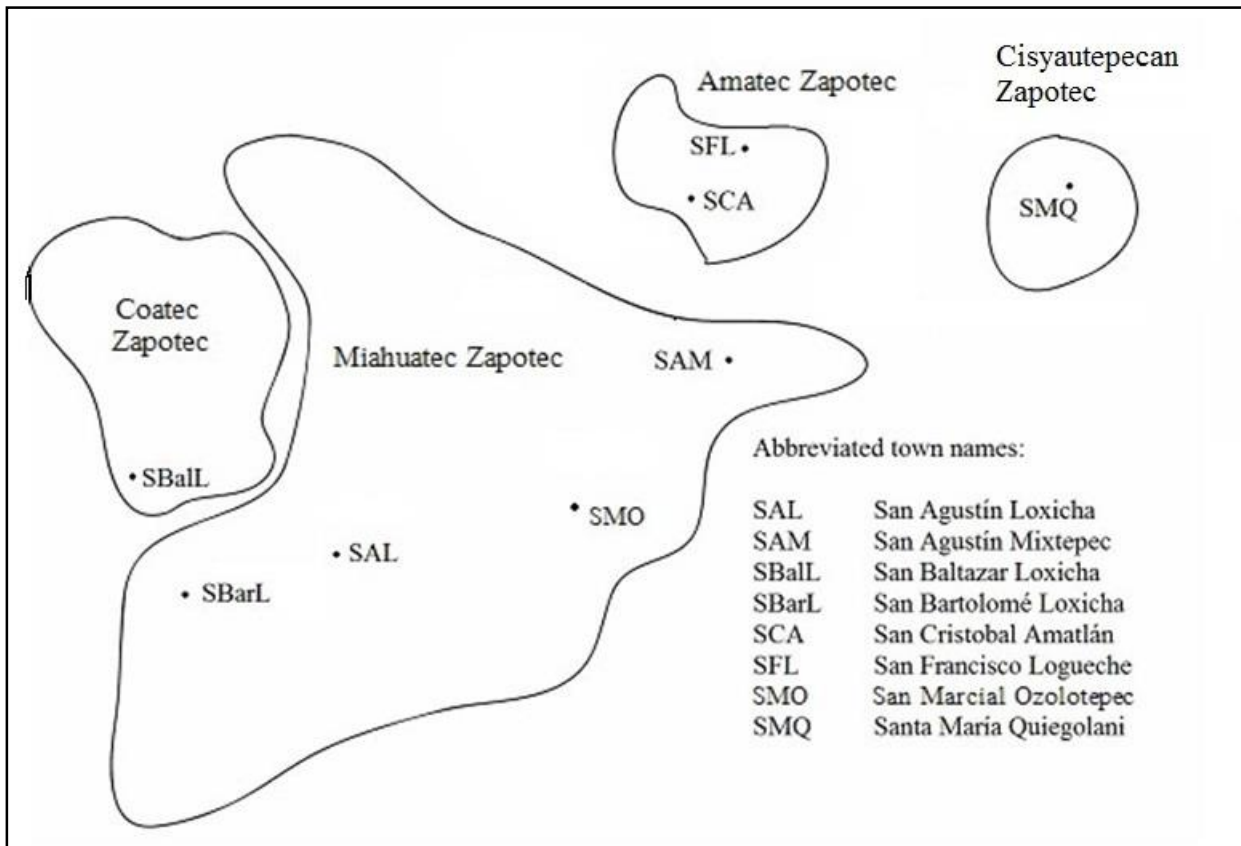


Figure 1: Relative location of varieties used in the present study

Data from Coatec and Miahuatec Zapotec come from the author's own fieldnotes. These were supplemented with published and unpublished sources for Cisyautepecan (Black 1994 & 2000) and Amatec Zapotec (Riggs 2005, n.d.; Angulo 1922-30).

Amatec and Coatec have been proposed to share a closer genetic affiliation (Smith Stark 2007, Beam de Azcona 2014b, in preparation) as Macro-Coatecan languages. Miahuatec and Cisyautepecan belong to separate genetic groupings. All these languages have been neighbors for several centuries, during which time they have converged with respect to certain aspects of their vocabulary and grammar. The inclusory constructions described in this paper are an example of one such diffused trait.

Another diffused SZ trait not found in other Zapotec groups is the lack of number marking. There are no plural affixes or clitics and even pronouns do not usually include number distinctions. Marlett and Pickett (2001) found that SZ languages were extreme in this regard compared to other Zapotec languages. Number can only be explicitly indicated through the use of independent quantifiers. While exceptions to this lack of number-marking exist in languages on the periphery of the SZ area, the generalization holds for all the varieties considered here.

The fact that independent quantifiers must be used to indicate the number of pronouns in SZ languages has serious implications for IC's since *virtually* all IC's described in the literature thus far have either an independent pronoun or a bound pronominal as the superset term. The superset of any IC necessarily has non-singular reference but in SZ

languages number is indicated almost<sup>7</sup> exclusively through the use of quantifiers, and therefore the superset of SZ IC's must be indicated, fully or partially, with a quantifier. Indeed, Central Zapotec (CZ) languages, which have pronouns with number distinctions, have IC's of a very different type, and more closely resemble IC's described for languages elsewhere in the world. Compare the CZ examples in (7) and (8) to the SZ example in (9).

- (7) Ca-da'uw=ënn gueht cëhnn Gye'eihlly. [SLQZ] (Munro 2000)  
 PROG-eat=1P tortilla with Mike  
 1. 'We're eating tortillas with Mike.' 2. 'I'm eating tortillas with Mike.'
- (8) ...b-zalloh ca-gyèi'ny=ënn zèèi'ny ri'cy cëhnn Beed. [SLQZ] (Munro 2000)  
 ...PERF-start PROG-do.1P=1P work there with Pedro  
 '...I (\*we) started working there with Pedro.'
- (9) Re y-oón [xi'n Pedr]<sub>NP-POS</sub> ngw-a zi'n. [SAL]<sup>8</sup>  
 all COL-three offspring Peter COMPL-go job  
 'Peter and his two sons went to work.'

Examples (7) and (8) from San Lucas Quiavini (SLQZ) have a plural subject pronoun indicating the superset and a comitative phrase indicating the subset. The SZ example in (9) has a quantifier phrase 'all three' referring to the superset while the possessed noun phrase 'Peter's children' refers to the subset. The possessor of that subset, Peter, is also included as a separate subset. In the CZ examples a plural pronoun indicates a non-singular entity, while in the SZ example the non-singularity is indicated only by the quantifier-numeral combination. In both instances this non-singular number counts the head of a later-occurring NP (the object of a loan preposition in SLQZ, the head of a possessed noun phrase in SAL) and additionally an entity indicated by other means--- the first person singular can be deduced from the use of the first person plural pronoun in SLQZ (Singer's *central member of the superset*) while the inclusion of the possessor is understood as a feature of this construction in SZ.

## 2. Overview of SZ IC's.

All SZ IC's contain, minimally, a quantifier and a possessed noun phrase. Both the possessor and the possessed noun count towards the number expressed by the quantifier. There are two main variations on the IC found in SZ languages, which I will refer to as nominal-initial and quantifier-initial. Whether or not these initial elements are actually heads of the phrases equivalent to the construction itself is an issue addressed further below. Both types of IC function as noun phrases. They are found in NP slots such as the core arguments of a clause or possessors of nouns (see Black 1994:326).

<sup>7</sup> There is a tendency for some speakers, probably due to contact with Spanish, to use the respectful second person pronoun with plural reference. Likewise, third person human pronouns in these languages include categories that indicate social status and there is a tendency for certain categories to be used more often when indicating a plural entity.

<sup>8</sup> The tonal analysis of this variety is not complete and so tone is marked only sporadically here, i.e. an unmarked vowel is not necessarily a low tone but can be a vowel whose tone is not reliably recorded in the data.

The Quiegolani variety (SMQ) of Cisyautepecan Zapotec lacks quantifier-initial IC's. It mainly has pronoun-initial IC's of the type shown in (10).

- (10) Tempran r-a-xee noo y-rup [xnaa noo]<sub>NP-POS</sub>. [SMQ Black 1994:342]  
 early HAB-go-rise 1E POT-two mother 1E  
 'Early my mother and I would get up.'

A minor variation also found in Quiegolani has an initial common or proper noun rather than a pronoun. According to Black (1994: 335), a proper name may function inclusorily in Quiegolani because, as in other Mesoamerican languages, a proper name can denote others associated with the named person, in addition to him/her, e.g. José could mean either José alone or 'José & co.'. According to Black's analysis, "proper names and common noun phrases become simply special types of third person pronouns."

- (11) [SMQ Black 1994:342]  
 Biki z-a g-un kompanyar *Gecha* y-rup [x-pëəd *Gecha*]<sub>NP-POS</sub>.  
 Virginia PR-go POT-do accompany Lucrecia POT-two POS-baby Lucrecia  
 'Virginia went to accompany Lucrecia<sub>i</sub> and her<sub>i</sub> baby.'

Nominal-initial IC's are noun phrases headed by their initial pronouns or nouns. (For arguments that nominal-initial IC's are indeed single constituents, see Black 1994:358-367). Coatec Zapotec has only quantifier-initial IC's, of the type shown in (12). This version of the IC usually has a numeral, or occasionally has a non-numeral quantifier, as the initial element.

- (12) Ngw-da téh [cheh' l ár]<sub>NP-POS</sub> má. [SBalL, Cazador.011]  
 COMPL-eat all spouse 3HF 3A  
 'He ate it with his wife (and family).'

This form of the construction is indeed a single constituent and functions in NP slots the same as nominal-initial IC's, but in these left-headed languages one expects the phrase-initial element to be the head and quantifiers are not nominals in these languages. An argument can be made that all or most Zapotec numerals, and perhaps non-numeral quantifiers as well, are verbs, or at least have a verbal origin historically (see the verbal analyses given in Córdova 1578:174ff and Black 1994: 339-349). The evidence for this is morphological and syntactic.

Zapotec numerals, and sometimes non-numeral quantifiers, are inflected for certain TAM categories usually found on verbs. For example, completive aspect marking on numerals can render ordinal meaning or can be used to quantify completed units of time, as in (13), or other nouns whose existence is established, i.e. their coming into existence is a completed event, as in (14), where the subject of 'two' is a relativized noun phrase. However, the full inflectional paradigm found with regular verbs does not exist for numerals in all languages. Miahuatec numerals do have a full paradigm or something close to it, but Coatec numerals have fewer forms than regular verbs.

- (13) Mb-rop=la [beés [nd-a-dé nh-kwi=noó xa' ná']<sub>REL-CL</sub>]<sub>NP-REL</sub>. [SBarL]  
 COMPL-two=already time PR-go-X X-trick=COM 3V 1S



‘S/he’s tricked me twice now<sup>9</sup>.’  
 (Lit. ‘The times that s/he has tricked me have been two already.’)

- (14) Mb-rop      [[xi'n      xa']<sub>NP-POS</sub>      xa-góot]<sub>NP-REL</sub>.      [SBarL]  
 COMPL-two      offspring      3V      3V-female  
 ‘Two of his children who are female have come into existence<sup>10</sup>.’  
 (Lit. ‘His children who are female humans are two in number.’)

When counting in Zapotec, numbers must be followed by nouns being counted. The closest equivalent to abstract numerical counting in European languages, as when teaching children using fingers, makes use of the impersonal pronoun encliticized to the number. This pronoun could be analyzed as a dummy subject, since the near-obligatory requirement for numbers to be followed by nouns is parallel to the near-obligatory occurrence of subjects with verbs. The order of these items is also the same. Just as verbs typically precede their subjects, quantifiers precede the nouns they quantify, whereas adjectives follow nouns (though this is also true of many languages, e.g. Romance, in which quantifiers are not verbs). Numerals can also be followed by adverbial clitics (see ‘already’ in [13]) which typically attach to verbs.

Morphologically, quantifiers, especially numerals, are inflected like verbs, sometimes with similar semantic effects (e.g. actions completed vs. units of time completed), though often the combination of TAM categories with quantifiers yields special semantics not natural with other verbal concepts. Syntactically, quantifiers require an argument the same as intransitive verbs. When looking at the internal syntax of the constituent headed by the quantifier, a verbal analysis seems well-justified. The main challenge for this analysis is the external syntax. Quantifier-initial phrases, i.e. phrases seemingly headed by quantifiers, often function in NP slots, e.g. the subject in (15).

- (15) Mb-rò'      [chòp      xà'gôt]s.      [SBarL]  
 COMPL-go.out POT.two 3V-female  
 ‘Two women went out.’

IC’s are just one type of construction involving quantifiers in SZ languages. In Quiégolani, the easternmost language considered here, quantifier constructions have an initial nominal head (see the first person exclusive pronoun in [10] and ‘Lucrecia’ in [11]), almost without exception.<sup>11</sup> In Coatec, the westernmost language considered here, the head nominal is always absent (in [12] note that no nominal precedes the quantifier ‘all’). In Miahuatec and Amatec both types of construction are possible, i.e. explicit nominal heads are optional. This variation is summarized in Table 1.

Coatec	Miahuatec & Amatec	Cisyautepecan
Quantifier NP <sub>SUBSET</sub>	(NP <sub>SUPERSET</sub> ) Quantifier NP <sub>SUBSET</sub>	NP <sub>SUPERSET</sub> Quantifier NP <sub>SUBSET</sub>

Table 2

<sup>9</sup> Spanish: ‘Van dos veces que me ha engañado’.

<sup>10</sup> Spanish: ‘Van (o 'ya son') dos de sus descendientes que son mujeres’.

<sup>11</sup>Although not an IC, there is an exception for a similar quantifier construction given by Black (1994: 368), see example (vii) of her footnote 26.

The problem of whether Zapotec numerals are verbs is a difficult one, and although the analysis of this part of speech has syntactic implications for the analysis of IC's, there is surely more to be said on the subject that can be stated here, and future analyses may well go in another direction. What can be stated confidently is that numerals in SZ languages have more verbal properties than in some other kinds of Zapotec. For example, some Valley Zapotec varieties do not inflect numerals and have no trouble counting with abstract numerals that lack overt arguments. Similar to the problem of relational nouns becoming prepositions, Zapotec numerals are surely in the process of deverbalizing and the process is further along in some languages than in others.

Because the IC functions in an NP slot, it is convenient to view the Cisyautepecan version of the IC with an initial head nominal as conservative, even though quantifier-initial phrases that aren't IC's also can fill NP slots. Regarding the Cisyautepecan form of the IC as conservative is also helpful when making typological comparisons since supersets are typically realized as pronouns in the typological literature on IC's. The analysis proposed here is that in SZ IC's, quantifiers head appositive clauses<sup>12</sup> which modify head nominals within a larger noun phrase. The head nominal itself may be either explicitly stated or elided and implicit. The IC as a whole counts as an NP, whether the nominal head is present or not. This analysis is diagrammed in (16).

---

<sup>12</sup> I previously considered an analysis in which the quantifier-headed clause is treated as a relative clause. Black (1994:367) also notes the parallels between these structures and relative clauses. However, SZ languages, unlike some other Zapotec languages (see Foreman & Munro 2007), do not use resumptive pronouns and instead have gaps in their relative clauses. To analyze the IC's as relative clauses the possessed noun phrases following the quantifier would either count as some sort of resumption, or else as a separate argument, with the head nominal of the IC being the relativized subject of the quantifier and the possessed noun phrase being a separate argument, and one would not want to argue that they are objects of transitive number verbs. However, although IC's are not true relative clauses, in SZ it is possible to achieve a relative effect through apposition, more similar to the IC structure. The following Coatec example was translated by the speaker with a Spanish relative clause, but the Zapotec structure is that of two adjacent clauses, or rather a clause which follows a noun phrase with which one of its own arguments is coreferential.

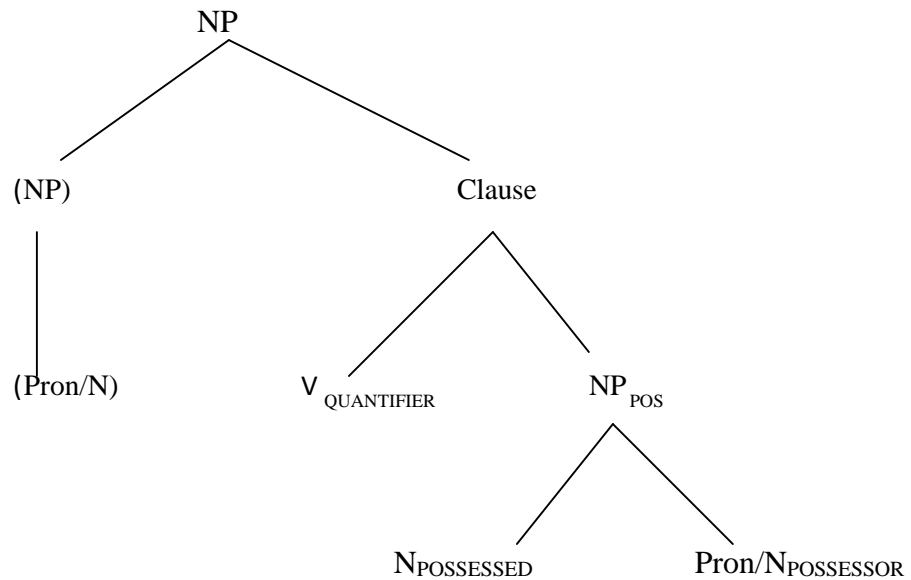
- (i) [Kwa'd ntée ñáa má<sub>i</sub>] [ndé má<sub>i</sub> zóo wée.] [SBarL: Cazador.066]  
 many forms HAB-appear 3A HAB-be 3A place DET

'Varias formas se ve el animal que está allí.'

'In various conditions appear the animals that are there.'

Lit.: 'In various conditions appear the animals. The animals are there.'

(16) Syntactic structure of SZ IC's



Having introduced the proposition that IC's in SZ languages function as NP's, we can now take a closer look at the internal composition of the IC. There are two main semantic components: the superset (§2.1) and the subset (§2.2).

2.1 SZ Supersets.

The nominal head of the IC (if explicitly stated) and the quantifier head of the appositive clause give information about the whole set of referents of the IC, which is termed the “superset”. The present section takes a more careful look at the supersets of SZ IC's on a language by language basis.

SZ IC's were first analyzed by Cheryl Black (1992, 1994, 2000) for Quiégolani Zapotec. The pronoun-initial type of IC found in Quiégolani is most similar to previously described IC's in other language families, in which a non-singular pronoun or person-marking morpheme serves as superset. The difference is simply that in SZ the pronoun and the indication of non-singular number are indicated with two separate words. For example, the pronoun *noo* is used in (17) with the number ‘two’ to refer to a plural first person, and without a numeral to refer to a singular first person.

(17) [SMQ] (Black, 2000)

Sabt w-a xee noo,  
Saturday COMPL-go rise 1EX

w-xa-ndxen [noo [y-rup [x-patron noo]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>.  
COMPL-eat-breakfast 1EX POT-two POS-patron 1EX

‘Saturday I got up and ate breakfast with my patron.’

Lit. “Saturday we two (including) my patron, got up and ate breakfast.”

That in SZ number is indicated independently from any pronoun referring to the superset, is what makes SZ IC's unique compared to IC's in Oceanic, Australian, and other languages. Since the quantifier used in SZ can be a numeral, the number specified in SZ

IC's may also be more specific than in other IC's (with the exception of those with dual pronouns).

In Quiegolani IC's we see two superset words then: the pronoun gives certain grammatical information (person, animacy, social status) about the superset, and the quantifier identifies the number of the superset. The quantifier becomes even more important typologically in Coatec. With the deletion of the head nominal, the quantifier is the only word which gives information about the superset as a whole. It is also now the only explicit head of a constituent which is equivalent to the entire IC:

- (18) Mbi-'d [[*choón* [*xin meé*]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> tí n. [SBalL]  
 COMPL-come POT.three offspring 3HR job  
 'He and his two sons came to work.'  
 Lit. '(They) three (including) his sons came to work.'

- (19) Ngw-áa [[*téh* [*xin meé*]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> ley. [SBalL]  
 COMPL-go all offspring 3HR rosary  
 'She went with all her children to (say) the rosary.'<sup>13</sup>  
 Lit. 'All (of them including) her children went to say the rosary.'

While nominal elements must refer to the superset in Quiegolani, and are always absent in Coatec, they are optional in Miahuatec and Amatec Zapotec, which are geographically intermediate between Quiegolani and Coatec. In Miahuatec Zapotec, including the SAM variety exemplified below, quantifier-initial IC's appear more common, as in (20), but nominal-initial IC's are also possible, as in (21).

- (20) *Nhé* *yu'g* [[*yoón* [*beéhl meé*]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> níit-yeht. [SAM]  
 AUX cook COL.three sister 3HC water-tortilla: food|  
 'She and her two sisters are cooking the food.'  
 Lit. '(They) three (including) her sisters are cooking the food.'

- (21) Ngw-a [*meé* [*reh* [*x-mbal meé*]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> lni. [SAM]  
 COMPL-go 3HC all POS-compadre 3HC party  
 'S/he went with all his/her compadres to the party.'  
 Lit. 'They (including) all his/her compadres went to the party.'

In some cases, whether an inclusory reading is intended can be ambiguous. However, where an inclusory reading is possible it is strongly preferred, and if a non-inclusory reading is intended it is typically made clear from the context. In Miahuatec there is also a morphological way of disambiguating.

In addition to potential mood, completive aspect, and future modal tense, in Miahuatec and some other SZ languages there is an additional form of numerals which is not found on (other) verbs and which is here termed the "collective". This form of the number one means 'whole,' of the number two means 'both,' and with numbers of three and greater this form

<sup>13</sup> In Spanish this was translated with another inclusory construction: 'Fue ella con todo y sus hijos al rosario.'

is most often used with the universal quantifier ‘all’ to refer to ‘all three,’ ‘all four,’ and so forth. The use of this form thus indicates that the number specified is the total number of some group and a partitive meaning is excluded. In Quiégolani the potential mood form is apparently used with these same semantics (Black 1994:329). In Miahuatec IC’s the collective is the form most often used (and likewise the potential form in Quiégolani IC’s), and it is possible to indicate a non-inclusory meaning in Miahuatec by using different inflection on the numeral. In the Miahuatec example in (22) the collective form of ‘three’ is used in an IC but in (23) the potential form is used for a non-inclusory reading.

Inclusory Construction [SBarL]

(22) [[*Reéh* *yoón* [*xi'n* [*m-gol* *Bdoónh*]]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>  
 all COL.three offspring ANC-elder Abdón

ngw-á [roó nit-do']<sub>NP-LOC</sub>.  
 COMPL-go mouth water-holy

‘(All three of them), Mr. Abdón and two of his children went to the shore.’  
 Lit. ‘All three (including) Mr. Abdón’s children went to the shore.’

Non-inclusory meaning [SBarL]

(23) [Tzon [*xi'n* [*m-gol* *Bdoónh*]]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> nd-aá [roó nit-do']<sub>NP-LOC</sub>.  
POT.three offspring ANC-elder Abdón HAB-go mouth water-holy  
 ‘Three of Mr. Abdón’s children went to the shore.’

In sum, the supersets of SZ IC’s are indicated minimally by a quantifier, which may bear special inflectional marking, and sometimes additionally a pronoun referring to the superset. While the quantifier is an appositive modifier of the pronoun, ellipsis of the pronoun may lead to the quantifier being the only explicit indication of the superset entity.

2.2 SZ Subsets.

The subsets of SZ IC’s are always possessed noun phrases. In the existing global literature on IC’s it is common for subsets to be possessed noun phrases, but not obligatory. IC’s are often ambiguous and this requirement that the subsets of SZ IC’s be possessed NP’s creates ambiguity as to whether an inclusory reading is intended versus a quantified possessed NP. Morphological means of disambiguation were discussed in the previous section and further methods for disambiguating are addressed in §2.2.1.

In (e.g. non-SZ) languages with pronominal supersets, the referents not overtly mentioned in the subset can be inferred from the pronominal, i.e. the central member of the superset (Singer, 2001a). For example, [1.DUAL Susan] clearly includes the first person singular, which is the central member of the first person dual superset. A quantifier does not have a central member which can be inferred, so for SZ IC’s whose supersets consist of a single quantifier, there has to be another way of inferring the remaining referent(s) not included in the subset. In SZ IC’s the possessor of the possessed noun phrase is inferred as being the remaining referent of the quantifier besides the subset which is the head of the possessed noun phrase. §2.2.2 considers the special role of noun possession in SZ IC’s.

2.2.1. Ways of disambiguating IC's from quantified possessed noun phrases. It is often the case in languages that the same form could either be an IC or instead a coordinating or comitative construction, as in (24).

- (24) (Colloquial) Czech (Lichtenberk, 2000)  
 Včera sme šli s Jirkou do kina.  
 yesterday AUX.1PL went 'with' Jirkou.INSTR to cinema.GEN  
 a. 'Yesterday Jirka and I went to the cinema.'  
 b. 'Yesterday we and Jirka went to the cinema.'

Just as in some languages a pronoun may be ambiguous as to whether it is functioning inclusorily or not, in SZ there may be ambiguity as to whether a possessor counts towards the quantity indicated by the superset or not. If a non-inclusory meaning is intended this may be indicated by the context, such as by stating clearly that the possessor was not involved, e.g. "Mary's sisters did it while Mary was out."

In §2.1 it was shown how inflection on quantifiers can help to distinguish between inclusory and non-inclusory constructions. It is also possible to use coordinating or comitative markers to disambiguate when a non-inclusory meaning is intended. SZ languages have a native coordinating conjunction (Coatec *na*, Miahuatec *no*) and to form a comitative construction this same word can be used or instead a borrowed Spanish preposition (*kon*, *konh*) 'with.'

- (25) Pedr ngw-a zi'n kon [[rop [xi'n Pedr]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>. [SAL]  
 Peter COMPL-go job with COL.two offspring Peter  
 'Peter went to work with his two sons.'

When a quantifier-initial phrase follows such a comitative/coordination marker, the NP's possessor does not count towards the number indicated by the quantifier if it is coreferential with the noun on the other side of the coordination/comitative marker, as in (25) and (26). If there is no such coreference, then the possessor may indeed count towards the superset, as in (27).

- (26) [Mgol Bdoónh] no [[rop [xi'n xa']<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> nd-aá roó nit. [SBarL]  
 elder Abdón with COL.two offspring 3V HAB-go face water  
 'Mr. Abdón<sub>i</sub> with two of his<sub>i</sub> children went to the beach.'

- (27) Mgol Bdoónh no [[rop [xi'n Mari]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> nd-aá roó nit. [SBarL]  
 elder Abdón with COL-two offspring Mary HAB-go face water  
 'Mr. Abdón went to the beach with Mary and her child.'

Another way of indicating a non-inclusory meaning, in which the possessor is not counted by the quantifier, is by adding a relative clause to emphasize the possession. Compare (28) and (29).

Inclusory meaning preferred

- (28) Nd-aá [[rop [bda'n xa']<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>. [SBarL]

HAB-go COL.two cross-sex.sibling 3V  
 ‘He and his sister went.’

Only non-inclusory meaning possible

(29) Nd-aá [[rop [bda’n xa’]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub> jwa’n ndxáp xa’.[SBarL]  
 HAB-go COL.two cross-sex.sibling 3V REL HAB-have 3V  
 ‘The two sisters that he has went.’

2.2.2. The role of noun possession in SZ IC’s. Noun possession is an obligatory feature of SZ IC’s, and a common feature of IC’s worldwide. In order to more easily discuss its typological significance with regard to this construction type, it is necessary to introduce some new terminology to help us distinguish between subsets which are explicitly stated, and those which must be extrapolated from other elements in the construction. There are three elements represented in IC’s worldwide: the superset, one or more “overt” subsets, and one “covert” subset.

The “overt subset(s)” is a subset that is directly indicated by a noun phrase, which may or may not form a phrase together with the superset (cf. Lichtenberk’s “phrasal” vs. “split” IC’s). It is not embedded within any other subset noun phrase, although it may be embedded within a larger noun phrase that is equivalent to the IC itself.

The “covert subset” is the remaining subset that together with any and all overt subsets completes the set of referents of the superset. The identity of the covert subset must be extrapolated from the morphosyntactic material that is present to indicate the superset or the overt subset. When the superset is represented by a pronoun, as in Australian languages (Singer 2001a), the semantically central person can be inferred as the covert subset. So, if a first person plural pronoun is the superset, the first person singular can be inferred and is thus interpretable as the covert subset. When only a quantifier is present in reference to the superset, as in Coatec, the covert subset can be inferred by speakers familiar with the construction type from the possessor of the overt subset.

Lichtenberk (2000) categorized IC’s based on whether superset and subset together formed a phrase and by whether or not a coordination or comitative marker was present. How a covert subset is indicated is a third typological dimension which should be taken into consideration in future work, which is why it is useful to have a term denoting this subset specifically.

Possessed noun phrases are not uncommon overt subsets in IC’s worldwide since those who are mentioned together in a group are likely to share some relationship that can be expressed linguistically through possession (cf. Singer’s [2001a:69] comments on kinship dyads in IC’s). An unusual feature of SZ IC’s is that the possessor of the overt subset obligatorily counts as the covert subset and is indeed the only indication of the covert subset in quantifier-initial IC’s.

While quantifier-initial SZ IC’s are unusual in not indicating the covert subset as the central member of a pronominal, even languages that use pronouns to refer to supersets may further specify the covert subset through possession the same as SZ languages do. In example (30) the possessor perhaps gives us more information about the covert subset than is indicated by the superset pronominal. Here the superset pronominal could presumably refer to any third person plus ‘her daughter,’ but in fact the girl’s mother is the remaining subset referred to by the superset pronominal.

- (30) Kathlamet (Boas 1901:158.9)  
 A'qa guā'nEsum qasxalō'kcaitx agā'xan.  
 then always they two picked berries her daughter  
 'Then she and her daughter always went picking berries.'

Since possessed subsets occur so frequently in IC's globally, possessors may be inferred as covert subsets when the covert subset is not indicated fully through some other means. If the reference of the superset pronoun includes any speech act participants (i.e. first or second persons), then these are likely to be the covert subset, though they may additionally occur as possessors of overt subsets, perhaps reinforcing their interpretation as covert subsets. If instead the superset is referred to by a third person pronoun (as in [30]), or by a quantifier with no person marking, then the covert subset must be indicated in some other way, and the possessor of the overt subset is likely to be interpreted as the covert subset.

Like lone quantifiers, third person pronominals force the hearer to rely on other elements to draw a conclusion as to the identity of the covert subset. Besides the common instance of covert subsets as possessors of overt subsets, topicality is another likely tool for making such conclusions. Though by nature a third person pronoun does not have a central member, listeners are able to recover the reference of a third person pronoun based on topicality. Example (41) further below seems to be such an example, in which some topical person, unnamed in the current clause, can be inferred as the remaining, i.e. covert, referent of a third person superset pronominal. Future typological work could establish three types of covert subsets: the central member of the superset, the possessor of an overt subset, and a topical covert subset.

### 3. Related quantifier constructions.

Quiegolani IC's are treated extensively in Black (1994:318-393) where they are called "special number marking constructions." However, not all of the examples given there would qualify as IC's. An examination of these constructions should help delimit the boundaries of what can be considered true IC's.

The special number marking constructions in Quiegolani all consist of a nominal head followed by a quantifier and then one or more noun phrases. Many of these constructions are not truly inclusory because the NP's which follow the quantifier have the exact value indicated by the quantifier, whereas the true inclusory construction has following the quantifier a possessed noun phrase whose head is equal to less than the complete value of the quantifier. Likewise, some of the Quiegolani constructions have post-quantifier noun phrases which are equal to the whole set of referents of the nominal which precedes the quantifier.

Including the IC, there are four construction types among Black's special number marking constructions. What I'll call the Enumerated NP Construction (ENC) enumerates a noun phrase by repeating the head nominal in an appositive clause which specifies the number of the head nominal.

#### Enumerated NP Construction

- (31) S-ya [men [y-rup men]<sub>CL</sub>]<sub>NP</sub>. [SMQ]  
 PR-go 3 POT-two 3  
 'They were both going.'  
 Lit. 'They, the two of them, were going.' or 'They were going, they being two.'





The proposed QLC's in (35-40) differ from IC's in that there are two overt subset NP's which do not share a possessive relationship. Together these two overt subsets add up to the total expressed by the quantifier and the whole set of referents of the superset nominal. There are no covert subsets in the QLC or any of these quantifier constructions other than the IC. The QLC and the QAC do not connect the two subjects of the quantifier clause with a coordinating/comitative marker. Like the IC, the QLC occurs in several SZ languages and may occur with or without a superset pronoun preceding the quantifier. Again, the distribution is that the superset pronoun is obligatory in Quiégolani (35-36), optional in Miahuatec and Amatec (37-39), and absent in Coatec (40).

#### Quantified List Construction

- (35) Ts-a [de [y-rup de Susan]<sub>CL</sub>]<sub>NP</sub>. [SMQ] (Black, 2000)  
 POT-go 2 POT-two 2 Susan  
 'You can go with Susan.'  
 Lit. 'You guys, you and Susan being two, can go.'
- (36) R-oo [men [y-rup men Biki]<sub>CL</sub>]<sub>NP</sub> nisgaal. [SMQ] (Black, 2000)  
 HAB-drink 3H POT-two 3H Virginia soda  
 'She and Virginia drink soda pop.'  
 Lit. 'They, she and Virginia being two, drink soda.'
- (37) Ne' m-dil [me' [rop me' soltato]<sub>CL</sub>]<sub>NP</sub>. [SFL] (Angulo, 1922-1930)  
 here COMPL-fought 3 COL-two 3 soldier  
 'Here the two of them fought, they and the soldiers.'  
 Lit. 'They, they and the soldiers being two (groups), fought here.'
- (38) Töö me' [[re' ce' [cic]<sub>NP</sub> [[kwan]<sub>NP</sub> tnoo me']<sub>NP-REL</sub>]<sub>CL</sub>]<sub>NP</sub>. [SFL] (ibid)  
 sell 3 [all many [pineapple]<sub>NP</sub> [[thing]<sub>NP</sub> bring 3]<sub>NP-REL</sub>]<sub>IC?</sub>  
 'They sell all the many pineapples and other things they bring<sup>14</sup>.'
- (39) N-yaad [[rop [nu]<sub>NP</sub> [[xexhap na]<sub>NP-POS</sub>, Chik]<sub>NP</sub>]<sub>CL</sub>]<sub>NP</sub>. [SCA]<sup>15</sup>  
 HAB-come COL-two 1PE<sup>16</sup> father-in-law 1S Frank  
 'We were coming, both me and my father-in-law, Francisco.'
- (40) Ngwáa kwa n [toóp náa mbál]<sub>CL</sub>]<sub>NP</sub> lóo. [SBalL]  
 COMPL-go M-look.for two 1S compadre 2F  
 'The two (of us), the compadre (and) I went to look for you.'

<sup>14</sup> The structure of the relative clause is ambiguous. It may be [pineapples] (and) [things that they bring] or it may be [[pineapples and things] that they bring].

<sup>15</sup> (Riggs 2005)

<sup>16</sup> While this pronoun may appear to denote the superset, it may not. Although the first person inclusive necessarily has plural reference, the first person exclusive may be used for either singular or plural reference in SZ languages. For some speakers and varieties, the 1e pronoun is used with singular reference more often than the 1s pronoun itself.

One significant difference between natural examples of the QLC and IC in SZ languages is that QLC's have only been found with non-numeral quantifiers and with the number two, but not with higher numbers. Elicited examples with 'three' were rejected by a Miahuatec speaker as unnatural.

Table 2 shows the differences between the three quantifier constructions proposed in this section and the IC which is the focus of this paper. Of these, the QLC is the most similar to the IC. In the following section I propose that these two constructions share a common origin.

ENC	NP <sub>1</sub> Quantifier NP <sub>1</sub>
QAC	NP <sub>1</sub> Quantifier NP <sub>1</sub> NP <sub>2</sub>
QLC	(NP <sub>SUPERSET</sub> ) Quantifier NP <sub>SUBSET</sub> NP <sub>SUBSET</sub>
IC	(NP <sub>SUPERSET</sub> ) Quantifier [NP <sub>POSSESSED</sub> NP <sub>POSSESSOR</sub> ] <sub>SUBSET</sub>

Table 2

#### 4. Evolution of SZ IC's.

In this section I describe how SZ IC's possibly developed from QLC's in §4.1 and how they are showing similar grammaticalization paths as Oceanic IC's (Bril 2004) towards comitative constructions in §4.2.

##### 4.1. QLC's become IC's through elision of coreferential elements.

A number of languages with IC's also have constructions similar to the QLC. For example, Chukchi has IC's, as shown in (3) and (41) and something like the QLC, as shown in (42), where both subsets of the superset pronoun are given overtly.

- (41) Chukchi (Dunn 1999: 172)  
 naqam ətrʔec ətri      ŋew-ʔəttʔ-ə-qej.  
 but    only 3PL.ABS woman-dog-E-DIM  
 'And it was just him and the bitch.'  
 Lit. 'but it was only them (including) the bitch.'

- (42) Chukchi (Dunn 1999: 173)  
 ii    ləye-taŋ-qonpə    ye-tumye-w-linet    ətri    jokwajo    ʔi-nə.  
 yes INTS-INTS-always PF-befriend-3PL    3PL.ABS eider.duck.3SGABS wolf-3SGABS  
 'Yes, and the wolf and the duck befriended each other forever.'

The term "quantified list construction" is more appropriate to the SZ construction than to the similar Chukchi construction, since a quantifier quantifies the superset in SZ. However number is also indicated by the superset pronoun used in Chukchi and in other languages with a similar construction type. Since IC's and QLC's are similar constructions and are often found in the same languages, connections between the two constructions are worth consideration.

A Zapotec-specific phenomenon (Butler 1976; Avelino et al. 2004; Munro and Lopez et al. 1999: 20; Beam de Azcona 2004: 335-339) whereby a noun phrase is deleted when it is coreferential with a possessor NP somewhere else in the clause, is a transparent way to derive an IC from a QLC. (43) and (44) show instances where a subject NP is omitted that is coreferential with the possessor of the following object.

(43) g<sup>w</sup>šeé            Ø-čib \_\_\_\_\_ za-jo.            [Yatzachi Zapotec (Butler 1976)]  
 tomorrow        POT-wash        clothes.of-1I  
 ‘Tomorrow we will wash our clothes.’

(44) ...g-weey \_\_\_\_\_ x-pëëd        noo.            [SMQ (Black 1994)]  
 ...POT-take        POS-child        1E  
 ‘...that I can take my daughter.’

Black concludes that IC’s with a single possessed noun phrase for an overt subset exhibit the same phenomenon as seen in transitive sentences with coreferential subjects and possessors of objects. In (45) is a hypothetical example of a QLC in an earlier form of SZ. If such a form did exist, and if the kind of ellipsis seen in (43-44) did apply, the result would be the structure shown in (46), which is the type of IC seen today in Quiegolani and sometimes in Miahuatec and Amatec.

(45) Hypothetical early Southern Zapotec QLC  
 Ngw-á        [meé [tyoóp    meé [xin        meé]NP-POS]CL]NP    yéh.  
 COMPL-go 3HR    POT.two 3HR offspring 3HR            sweathouse  
 ‘S/he went to the sweathouse with her kid.’  
 Lit. ‘They, she and her kid being two, went to the sweathouse.’

(46) Hypothetical early Southern Zapotec IC/elided form of QLC  
 Ngwá        [meé [tyoóp \_\_\_\_ [xin        meé]NP-POS]CL]NP    yéh.  
 COMPL-go 3HR    POT-two    offspring 3HR            sweathouse  
 ‘S/he went to the sweathouse with her kid.’  
 Lit. ‘They, (she and) her kid being two, went to the sweathouse.’

Note that while I have not seen a modern example of a QLC exactly like (45), with only human subsets and one subset being coreferential with the possessor of the other subset, (47) does have this exact same structure and differs only semantically, by having subsets of different levels of animacy. There are restrictions in SZ, as in other languages of the world, regarding the animacy of the referents of IC’s. A description of animacy-related requirements has been withheld here due to space considerations, but there is evidence in Miahuatec that when referring to a group of mixed animacy the higher animate subset must precede the subset of lower animacy. This is likely the reason that the first overt subset in (47) has not been deleted, despite its coreference with the possessor of the following overt subset.

(47) [[Rop    xa’ [m-bak    xa’]NP-POS]CL]NP    nd-aá.            [SBarL]  
 COL.two 3V    ANC-dog    3V            HAB-go  
 ‘He went with his dog.’

Lit. ‘(They,) he and his dog being two, went.’

Quiegolani examples cited by Black show that there is some optionality with regards to the deletion of a subset NP that is coreferential with the possessor of a following subset NP. The two examples in (48) are virtually identical except that one is a QLC and the other an IC. (Recall the discussion from §2 which described how proper names may be used with associative plural reference in Quiegolani, referring to a group including the named person).

(48) [SMQ] (Black 1994: 342)

- a. Biki z-a g-un kompanyar  
Virginia PR-go POT-do accompany

[Gecha [y-rup Gecha [x-pëəd Gecha]]<sub>NP-POS</sub>]<sub>CL</sub>].

Lucrecia POT-two Lucrecia POS-baby Lucrecia

‘Virginia went to accompany Lucrecia and her baby.’

Lit. ‘Virginia went to accompany Lucrecia & Co., Lucrecia (and) her baby being two.’

- b. Biki z-a g-un kompanyar [Gecha [y-rup [x-pëəd Gecha]]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>.  
Virginia PR-go POT-do accompany Lucrecia POT-two POS-baby Lucrecia.

‘Virginia went to accompany Lucrecia and her baby.’

Lit. ‘Virginia went to accompany Lucrecia & Co., Lucrecia’s (and) baby being two.’

#### 4.2. Grammaticalization of SZ quantifier constructions.

Bril (2004) identified how superset pronouns have come to be used as markers of coordination and comitative constructions in Oceanic languages (see §1.2). As seen throughout this paper, superset pronouns are optional in SZ IC’s and the more essential word representing the superset is the quantifier. One quantifier, ‘two’, is grammaticalizing in the same way Bril observed for superset pronouns. In the following examples from Coatec, Amatec and Quiegolani, the form of the number used in QLC’s and IC’s, i.e. the collective or the potential (depending on the language), is translated as ‘with’ or ‘and’. These examples differ from typical IC’s and QLC’s in having different constituent order. Subset NP’s follow the quantifier in all the previous examples cited in this article, but here one subset of the pair quantified by ‘two’ precedes the quantifier and the other follows. Note also that in all the quantifier constructions found in Quiegolani and listed in Table 2, the quantifier-headed clause is in apposition to a noun phrase which either denotes a superset or is coreferential with one of the NP’s that follows the quantifier. In the examples with grammaticalized ‘two’ there is no such apposition. The NP that precedes ‘two’ is coordinate with the NP that follows ‘two’ and is neither wholly nor partially coreferential with it.

- (49) Nal za na tub kwent cheen konej rop mbew. [SCA; (Riggs n.d.)]  
now give 1S one story of rabbit COL.two coyote  
‘Now I will give the story of the rabbit with the coyote.’

- (50) [[Laa [xnaa noo]<sub>NP-POS</sub>]<sub>NP</sub> y-rup [xuz noo]<sub>NP-POS</sub>]<sub>NP</sub> r-laa-w. [SMQ; Black 1994]  
 FOC mother 1E POT-two father 1E HAB-do-3I  
 ‘My mother and my father did it.’
- (51) Xna-ydoo x-pee Manwel [SMQ]  
 mother-church POS-son Manuel
- n-ak [[Katalina]<sub>NP</sub> y-rup [Tomas]<sub>NP</sub>]<sub>NP</sub>.  
 s-become Catherine POT-two Thomas  
 ‘The godparents of Manuel’s son are Catherine and Thomas.’
- (52) Ná nhbítê-ta’ mẽ gôtz tǒp xmbál mẽ. [SBalL]  
 NEG IRR:return=anymore 3HR female two POS:compadre 3HR  
 ‘Now the woman and her compadre did not come back.’

Comparing these coordinate noun phrases to the quantifier-initial type of IC’s and QLC’s which are common in all the SZ varieties studied here except Quiegolani, the most striking difference is that one subset noun phrase occurs preceding the quantifier when normally it would follow it, along with the other subset NP if there were one. It is common to front subjects to preverbal position to mark focus and topic in Zapotec languages. Such fronting is enough to mark focus but such fronted constituents may optionally occur with a focus particle. In (53) a focus particle precedes a pronoun which occurs earlier in the sentence than the quantifier. Since these pronouns do not indicate number there is potential ambiguity as to whether they refer to the superset or some subset of the IC. In (53) the whole IC, which is the subject of the larger clause, is focused at the beginning of the sentence. However, in (54) only one subset of a would-be quantifier construction is fronted to focus position, while the quantifier and the remaining subset are left following the main verb. It is difficult to say whether this should still be regarded as an IC or QLC with dislocation of one subset NP, or whether the fronting of this NP is enough to give this example a different status as a grammaticalized use of the number ‘two’ to mean ‘with.’ The example is shown twice with the bracketing reflecting the two different analyses, one in which the IC (in italics) is a discontinuous noun phrase, and the other in which the same elements are treated as subject and comitative oblique. The focusing of one subset NP may create a pivotal ambiguity which enables the reinterpretation of the superset numeral to take place.

- (53) [Le’eh [meé [roóp [beéh’l meé]<sub>NP-POS</sub>]<sub>CL</sub>]<sub>NP</sub>]<sub>NP</sub> nhé yu’g níityehët. [SAM]  
 [FOC 3HC two sister 3HC] AUX cook |water-tortilla: food|  
 ‘She and her sister are preparing the meal.’  
 Lit. ‘They, she and her sister being two, are cooking food.’
- (54) [*M-beéh’t*]<sub>NP</sub> ngw-a g-aáw [[roóp x-na’a]<sub>CL</sub>]<sub>NP</sub>. [SAM]  
 [*M-beéh’t*]<sub>NP</sub> ngw-a g-aáw [roóp x-na’a]<sub>PP</sub>.  
 ANC-child COMPL-go POT-eat COL-two POS-mother  
 ‘The child went to eat with his/her mother.’

Subset NP's which have been fronted with respect to the quantifier are overt. If they possess or are coreferential with the possessor of a subset NP, in fronted position they are being singled out for focus and are not solely in the possessor function. In (49) if there were no fronting with respect to the quantifier this would be a QLC with two overt subsets, but in (54) the fronted subset NP is indeed the possessor of the subset NP that remains in unfocused position following the quantifier. It is significant that there is no resumptive pronoun following the subset 'mother' because there is a single instance of an NP referring to 'child' and it is not embedded in an IC as the possessor of the overt subset NP. Now the possessor NP has become overt itself and there is no subsequent coreferential mention of it. One might argue that this is no longer an example of an IC because the second subset is no longer covert. An alternative way to view this is that the second subset is even more covert, not even being stated inside the IC, and its only mention being now in a separate phrase is reminiscent of Lichtenberk's split (as opposed to phrasal) IC category.

## 5. Conclusion.

This examination of SZ IC's has attempted to contribute to the global typology of IC's 1) by making it known that supersets may be represented by words or phrases other than pronouns, and indeed that superset terms, even ones of very different lexical categories, may show similar grammaticalization paths to those already described for superset pronouns; and 2) by introducing the terms "overt" and "covert" subset to distinguish between these two semantic elements of IC's.

New descriptions of IC's should identify what type of words or markers are used to indicate the superset. If a quantifier represents the superset, its part of speech and role in IC phrase structure should be identified, since these can be verbs, adjectives, or special categories to themselves. It should also be stated whether or not multiple markers are used, and which markers are optional or obligatory. In terms of semantics, future descriptions of IC's should tell us how the covert subset may be deduced, e.g. as the central member of a pronoun, an element with another syntactic function embedded within the IC or larger clause, through topicality, etc.

Finally, syntactic descriptions of IC's should describe the functional roles that IC's play within the larger clause, and whether IC's themselves constitute a single phrase or not. Lichtenberk (2000) showed that the superset and subset can be part of a single phrase or can belong to different phrases, and Southern Zapotec data show that a significant phrasal relationship may also exist between the overt and covert subsets. Also seen in the Zapotec data is a related construction in which only overt subsets appear, with these together representing the whole group referred to by the superset. Future descriptions of languages with IC's should also point out whether a construction like the QLC exists and how it is realized.

## References

- Aissen, Judith L. 1989. "Agreement controllers and Tzotzil comitatives." *Language*, 65: 518-536.
- Angulo, Jaime de. 1922-1930. "Zapotecan texts." Archived with the American Philosophical Society in the American Council of Learned Societies Committee on Native American Languages Collection: Freeman and Smith 3943.

- Avelino, Heriberto, John Foreman, Pamela Munro, and Aaron H. Sonnenschein. 2004. "Covert subjects in Zapotecan." Paper presented at the SSILA annual meeting, Boston.
- Beam de Azcona, Rosemary G. 2004. *A Coatlán-Loxicha Zapotec Grammar*. Ph.D. dissertation. UC Berkeley.
- . 2014a. "Algunos rasgos lingüísticos del zapoteco sureño" in Rebeca Barriga Villanueva and Esther Herrera Zendejas, eds. *Lenguas, Estructuras y Hablantes: Estudios en Homenaje a Thomas C. Smith Stark, Vol. 2*. México, D.F.: Colegio de México. 643-660.
- . 2014b. "Algunas Isoglosas de la Sierra Sur", Trabajo presentado en el Coloquio de Lenguas Otomangues y Vecinas VI in honor of Mario Molina Cruz. Oaxaca: Biblioteca de Investigación "Juan de Córdova". Powerpoint available at [https://www.academia.edu/6872282/Algunas\\_isoglosas\\_de\\_la\\_Sierra\\_Sur](https://www.academia.edu/6872282/Algunas_isoglosas_de_la_Sierra_Sur).
- . In preparation. *El zapoteco de la Sierra Sur: Estudios comparativos dentro de un grupo areo-genético*.
- Black, Cheryl A. 1992. "Comitative coordination using quantifiers in Quiégolani Zapotec." In H. Andrew Black and James McCloskey (eds.), *Syntax at Santa Cruz*, vol. 1, 1-33. Santa Cruz: University of California.
- . 1994. *Quiégolani Zapotec Syntax*. Ph.D. dissertation. UC Santa Cruz.
- . 2000. "Number Marking Innovations in Zapotec." Paper presented at the Indigenous Voice of Oaxaca conference, UCLA
- Boas, Franz. 1901. "Kathlamet Texts." *Bureau of American Ethnology*, Bulletin 26. Washington: Government Printing Office.
- Bradley, David. 2001. "Counting the Family: Family Group Classifiers in Yi (Tibeto-Burman) Languages" *Anthropological Linguistics* Vol. 43 No. 1, pp. 1-17.
- Bril, Isabelle. 2004. "17. Coordination strategies and inclusory constructions in New Caledonian and other Oceanic languages". In Martin Haspelmath (ed.), *Coordinating Constructions*, 499–533.
- Butler, Inez M. 1976. Reflexive constructions of Yatzachi Zapotec. *IJAL* vol. 42, no. 4.
- Corbett, Greville G. 2000. *Number*. Cambridge University Press.
- Córdova, P. Fray Juan de. 1987 [1578]. *Arte del Idioma Zapoteco*. México: INAH.
- Dienst, Stefan. 2006. *A Reference Grammar of Kulina*. Ph.D. dissertation. La Trobe University.
- Dunn, Michael John. 1999. *A Grammar of Chukchi*. Ph.D. dissertation. Australian National University.
- Evans, Nicholas D. 1995. *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Foreman, John and Pamela Munro. 2007. "Reconstructing a Typologically Marked Strategy for Zapotec Relative Clauses". In the proceedings of the 2004 meeting of the High Desert Linguistics Society, Albuquerque.
- François, Alexandre. 2000. *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu*. PhD Dissertation, Université Paris IV Sorbonne.
- Gaby, Alice. 2006. *A grammar of Kuuk Thaayorre*. Ph.D. dissertation. University of Melbourne.
- Haspelmath, Martin. 2007. "Coordination." In Timothy Shopen (ed.) *Language typology and syntactic description, Vol. II: Complex Constructions*. 2nd ed, 1-51. Cambridge: Cambridge University Press.



- Ladusaw, William A. 1989. "Group reference and the plural pronoun construction."  
In *Papers on the plural pronoun construction and comitative coordination*.  
Syntax Research Center, University of California, Santa Cruz.
- Lichtenberk, Frantisek. 2000. "Inclusory Pronominals." *Oceanic Linguistics*, Vol. 39:1.
- Marlett, Stephen A. and Velma B. Pickett. 2001. "Pluralization in Zapotec Languages."  
<http://www.sil.org/mexico/zapoteca/G025a-ZapotecPlurals.htm>
- Munro, Pamela. 2000. "San Lucas Quiavini Zapotec *nèe* and *cèhnn* 'with': Preposition,  
Adverb (?), Applicative." Paper presented at the meeting of the American  
Anthropological  
Association, San Francisco.
- Munro, Pamela and Felipe H. Lopez with Olivia V. Méndez, Rodrigo Garcia, and  
Michael R. Galant. 1999. *San Lucas Quiavini Zapotec Dictionary*. UCLA  
Chicano Studies Research Center Publications.
- Riggs, David. 2005. *Amatlán Zapotec texts*. ms.
- Riggs, David and Sylvia. N.d. *Amatlán Zapotec lexicon*. ms.
- Schwartz, Linda 1988a. "Conditions for verb-coded coordinations." In Michael Hammond,  
Edith A. Moravcsik and Jessica R. Wirth (eds.) *Studies in Syntactic Typology*,  
Typological Studies in Language 17, 53-73. Amsterdam: John Benjamins.
- 1988b. "Asymmetric feature distribution in pronominal 'coordinations'." In M.  
Barlow and A. Ferguson (eds.) *Agreement in Natural Language: Approaches,  
Theories, Descriptions*, Center for the Study of Language and Information, 237-249.  
Stanford: Stanford University Press.
- Singer, Ruth. 2001a. *The Inclusory Construction in Australian languages*. University of  
Melbourne Honours Thesis.
- 2001b. "The Inclusory construction in Australian languages", *Melbourne Papers  
in Linguistics and Applied Linguistics*: 81-96.
- Smith Stark, Thomas C. 2007. "Algunas isoglosas zapotecas." in Cristina Buenrostro  
et al. *Clasificación de las lenguas indígenas de México*, Memorias del III  
Coloquio Internacional de Lingüística Morris Swadesh, México D.F., 69-133.

