

# Aspectual Morphemes in CVZ Number Marking

James Faville  
December 22, 2017

## Abstract

We show that the definite future and irrealis morphemes in Colonial Valley Zapotec (CVZ) can given denotations as a nonmodal future tense and a generic modal quantifier, respectively, and that these denotations can account for their interpretation as prefixes to numerals. We then discuss remaining issues with this analysis and pose questions for future research.

## 1 Introduction

### 1.1 Colonial Valley Zapotec

Colonial Valley Zapotec (CVZ) is the language of a corpus of texts written in the Oaxaca Valley of Mexico between 1565 and 1808.<sup>1</sup> CVZ represents the historic form of modern Valley Zapotec languages such as San Lucas Quiavini Zapotec (ISO code [zab]), which are VSO members of the Otomanguean family. CVZ was written using the Roman alphabet but orthography was not standardized and the relation between CVZ orthography and phonology is not always clear (c.f. Broadwell 2014). The CVZ corpus consists primarily of documents like testaments and wills written by native speakers. In this paper we also refer to data from the adaption of a Spanish catechism into Zapotec which was commissioned by Fray Leonardo Levanto, cited as Levanto (1766).

### 1.2 Aspectual Morphemes on Numbers and Quantifiers

Numbers in Colonial Valley Zapotec are often accompanied by temporal/aspectual prefixes. The irrealis morpheme *qui-* glossed as IRR appears in (1) below as an aspectual prefix in Cordova's (1578) Zapotec grammar and in (2) as a prefix to a numeral.<sup>2</sup>

- (1) C-olla=ya  
IRR-sing=1  
'I will sing.' (Broadwell 2015, ex 3c; Cordova 1578a:19)

<sup>1</sup>Thanks are owed to my professor Brook Danielle Lillehaugen, George Aaron Broadwell, Felipe H. Lopez, Janet Chávez, my classmates, especially Emily Drummond and Bridget Murray who provided feedback on a draft of this paper, Shizhe Huang and Peter Klecha who also provided feedback on a draft of this paper, and many others. All mistakes are my own.

<sup>2</sup>This paper uses the following glossing abbreviations: 1 = 1st person; 2 = 2nd person; 3 = 3rd person; DEF = definite future; HAB = habitual; IRR = irrealis. Hyphens are not original to the text, but have been inserted at morphological boundaries to aid comprehension of the examples.

- (2) anaa r-onatijaga tes p<sup>o</sup> ernades testico Fracisgo luis p<sup>o</sup>  
now HAB-listen witness Pedro Hernandez witness Fracisgo Luis Pedro  
mural testico **que-ona bene xiini**  
Morales witness IRR-**three person child**  
'Now listen the witnesses Pedro Hernandez, witness, Fracisgo Luis Pedro  
Morales, Morales, witness, and **all three children**'  
(Plesniak 2013, ex 12; A1697;62v-22-24)

A similar alternation is seen in the definite future morpheme *ze-* and the perfective morpheme *hue-*. Note that all of these morphemes experience orthographic irregularities.

The uses of the future tenses are described further below. The goal of this paper will be to give a formal semantic denotation to each future tense marker which consistently accounts for its interpretation in both verbal and numeric environments. Although these morphemes also appear on quantifiers, we leave a thorough investigation of this phenomenon to future research. It seems very likely, however, that the analysis developed here can be extended to account for environments of non-numeric quantification without much difficulty.

## 2 Methodology

When a morpheme is observed in two very different environments, a natural question to ask is whether there is only one morpheme or whether there are rather two distinct (though possibly diachronically related) morphemes with homophonous realisations. How would one go about answering this question?

If a single denotation can reasonably cover the interpretation of the morpheme in both environments, the most parsimonious account would be one of identity rather than homophony. If this cannot be done without excessive stipulation or detriment to the semantic theory, then an analysis of homophony is preferable.

We propose the following methodology for carrying out this investigation:

- Describe adequately the interpretation of the morpheme(s) in each environment.
- Choose one environment, and form a denotation which accounts the interpretation of the morpheme here.
- Assume that there is a single morpheme which has an identical denotation in each environment. Attempt to minimally change the denotation formed above to account for its interpretation in the second environment while preserving an account of its interpretation in the first environment.
- Reflect on whether these changes are reasonable.

This paper uses the above methodology to examine the question of whether the aspectual morphemes *ze-* ‘DEF’ and *qui-* ‘IRR’ in CVZ also appear as prefixes to numerals (and similar lexical items such as quantifiers, c.f. Black 2008), or whether each has a homophonous counterpart with a distinct meaning.

Section 3.1, the distinction between *qu-* and *z-* as verbal prefixes is discussed, and a specific denotation for each morpheme is proposed. In Section 3.2, the interpretation of *qu-* and *z-* as numeric prefixes is discussed. In Section 4, the semantic formalism used in the rest of the paper is outlined. In Section 5, we write denotations for the verbal prefixes, and in Section 6 we extend these to account for numeric uses.

### 3 Description

The table in (3) below summarizes the uses of the aspectual morphemes on verbs and numerals. The interpretation of *ze-* and *qui-* is discussed below, but a full description of the perfective *hue-* is outside of this paper, and it is only included in this table for completeness. An analysis of *hue-* is given in Lillehaugen (2011).

<i>Uses of CVZ Aspectual Morphemes</i>			
Morpheme	Gloss	Verbal Use	Numeric Use
(3) <i>hue-</i>	PFV	perfective aspect	marks ordinality
<i>ze-</i>	DEF	definite future	marks additionality
<i>qui-</i>	IRR	irrealis future	marks entirety

#### 3.1 Verbal

The irrealis morpheme *qui-* can be used in various nonverbal environments. (4) and (5) below demonstrate that in addition to denoting a future time as in as (1), *qui-* can denote a deontic modality which we indicate by ‘should’ in the translation

- (4) ...qui-raa-lii benni Christianos ge-zeete=ni c-oo-quipe=ni.  
 IRR-all-straight person Christians IRR-study=3: IRR-be.contained-head=3

(Levanto 1766:1)  
 ‘...all true Christians should know it [the laws of our Holy Mother Church] and it should be contained in their heads.’

- (5) qui-yapa napea que-nni=na laaya...  
 IRR-be.kept command: IRR-do=1PL sacred.thing (Levanto 1766:2)  
 ‘This command should be kept: we should do the sacred thing...’

The definite future in modern Zapotec languages is used to discuss future events

of whose occurrence the speaker is certain. Munro (2007) illustrates this by the following contrast between the irrealis and the definite futures in Tlacolula Valley Zapotec (TVZ).<sup>3</sup>

- (6) a. Z-eheh Jwaany Ba'ahc.  
DEF-go Juan Tlacolula  
'Juan will surely go to Tlacolula.'  
b. S-tòò'oh Gye'eihlly ca'rr.  
DEF-buy Mike car  
'Mike will surely buy the car.' (Munro (2007), ex 2-3)
- (7) a. Ch-ia Jwaany Ba'ahc.  
IRR-go Juan Tlacolula  
'Juan will go to Tlacolula.'  
b. Y-tòò'oh Gye'eihlly ca'rr.  
DEF-buy Mike car  
'Mike will buy the car.' (Munro (2007), ex 2-3)

We will later show that this meaning can be derived from giving *ze-* the semantics of a nonmodal future tense. A larger corpus study is needed to confirm that this meaning encompasses all CVZ usage of the definite future, but it appears that this is the case. The CVZ sentence (8) below is compatible with an analysis of *ze-* as a straightforward future tense.

- (8) áca=ti zo-aca c-ochaga-ñaa=ni  
NEG=TI DEF-can scirr-join-hand=3  
'they cannot get married' (Anderson and Lillehaugen 2015, ex 44;  
Vellon 121; 25)

The translation of (8) which a nonmodal future meaning for *ze* would predict is 'they will surely not be able to get married', which seems very plausible.

- (9) C-olla=ya  
IRR-sing=1  
'I will sing.' (Broadwell 2015, ex 3c; Cordova 1578a:19)

### 3.2 Numeric

On numbers, the definite future denotes additionality and is translated as 'another' or 'more', as in (10) and (11) below.

---

<sup>3</sup>Munro (2007) also discusses a similar morpheme *z2* used on verbs of motion which has sometimes been identified with the definite future, but whose semantics are more complex and which has a distinct form in CVZ.

- (10) ze-tuba cue  
DEF-two lands  
(Munro & Sonnenschein 2007, ex 40c, Ti649)  
'another two lands'
- (11) ce-tobi quiba-yaga qu-e=ya lao lachella=ya  
DEF-one metal-wood IRR-go=1s face spouse=1s  
'I will go before my spouse with another axe' (Plesniak 2013, ex3;  
Te614:1-33)

Munro and Sonnenschein (2007) claim that the irrealis future denotes definiteness on number and can be translated as 'the'. Plesniak (2013) argues against this analysis on the basis of examples like (2), repeated below as (12) for convenience, in which the three children referred to be *que-ona bene xiini* have not been previously introduced into the discourse.

- (12) anaa r-onatijaga tes p<sup>o</sup> ernades testico Fracisgo luis p<sup>o</sup>  
now HAB-listen witness Pedro Hernandez witness Fracisgo Luis Pedro  
mural testico **que-ona bene xiini**  
Morales witness IRR-**three person child**  
'Now listen the witnesses Pedro Hernandez, witness, Fracisgo Luis Pedro  
Morales, Morales, witness, and **all three children**'  
(Plesniak 2013, ex 12; A1697;62v-22-24)

Instead, Plesniak argues that the irrealis future conveys entirety, and can be roughly translated as 'all'. Black (2008) ascribes the same interpretation to the potential aspect on numbers in modern Isthmus Zapotec and Quiegolani Zapotec, making this a very attractive analysis.

However, the (13) below is problematic for an entirety analysis, since *qui-* appears to be interpreted as an ordinal.

- (13) ti-nij[=a] na-pa=ya cache cuee layoo bini quihe-tobi cuee naa  
hab-say=1s STA-have=1s seven plot land seed IRR-one plot STA-be  
saquieçaa co-ropa cuee naa sanaquiegaçea co-yona...  
Saquieçaa PFV-two plot STA-be Sanaquiegaçea PFV-three plot  
'I have seven plots of arable land: the [first?] plot is Saquieçaa, the  
second plot is Sanaquiegaçea, the third...' (Te614;34-34)

Lillehaugen (2011) notes that this construction would be expected to use the perfective, and other examples of the irrealis on numerals which we have seem do appear compatible with a reading of 'whole' or 'only', both of which are possible meanings for 'all one'. We will leave an investigation of (13) and the objections it poses to our analysis if the use of the irrealis as an ordinal proves to be robust for future research.

## 4 Formalism

We assume that modal quantification is quantification over a domain of possible worlds whose composition is determined compositionally by the lexical semantics of morphemes or contextually. Kratzer (1981) argues that a BEST function is an integral component of modal quantification, since embedded propositions under modal quantification only have to hold of the most salient worlds under consideration. A schema for modal quantification is given in (14) below, where M is a modal domain of possible worlds and P is some proposition that holds of possible worlds.

$$(14) \quad \textit{Modal Quantification} \\ \forall w' \in \text{BEST}(M) [ P(w') ]$$

We follow Klecha (2013) in analyzing futures like the English *will* to be modals in that they quantify over possible worlds which could be the actual world (as the future is usually not known). We also suppose the possibility of a nonmodal future, which simply denotes a relation between two time intervals.

Sample denotations for these two types of futures are given below: syntactically, the modal future introduces quantification over the VP, and the nonmodal future would probably be expressed as a feature on T. *u* corresponds to evaluation time: we leave for future syntactic/semantic analysis the question of whether this order of composition is feasible.

$$(15) \quad \textit{Modal Future} \\ \lambda P_{st} \lambda u_i . \forall w' \in \text{BEST}(\text{fut}(w, t)) [ P(w') ]$$

$$(16) \quad \textit{Nonmodal Future} \\ \lambda u_i \lambda t_i . t > u$$

We follow Landman (2008) in supposing that numerals denote functions which return the order of sets and compose with other nominal modifiers by Predicate Modification, though this analysis could work equally as well if numerals denote sets and compose via intersection. The denotation of the numeral “two” is given below.

$$(17) \quad \llbracket 2 \rrbracket^{i,c,g} = \lambda x . |x| = 2$$

In order for our denotation to be consistent between the verbal and numeric uses of *qui-* and *zi-*, we will suppose that the type clash between entities and predicates of intervals is not an obstacle to semantic composition. We also assume two types of type-shifting within NumP are possible. The first is shifting from *et* to  $\langle et, t \rangle$  by changing a function on entities to a function on sets by passing the most salient member of that set to the original function. This could perhaps be accomplished by a null morpheme analogous to “of” in English

complex cardinals like "three of them" (c.f. Landman 2008). The second is shifting from  $\langle et, t \rangle$  to  $et$  by changing a function on sets to a function on (plural) individuals by passing the set of all entities who are part of that individual to the original function.

## 5 Verbal Analysis

The distinction between the definite and irrealis futures can be captured by treating the irrealis as a modal like English *will*, but the definite future as a nonmodal temporal operator that asserts temporal subsequence. These denotations are given in (18) below. Note that the modal domain over which IRR quantifies is assumed to be contextually determined since it could be either epistemic or deontic, as discussed in Section 3.

- (18) *Denotation of Future Tenses*
- a.  $\llbracket \text{DEF} \rrbracket^{i,c,g} = \lambda u_i \lambda t_i . [ t > \iota u ]$
  - b.  $\llbracket \text{IRR} \rrbracket^{i,c,g} = \lambda M_{st} \lambda P_{st} . [ \forall w' \in \text{BEST}(M) [ P(w') ] ]$

These denotations capture the crucial component of the meaning of the definite future, which is that in veridical contexts speakers must have absolute knowledge of the event. If they did not possess this knowledge, they would be forced to use the modal future since there are multiple possible worlds whose futures could correspond to the future of the actual world.

## 6 Numeric Analysis

We now adapt the denotations above to account for numeric uses by replacing functions of intervals and worlds with functions of entities, and the BEST function which picks out a most salient set of worlds with the analagous  $\iota$  operator on entities which picks out a most salient entity.

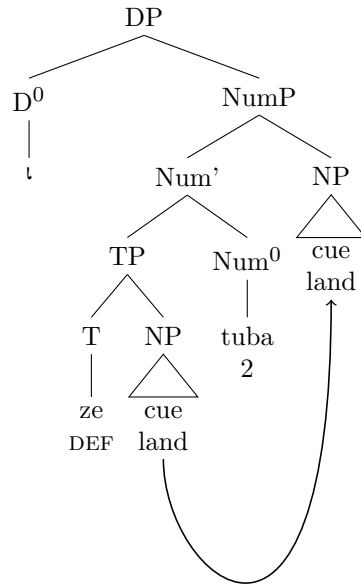
- (19) *Denotation of Future Tenses on Numbers*
- a.  $\llbracket \text{DEF} \rrbracket^{i,c,g} = \lambda y_e \lambda x_e . [ x > \iota y ]$
  - b.  $\llbracket \text{IRR} \rrbracket^{i,c,g} = \lambda M_{et} \lambda P_{et} . [ \forall y \in \iota M [ P(y) ] ]$

### 6.1 Definite Future on Numbers

Using the structure in (20) below, we can derive the meaning of the nominal phrase containing *ze-* (10). In (20), the NP *cue* moves from within the TP projected by *ze* to a right-branching Spec position in NumP, where it is pronounced, and it is interpreted in both positions. This derives the correct word

order and meaning, but we leave further motivation of this movement to future research.

(20) *Tree Diagram of (10)*



The meaning of the NumP is derived in (21) below. Note that the iota operator  $\iota$  is the function which picks out the most salient individual in a set (refer to the previous discussion of type-shifting).

- (21)
- a.  $[[\text{DEF}]^{i,c,g}] = \lambda y_e \lambda x_e . [x > y]$
  - b.  $[[\text{DEF}]^{i,c,g}] = \lambda P_{et} \lambda x_e . [x > \iota P]$  (after type-shifting)
  - c.  $[[\text{TP}]^{i,c,g}] = \lambda x_e . [x > \iota \text{land}]$
  - d.  $[[\text{Num}']^{i,c,g}] = \lambda x_e . [(x > \iota \text{land}) \wedge |x| = 2]$
  - e.  $[[\text{NumP}]^{i,c,g}] = \lambda x_e . [x \text{ is land} \wedge (x > \iota \text{land}) \wedge |x| = 2]$

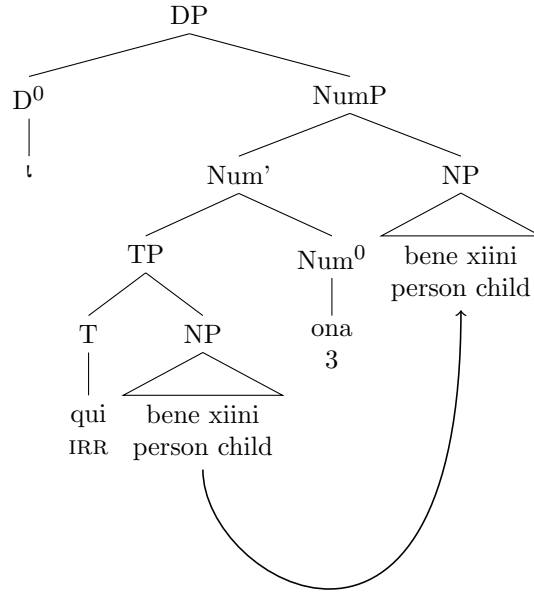
The meaning obtained from this composition is roughly "the two lands which are all subsequent with respect to some contextually relevant ordering to the members of the group of lands most salient prior to this utterance".



## 6.2 Irrealis Future on Numbers

We now give a similar derivation of the meaning of the nominal phrase containing *qui-* in (2).

(22) *Tree Diagram of (2)*



- (23)
- $[[\text{IRR}]]^{i,c,g} = \lambda M_{et} \lambda P_{et} . [ \forall y \in \iota M [ P(y) ] ]$
  - $[[\text{TP}]]^{i,c,g} = \lambda P_{et} . [ \forall y \in \iota \text{child} [ P(y) ] ]$
  - $[[\text{TP}]]^{i,c,g} = \lambda x_e . [ \forall y \in \iota \text{child} [ y \leq x ] ]$  (by type-shifting)
  - $[[\text{Num}']]^{i,c,g} = \lambda x_e . [ \forall y \in \iota \text{child} [ y \leq x ] \wedge |x| = 3 ]$
  - $[[\text{NumP}]]^{i,c,g} = \lambda x_e . [ x \text{ is child} \wedge \forall y \in \iota \text{child} [ y \leq x ] \wedge |x| = 3 ]$

The meaning obtained from this composition is roughly "the group of three children which contains all members of the group of children most salient prior to this utterance".

## References

- [1] Altshuler, Daniel, and Carter, Sam. 'Now' as a dyadic connective. In *Proceedings of Semantics and Linguistic Theory 27* (in prep.)
- [2] Anderson, Carolyn Jane, and Brook Danielle Lillehaugen. 2015. The morphosyntax of negation in Colonial Valley Zapotec. *Annual meeting of the Society for the Study of the Indigenous Languages of the Americas*, Portland.
- [3] Black, Cheryl A. 2000. Number marking innovations in Zapotec. In *The Indigenous Voice of Oaxaca* conference. University of California, Los Angeles.
- [4] Broadwell, George Aaron. 2014. Phonological and orthographic distinctions in early Zapotecan manuscripts: The influence of Nahuatl. Ms. University of Florida. [https://www.academia.edu/1561180/Phonological\\_and\\_orthographic\\_distinctions\\_in\\_early\\_Zapotecan\\_manuscripts\\_The\\_influence\\_of\\_Nahuatl](https://www.academia.edu/1561180/Phonological_and_orthographic_distinctions_in_early_Zapotecan_manuscripts_The_influence_of_Nahuatl)
- [5] Broadwell, George Aaron. 2015. The historical development of the progressive aspect in Central Zapotec. *International Journal of American Linguistics* 81.2: 151-185.
- [6] Cordova, Juan de. 1578a. *Arte en lengua zapoteca*. Mexico: Casa de Pedro Balli. Republished by Nicolás León as *Arte del idioma zapoteco de Juan de Córdoba* (Morelia, Michoacán: Imprenta del Gobierno en la Escuela de Artes, 1886). A facsimile edition of León's edition published as *Arte del idioma zapoteco* (Mexico City: Ediciones Toledo and INAH, 1987)
- [7] Klecha, Peter. 2014. Diagnosing modality in predictive expressions. *Journal of Semantics* 31.3: 443-455.
- [8] Klecha, Peter. 2015. Modality and embedded temporal operators. Ms. Ohio State University.
- [9] Kratzer, Angelika. 1981. The notional category of modality. *Pages 38-74 of: Eikmeyer, Hans-Jurgen, & Rieser, Hannes (eds), Words, Worlds, and Context*, vol. 6. Walter de Gruyter.
- [10] Landman, Fred. 2008. *Indefinites and the Type of Sets* 7. John Wiley & Sons.
- [11] Lillehaugen, Brook Danielle. 2011. The perfective aspect in Colonial Valley Zapotec. *Linguistic Association of the Southwest annual meeting*, South Padre Island.
- [12] 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 Zaaпотeca*, beta version. Ticha Project: <https://ticha.haverford.edu/en/texts/levanto/>
- [13] Lillehaugen, Brook Danielle. 2011. The perfective aspect in Colonial Valley Zapotec. *Linguistic Association of the Southwest annual meeting*, South Padre Island.
- [14] Munro, Pamela. 2007. A definite mystery. In *Berkely Linguistics Society 1001*:91.
- [15] Munro, Pamela, and Sonnenschein, Aaron Huey. 2007. Four Zapotec Number Systems. University of California at Los Angeles, m.s.
- [16] Plesniak, Daniel. 2013. Towards a Fuller Understanding of the Interaction between Aspect and Numerals. Haverford College, m.s.