

Chapter 2

Definiteness in Cuevas Mixtec

Carlos Cisneros

University of Chicago

Languages vary widely in their morpho-syntactic strategies for marking definiteness within the noun phrase. Schwarz (2009; 2013) and Jenks (2015) find that these strategies often correspond to distinct characterizations of the semantics of definite descriptions. Many languages feature distinct mechanisms for expressing definite descriptions as either *unique* or *familiar*, such as by having two distinct classes of definite article or by contrasting definite bare nominals with some form of overt definite marking. Cuevas Mixtec shows that a language can also feature internal variation in the marking of either uniqueness or familiarity. Most nominals of this language are capable of taking on bare forms for the expression of uniqueness, while familiarity is expressed using overt definite articles. There are some nominals, however, which never combine with overt definite articles or which must take on definite articles in a larger set of semantic environments. The variation observed here seems to be tied to etymological factors within the nominal and the influence of an animacy hierarchy.

1 Introduction

Recent literature on the proper characterization of definiteness shows that languages vary widely in the strategies they utilize for its expression, from bare nominals to the occurrence of definite articles or even demonstratives. Schwarz (2009) and Jenks (2015) show that when languages feature more than one strategy for the expression of definiteness, the variation exhibited semantically corresponds to distinct notions of definiteness itself. One class of definite expressions will encode *uniqueness* of an individual, such that the descriptive content conveyed by the nominal can only be attributed to that individual. Another class of definiteness expressions will encode anaphoricity or *familiarity*, where the expression invokes an anaphoric link to a previously mentioned individual in



a discourse. Both Schwarz and Jenks find robust cross-linguistic evidence for the validity of a non-uniform approach to the characterization of definiteness, given the great diversity of languages that grammaticize the distinction between uniqueness and familiarity. However, this is far from being the whole story on the nature of definiteness encoding in the nominal domain. Despite the growth of investigation on cross-linguistic variation in the expression of uniqueness and familiarity, there does not yet seem to be thorough investigation on *language internal* variation in the expression of the distinction. This paper brings to light some pertinent details regarding a language with such internal variation, with hopes of contributing to the greater account of definiteness across languages.

Cuevas Mixtec is an Otomanguean language which displays at least two distinct strategies for expressing definiteness. The language features a set of definite articles that are derived from a noun classifier system. Definiteness may also be expressed by bare nouns, which may also have an existentially quantified or generic interpretation in some contexts. The example below demonstrates both strategies at work, where a nominal *īsū* ‘deer’ is interpreted as a definite description, referring to the entirety of the group of organisms that are named such. The occurrence of the definite article *tyí* generally restricts the interpretation of *īsū* ‘deer’ to a definite description, but it is optional in this context so long as another determiner type does not replace it.

- (1) *indyī’i syà’á [(tyí) īsū]*
end.COMPL account the.AML deer
‘The deer went extinct.’

The examples below show the optionality of the definite article *tyà* for the expression of definiteness on a nominal predicate. Within the village of San Miguel Cuevas and surrounding villages, certain festivals are organized by gender-based committees led by an administrator of the same gender. There is therefore a unique male and unique female administrator for the organization of these festivals. In the examples, a character named Juan is being presented as an administrator.¹ The absence of a definite article allows the nominal predicate to be interpreted as definite when uttered within the context of the male village festival committee (or indefinite otherwise). The presence of the article restricts the interpretation of the nominal predicate to a definite one, identifying Juan as the unique male administrator regardless of context.

¹There are two words for this occupation in Cuevas Mixtec, *mastoni* and *mārtòòn*, both of which seem to have originated as loanwords from other language groups. Both words appear in this paper.

(2) Context: Juan is presented as the male administrator within a meeting of men organizing a village festival.

- a. [tyà Juáàn] kúú m̄astóní
 the.SG.M Juan be.IPFV administrator
 ‘Juan is the administrator.’
- b. [tyà Juáàn] kúú [tyà m̄astóní]
 the.SG.M Juan be.IPFV the.SG.M administrator
 ‘Juan is the (male) administrator.’

Recent investigations into languages which feature multiple strategies for definiteness expression have shown that the opposition between the styles of definiteness marking corresponds to a distinction in the notions of definiteness that are being invoked. Schwarz (2009) shows that for the particular case of German, *weak* definite articles encode uniqueness, or the quality of uniquely satisfying the descriptive content of the nominal relative to a situation, while *strong* definite articles encode an *anaphoric* link to a previously mentioned individual. Schwarz (2013) and Jenks (2015) later show that when a language allows bare nominals to serve as definite descriptions, the notion of definiteness expressed similarly tends to be that of uniqueness, while the same language utilizes overt definiteness marking for creating anaphoricity. Cuevas Mixtec is shown in this paper to be very similar to other languages which allow for bare nominals to have definite interpretations, thereby supporting these previous findings. However, the language presents a more complicated picture by displaying internal variation in the correspondences between definiteness marking and the notion of definiteness involved. There seems to be a grouping of nominals into at least three types with respect to the strategy for encoding either uniqueness or anaphoricity. There are those which follow the pattern of reserving bare nominals for expressing uniqueness and utilizing overt marking for familiarity, those which follow an English pattern of utilizing overt marking for both uniqueness and familiarity, and those which cannot host definite articles at all.

In the rest of this paper, I cover the necessary background on the study of both definiteness and Mixtec to introduce the evidence for the claims made above. In §2, I briefly introduce the analysis of definiteness marking for languages which permit bare nominal definite descriptions by Schwarz and Jenks. I provide Schwarz’s examples from Standard German used to demonstrate grammatical sensitivity to the expression of uniqueness and familiarity. In §3, I then introduce some background information on Cuevas Mixtec, which will be necessary for reading the data. I provide a very brief typological introduction to the lan-

guage as well as some background on the speaker community located in western Oaxaca, Mexico, and in California. This is followed by an introduction to the particular orthography of Cuevas Mixtec that is in development, then by a grammatical sketch of the language covering basic word order and the basic grammar of noun classifiers. In §4, I then present evidence for the interpretation of the definite descriptions of the language as either encoding uniqueness or anaphoricity. These are semantic environments where the interpretation of a definite description is restricted to either a uniqueness definite or anaphoric definite, which mutually exclude each other. In §5, the evidence for the correspondence between definiteness marking strategy and notion of definiteness is used again to present evidence for internal variation with respect to that correspondence. Different nominals are compared with each other to establish their definiteness marking preference in the relevant semantic environments. The paper then concludes with a summary of the findings.

2 Definiteness background

This section introduces the key notions of definiteness that will be shown to characterize the definite descriptions of Cuevas Mixtec. Both notions correspond to early attempts at the characterization of English definite descriptions, or nominal expressions with *the*. Schwarz (2009) shows that both approaches are validated by cross-linguistic variation in the semantics of definite descriptions and the alternative strategies languages employ to encode definiteness. Cuevas Mixtec provides further support for an approach to the semantics of definiteness which considers internal variation in the number of strategies for composing definite descriptions.

2.1 Uniqueness and familiarity

There has been a long debate regarding the most proper semantic characterization of definite descriptions, and two approaches in this respect have been more prominent. There is a *uniqueness* approach, which claims that definiteness is the function of referring to an entity that is the unique bearer of the property denoted by the nominal description. The quality of uniqueness need not be absolute, but evaluated relative to some contextual domain or situation. Examples of felicitous uses of English definite articles expressing uniqueness include *the president of the United States* and *the Taj Mahal*, where each expression refers to a thing that uniquely satisfies the nominal description with respect to some domain. In

these cases, the domains are quite broad and seemingly absolute, at least when constrained to a small scope in time, but expressions like *the projector* are also interpreted as unique within smaller contexts, despite their non-uniqueness in the world at large. In the following example, *the projector* felicitously refers to a uniquely identifiable entity when uttered in the context of a lecture hall where there is a single projector. When constrained to such a context, there is nothing else for the expression to refer to.

- (3) Context: A presentation is about to start within the lecture hall of a school.

The projector is not being used today. (Schwarz 2013: 3)

It does not matter that there are other projectors in the greater building beyond the lecture hall, which represents a broader domain. There is a communicative mechanism whereby the speaker constrains a domain so as to ensure the uniqueness of the definite description's referent within it. Similarly, expressions like *the dog* or *the professor* can also be unique within small domains such as a family unit or a classroom. In predicate logic, the condition of uniqueness can be expressed as universal quantification over the equivalence of referents of a nominal predicate.

- (4) $\exists x[P(x) \ \& \ \forall y[P(y) \Rightarrow y = x]]$

'There is an x that is P and all y that are P are identical to x ' (Schwarz 2013: 3)

The second common approach to characterizing definiteness is the *familiarity* approach, which claims that definiteness is the function of referring to an entity that is familiar or salient to discourse participants. Researchers have touched on a number of ways that familiarity itself could be characterized, such as perceptual accessibility or salience in cultural institutions. Roberts (2003) distinguishes between two kinds of familiarity, *weak* and *strong*, which outline the distinct notions of familiarity according to linguistic input. Weak familiarity corresponds to a broad variety of mechanisms for identifying the referent of an expression beyond linguistic input. Strong familiarity is more precise by its characterization as the function of creating an anaphor to a previous linguistic expression in a discourse. The following example illustrates this usage, in which *the book* is an expression used to further comment on an entity already introduced earlier by *a book*.

- (5) *John bought a book and a magazine. The book was expensive.* (Schwarz 2013: 3)

As an anaphoric expression, it is important for the definite description to be preceded by an *antecedent*, served by the expression *a book* in the previous example. Without the antecedent, the definite description lacks a referent to refer to for the anaphoric usage, and the expression will become awkward, as in the example below.

- (6) *John bought a newspaper and a magazine. #The book was cheap.*

The anaphoric use of definite descriptions can be semantically modeled as an elaboration on their uniqueness usages with an additional condition. Schwarz (2009) claims that familiarity definites feature an additional index argument 1 which receives an interpretation from an assignment function g . The assignment function in turn maps the index to the individual introduced by an appropriate indefinite, essentially building a pronoun into the meaning of the definite description.

- (7) $\iota x.P(x) \ \& \ x = g(1)$

‘The unique x that is both P and identical to the individual interpreted from the assignment function g on the index 1’

Recent literature on definiteness has been more concerned with strong familiarity, and since this notion is more relevant to the discussion of definiteness in this paper, it will be referred to simply as *familiarity* throughout.

2.2 Weak and strong articles of German

Cross-linguistic investigations on definiteness in general have found good evidence for the adequacy of both approaches outlined above, with some languages even distinguishing the two characterizations of definiteness grammatically. Schwarz (2009; 2013) shows that various Germanic languages which feature two distinct classes of definite article exhibit a correspondence between the definite articles’ meanings and the two dominant analyses of definiteness. For example, Standard German features two distinct classes of definite article whose morphological differences are apparent by their interaction with prepositions. Standard German strong articles like *dem* in the example below resist morphological fusion with the preposition, while weak articles fuse with prepositions. The articles are otherwise similar in appearance and pronunciation.

- (8) German (Schwarz 2009: 14)
- a. *Hans ging zu dem Haus.*
Hans went to the_{strong} house
‘Hans went to the house.’

- b. *Hans ging zum Haus.*
 Hans went to.the_{weak} house
 ‘Hans went to the house.’

Schwarz finds a distinction in the meanings each class of definite article contributes. Weak articles are uniqueness definites that highlight a relatively unique individual and generally cannot be used to compose anaphora. In a sentence such as (9), the weak article establishes the relative uniqueness of the referent of *Mond* ‘moon’ in a broad domain such as Earth.

- (9) German (Schwarz 2009: 40)
Armstrong flog als erster zum Mond.
 Armstrong flew as first.one to.the_{weak} moon
 ‘Armstrong was the first one to fly to the moon.’

Schwarz also finds that strong articles are familiarity definites, which create an anaphoric link between a definite description and its antecedent, and they cannot create reference to an individual that has not yet been mentioned in the discourse. The strong article therefore creates an anaphoric link between the two utterances of *Buch* ‘book’ in the example below, such that the utterances refer to the same individual. For comparison, the weak article lacks the necessary anaphoric properties required to link the two utterances of *Buch* to a common referent, a book about sunchokes (*Topinambur*). It does not help either that the referent of *Buch* is not very unique in the context of the New York Public Library.

- (10) German (Schwarz 2009: 30)
In der New Yorker Bibliothek gibt es ein Buch über Topinambur.
 in the New York library exists EXPL a book about topinambur
Neulich war ich dort und habe #im / in dem Buch nach einer
 recently was I there and have in.the_{weak} / in the_{strong} book for an
Antwort auf die Frage gesucht, ob man Topinambur grillen
 answer to the question searched whether one topinambur grill
kann.
 can
 ‘In the New York Public Library, there is a book about topinambur.
 Recently, I was there and searched in the book for an answer to the
 question of whether one can grill topinambur.’

The strong article itself also becomes awkward when combined with nominals without an antecedent. In the example below with *Bürgermeister* ‘mayor’, there is no previous mention of a mayor to serve as an antecedent to the definite description.

(11) German (Schwarz 2009: 40)

Der Empfang wurde vom / #*von dem Bürgermeister eröffnet.*
the reception was by.the_{weak} / by the_{strong} mayor opened
'The reception was opened by the mayor.'

Schwarz thus presents strong evidence for a correspondence between the notion of definiteness (i.e. uniqueness, familiarity) and the morpho-syntactic realization of the definite article in Standard German. Similar results are robustly exhibited for the distinct types of definite article in another Germanic language, Fering, with data from Ebert (1971).

2.3 Bridging

Before moving on to the discussion of definiteness marking in languages outside of the Germanic family, it is worth noting a final set of discourse environments where preferences between weak and strong articles have been displayed. When Hawkins (1978) set out to understand the semantic source of grammatical differences between definite and indefinite descriptions, he laid out a preliminary taxonomy of the distinct uses of the definite article to be later accounted for in linguistic models. From the taxonomy, anaphoric uses were those which inspired the familiarity approaches to the semantics of definite descriptions. Additionally, *immediate situation* uses and *larger situation* uses were those which inspired the uniqueness approaches, differentiating between smaller and larger spaces within which uniqueness is evaluated. If the domain within which uniqueness is evaluated is a current and localized space where the utterance occurs, this usage may be described as an *immediate situation* use. If the domain is instead a broad or global one, considering large expanses of space beyond the utterance situation, this usage may be described as a *larger situation* use. Schwarz uses these discourse environments to test preference for weak or strong articles within nominal expressions and finds clear correspondences between semantic environment and article preference.

Hawkins also discussed a fourth usage which has seen mixed results in Schwarz's assessment of sensitivity to the presence of weak and strong articles. Cases of associative anaphora, or *bridging* (Clark 1975), constitute anaphoric uses of definite descriptions whereby the antecedent is not coreferential, but it refers to an item or circumstance which stands in some relation to the referent. The example below shows an anaphoric use of the definite description *the ceiling* where there is no previous mention of a ceiling. However, the existence of a room would

entail in common world knowledge the existence of a unique ceiling without explicit mention of one.

(12) *I looked into the room. **The ceiling** was very high.* (Clark 1975: 171)

Schwarz finds that cases of bridging in Standard German generally have no preference for either the weak or strong article. However, there are certain sub-cases of bridging that do demonstrate preferences, depending on the kind of relationship that is exhibited between the definite description and its antecedent. Weak articles seem preferred when the relationship between the definite description and the antecedent is that of a *part-whole* relationship, in which the referents of both expressions relate to each other as though one were an appendage of the other. This is demonstrable through the example of a fridge and its crisper. In the example below, the nominal *Gemüsefach* ‘crisper’ prefers co-occurrence with the weak article.

(13) German (Schwarz 2009: 52)

Der Kühlschrank war so groß, dass der Kürbis problemlos
the fridge was so big that the pumpkin without.problem
im / #**in dem** **Gemüsefach** untergebracht werden konnte.
in.the_{weak} / in the_{strong} crisper stowed be could
‘The fridge was so big that the pumpkin could easily be stowed in the crisper.’

Strong articles seem to be preferred when the relationship is something other than a part-whole relationship, such as a relation in which the antecedent refers to a producer of the referent of the definite description. This is demonstrable through the example of a play and its author. In the example below, the nominal *Autor* ‘author’ prefers co-occurrence with the strong article.

(14) German (Schwarz 2009: 53)

Das Theaterstück missfiel dem Kritiker so sehr, dass er in seiner
the play displeased the critic so much that he in his
Besprechung kein gutes Haar #am / *an dem* *Autor* ließ.
review no good hair on.the_{weak} / on the_{strong} author left
‘The play displeased the critic so much that he tore the author to pieces in his review.’

It is interesting why these scenarios should display such preferences that are dependent on the kind of relationship established between the definite description and its antecedent. To answer for the case of the part-whole relationship,

Schwarz (2009) suggests that the preference of definiteness marking associated with uniqueness derives from an analysis of part-whole relationships as expressing decomposable situations that entail unique parts. Given some situation, such as one in which there is a car, one can reasonably assume the existence of unique parts according to common knowledge, such as a car's license plate. Product-producer relationships then differ from part-whole relationships because of how detachable a producer can be from a product across possible situations, requiring some additional mechanism for the construction of bridging.

2.4 Cross-linguistic variation

Beyond the German data, Schwarz (2013) finds that many other languages which display similar internal variation in strategies of definiteness marking also associate these strategies with either uniqueness or familiarity readings. In a brief cross-linguistic survey of how the two notions of definiteness are expressed across languages, he shows that not only do Lakhota and Hausa feature two distinct types of definite article, they also display a parallel phenomenon to German in associating these articles with either uniqueness or familiarity readings. Schwarz also shows that another common strategy for the expression of definiteness across languages is to utilize bare forms of nominal expressions. Languages like Akan and Mauritian Creole widely feature bare nominals as definite descriptions in their grammars. Schwarz further notes that the interpretation of these definite bare nominals tends to be only that of uniqueness, parallel to the interpretation of weak articles in standard German. The example from Mauritian Creole below shows two bare nominals *later* 'earth' and *soley* 'sun' serving as definite descriptions, denoting two individuals that are uniquely characterized by their descriptions in a global domain.

- (15) Mauritian Creole (Wespel 2008: 150; source: O.M.49)

Later turn otur soley.

Earth revolve around sun

'The Earth moves around the Sun.'

In order to express familiarity, the same languages will employ overt modification on nominals, sometimes in the form of definite articles specifically reserved for familiarity uses. The parallels observed in the data from these languages and Standard German are even encountered in cases of bridging, where a grammatical sensitivity to part-whole and product-producer relationships is displayed. Part-whole relationships favor uniqueness-expressing, bare nominals as

definite descriptions, while product producer relationships favor overt definiteness marking. Jenks (2015) confirms Schwarz's observations on languages with definite bare nominals by presenting data on Thai, in which bare nominals indeed may have uniqueness readings, while familiarity is expressed through the use of demonstratives. He further claims that the findings are replicable for several other numeral classifier languages.

In the rest of this paper, it is shown that Cuevas Mixtec mostly patterns with Akan and Mauritian Creole in employing bare nominals for the expression of uniqueness, while familiarity is expressed using a series of definite articles which encode noun class. However, this generalization only serves for the distinction between uniqueness and familiarity for a large subset of the nominal inventory of the language. There are some cases of nominals for which bare forms are more restricted in their distribution, forcing overt definite articles to also take on uniqueness interpretations. This alternative pattern more closely resembles the strategy of definiteness marking in English, where there is a single definite article for the expression of both uniqueness and familiarity. The choice of which nominals are selected for either strategy appears to be systematic, as nominals displaying the English strategy tend to be predicates of humanity or personhood. Ultimately, the paper shows that languages like Cuevas Mixtec can display internal variation in the strategy for definiteness marking, with input from the lexical semantics of nominal predicates.

3 Background on Cuevas Mixtec

This section presents some historical and linguistic background on the language of interest for this paper, Cuevas Mixtec. It first very briefly introduces the Mixtec family of languages in a historical context. It then introduces some phonological details, along with the working orthography for Cuevas Mixtec in which the data are written up. Finally, a brief sketch of some word order patterns observed in Cuevas Mixtec is presented. Although the purpose of this paper is not to flesh out the phrase structures of the language, some familiarity with basic sentence structure is helpful for interpreting the data on definiteness expressions later.

3.1 Mixtec language family and Cuevas Mixtec

Mixtec is a family of languages which are indigenous to the Mixteca region of southern Mexico. Mixtec speakers are encountered in villages and cities throughout the Mixteca, which encompasses much of the western half of Oaxaca state

and includes parts of neighboring Guerrero and Puebla states, an area altogether covering roughly 10,000 square miles (Bradley & Hollenbach 1988). In 1988, the Mixtec speaker population was almost 250,000 people, although this population had grown to 477,995 people according to the Mexican census for the year 2010 (INEGI 2010). The language family has been described as being composed of about 20 mutually unintelligible languages and their variants² (Bradley & Hollenbach 1988). Each village features its own variant of Mixtec, with phonological, syntactic, and lexical idiosyncracies. Mutual intelligibility between variants is often restricted to villages in close geographic proximity, and often enough two villages that speak different Mixtec languages are near each other. There is no widespread or standard variety of Mixtec, although the variants have been able to be categorized into groups according to mutual intelligibility (Egland 1978: 25–37) and historical sound changes (Josserand 1983). For these reasons, grammatical descriptions of Mixtec languages highlight the village of origin for the variant of Mixtec described, as this paper does.

The Mixtec language family belongs to the greater Otomanguean language stock distributed throughout central and southern Mexico today (Rensch 1976). Features common to all Otomanguean languages include isolating morphology and significant representation of morphemic suprasegmental features, such as tone and voice quality. Because of the high presence of these features in Otomanguean, languages like Mixtec have been subject to a wealth of phonetic and phonological research. In contrast, research into Mixtec for the sake of syntactic (or semantic) description is much less abundant (Bradley & Hollenbach 1988). Within Otomanguean, Mixtec is further grouped with Triqui and Cuicatec into the Mixtecan language family, spread throughout the western half of Oaxaca, eastern Guerrero, and southern Puebla.

Although Mixtec speakers are often thought of as a single ethnic group by outsiders, Mixtecs themselves tend not to identify with each other in such a manner. The terrific linguistic diversity found within the Mixtec language family is reflective of an old culture of village-based ethnocentricity. Mixtec speakers in Mexico often identify with their home village as a source for ethnic identity (Spores & Balkansky 2013: 221–223). They much less identify with a broader Mixtec sociolinguistic heritage, and this is apparent in the history of resource competition and intercommunity conflict in the Mixteca region, recorded since before the

²The term *variant* here is a common substitute for *dialect* in discussion about languages of Mexico. The term *dialect* has certain political and derogatory connotations in the Mexican and Latin American context that are preferably avoided. The terms *variant* and *variety* replace *dialect* in order to disambiguate reference to the high degree of mutual intelligibility that one speech community has with another.

Spanish Colonial period. This trend would change somewhat for the thousands of Mixtec temporary laborers and immigrants who had moved to northern Mexico and the United States towards the end of the twentieth century. Mixtecs among the United States diaspora have accommodated broader ethnic affiliations with other Mixtecs, and even other Mexicans of Oaxacan origin, as a response to their alienating circumstances as migrant workers (Velasco Ortiz 2005; Spores & Balkansky 2013: 228–235). Many have organized and formed interest groups around issues pertaining to the plight of the broader Oaxacan migrant community in the United States and Northern Mexico. However, despite these new developments, Mixtec migrants retain strong hometown or village affiliations, and this phenomenon has gone hand in hand with Mixtec dialectal diversity for at least several centuries.



Figure 1: San Miguel Cuevas in northwest Oaxaca (personal elaboration)

This paper concentrates on data from Cuevas Mixtec, the particular variant of Mixtec spoken in the village of San Miguel Cuevas, or *ñũũ★ nũũ★ yũkũ*³ ‘the village on the mountain’ as it is named in the variant. This village is located in the municipality (or *municipio*) of Santiago Juxtlahuaca, southwest of the munic-

³The stars here are introduced later as marking the presence of floating tones.

ipal seat (Figure 1). This location might put Cuevas Mixtec in Josserand's (1983) classification as a variety of Southern Lowlands Mixtec. According to the 2010 Mexican Census, this village had a population of 522 inhabitants (242 male and 280 female) with 441 citizens over the age of three years of age that spoke an indigenous language (233 male and 208 female) (INEGI 2010). The local name for the variant of Mixtec spoken here is *tù'ūn ndá'ví* 'poor language', although there is a movement to replace this manner of referring to the language with *tù'ūn sàvì* 'rain language' or 'Dzahui's language'.⁴ Due to early twentieth century educational policy against the retention of indigenous languages in Mexico, few Mixtec speakers are trained in written forms of their languages (Velasco Ortiz 2005: 29), and San Miguel Cuevas has yet to see a standardized, written form of theirs.

Beyond San Miguel Cuevas, speakers of this variant are also found in the United States, having immigrated to take on jobs in the service industry, manufacturing, and agriculture. Most of these speakers are immigrants born in Mexico, and they are located in Delaware, the Portland metropolitan area of Oregon, and Fresno county in California. In Fresno, the speaker population is absorbed into the greater Mixtec or Oaxacan community, which also has significant numbers of Mixtecs from Yucuquimi de Ocampo and Santiago Tilantongo in Oaxaca, and Metlatónoc and Jicayán de Tovar in Guerrero. The variants of Mixtec spoken by members of different towns may differ to the extent that Spanish is preferred as a means of communication, and Cuevas Mixtec is therefore not widely spoken outside the home. Within the home, Cuevas Mixtec is spoken more frequently to varying degrees. Some local radio stations have accommodated some programming in several local varieties of Mixtec at special times, though I am not certain that they have had programming in Cuevas Mixtec. Local rap artist Miguel "Una Isu" Villegas has incorporated Cuevas Mixtec into the lyrics of several of his songs, and these songs are available on several media-sharing websites.

⁴*Dzahui* is the name of the Mesoamerican rain deity that appears in Mixtec codices and ancient stone carvings. The movement to rename all Mixtec languages as local translations of 'rain language' or 'Dzahui's language' has spread into much of the Mixteca region besides San Miguel Cuevas, though I have not been able to trace its origin or motivation. Some motivation may come from the fact that the veneration of rain deities is a practice of the native Mixtec religion which has survived the imposition of Catholicism in the colonial era. In San Miguel Cuevas, there is a special stone named Saint Michael which is provided offerings in exchange for the prospect of rain. I have also been told about a stone of similar purpose in Ixpantepec Nieves which has retained the name 'rain' or *Dzahui*.

3.2 Orthography of Cuevas Mixtec

For each example sentence of Mixtec throughout this paper, I present the data with three transcription tiers: transcription, morpheme gloss, and translation. Transcriptions are written using a variant of the official Mixtec orthography endorsed by the Academy of the Mixtec Language (*Academia de la Lengua Mixteca* or *Ve'e Tu'un Savi*), instead of phonetic symbols from the International Phonetic Alphabet (IPA). This is intended to facilitate reading for Mixtec scholars and other readers familiar with this orthography. Table 1 presents the current alphabet for Cuevas Mixtec, with corresponding IPA symbols for each entry. There are five oral vowels *a e i o u*, voiceless affricate and plosives *ch k ku p t ty*,⁵ voiced plosive *d*,⁶ prenasalized stops *mb nd ndy ng*, nasal stops *m n ñ*, voiceless fricatives *s sy x*,⁷ voiced fricatives *v y*, and liquids *l r*.⁸ There is additionally a glottal stop of ambiguous phonemic status, and this is written with an apostrophe.

Table 1: Cuevas Mixtec orthography and phone correspondences

<i>a</i>	<i>b</i>	<i>ch</i>	<i>d</i>	<i>e</i>	<i>i</i>
/a/	/b/	/tʃ/	/d/	/e/	/i/
<i>j</i>	<i>ju</i>	<i>k</i>	<i>ku</i>	<i>l</i>	<i>m</i>
/x/	/x ^w /	/k/	/k ^w /	/l/	/m/
<i>mb</i>	<i>n</i>	<i>nd</i>	<i>ndy</i>	<i>ng</i>	<i>ñ</i>
/ ^m p/	/n/	/ ^m t/	/ ⁿ c/	/ ^ŋ k/	/ɲ/
<i>o</i>	<i>p</i>	<i>r</i>	<i>s</i>	<i>sy</i>	<i>t</i>
/o/	/p/	/r/	/s/	/ç/	/t/
<i>ty</i>	<i>u</i>	<i>v</i>	<i>x</i>	<i>y</i>	'
/c/	/u/	/β/	/ʃ/	/z/	/ʔ/

Otomanguan languages are well known for the preponderance of suprasegmental features, and Cuevas Mixtec is no exception. Nasalized vowels are repre-

⁵The reader may notice that the letter *u* is used for representing both a vowel and a secondary feature of two consonants *ku* and *ju*. In the data, tone marking always occurs on vowel symbols, including those for /u/. This distinguishes the occurrence *u* as a vowel from its occurrences in consonant digraphs, which do not take tone marking.

⁶The voiced plosive *d* seems to only occur on one pronoun and is likely an allophone of the voiceless plosive *t*.

⁷The plosive *b* and the fricatives *j ju* occur in loanwords and proper names from Spanish.

⁸The tap *r* seems to only occur in some pronouns and may be an allophone of *ty*.

sented with an adjacent *n* as in *an en in un*. Cuevas Mixtec is also a highly tonal language, with three level tones that combine into nine possible contours. High tones, low tones, and mid tones are marked with acute accent *á*, grave accent *à*, and macron *ā*, respectively. With the help of more recent work on similar varieties of Mixtec (Carroll 2015), I have also been able to identify the presence of floating tones.⁹ I have not thoroughly investigated the distribution of floating tones, so they are inconsistently marked in the data. Where they are marked, the pending convention I have chosen is to represent their presence with a star *★*.

Many pronouns in the language have cliticized forms. For clitics, I stray from the Academy of the Mixtec Language and follow a convention observed in Bradley & Hollenbach's (1988) *Studies in the syntax of Mixtecan languages* series by representing clitics as orthographically detached from host words. The detachment is represented by horizontal space in the data, and therefore, no special marking for clitics is used.¹⁰ Finally, the data itself throughout this paper is marked for acceptability as a phrase or sentence of the language, unacceptability, infelicity, and spontaneous elicitation. Acceptable phrases and sentences are marked with a checkmark *✓*, semantically or grammatically anomalous phrases are marked with an asterisk ***, and infelicitous sentences are marked with a pound sign *#*. Spontaneously elicited phrases and sentences, or those which were produced by a speaker in speech or translation, are unmarked.

3.3 Word order patterns of Cuevas Mixtec

This subsection covers basic word order patterns encountered in the language in order to facilitate reading of definiteness data later. Basic sentence structure is presented first, and Cuevas Mixtec is shown to be a VSO language with certain conditions for optional or obligatory repositioning of verb arguments to a preverbal position. Some aspects of the structure of the noun phrase are presented afterwards in order to demonstrate the distribution of definite articles with respect to other modifiers later. The subsection then presents examples of the distribution of noun classifiers, which have occurrences as the definite articles of the language.

⁹Floating tones are applied to the first vowel of the following word, and their value depends on the tone value of the last vowel of the word they originate from. They manifest as high tones when the tone of the last vowel is low and as low tones otherwise. Therefore, the floating tone from *nùù★* 'face' will be high, while the floating tone from *chítú★* 'cat' will be low.

¹⁰The result of this convention is the lack of representation of data where clitics combine with truncated forms of host words. Truncation often occurs on long vowels or [V?V] strings after a clitic without a consonant is attached.

3.3.1 **Basic sentence structure**

Cuevas Mixtec is a verb-initial language, like many languages of the Mesoamerican area. The subject argument of a verb consistently follows the verb if it is a clitic pronoun. The object argument follows the subject in transitive sentences.

- (16) ✓ *kíxi* *rā*
 sleep.IPFV 3SG.M
 ‘He is sleeping.’

- (17) ✓ *kúnì* *rí* *tyikuii*
 want.IPFV 3SG.AML water
 ‘It wants water.’

Acceptable subject placement varies when the subject argument is not a clitic but a full nominal. SVO word order is often the preferred word order for sentences with non-clitic subject arguments uttered without a discourse context.¹¹ VSO word order in this case is often strange without a discourse context presented beforehand.

- (18) Context: ∅
 a. ✓ [*tyà* *Juáàn*] *isyūn* [*ūn kárró*]
 the.SG.M Juan buy.COMPL one car
 ‘Juan bought a car.’
 b. # *isyūn* [*tyà* *Juáàn*] [*ūn kárró*]
 buy.COMPL the.SG.M Juan one car
 (‘Juan bought a car.’)

- (19) Context: ∅
 a. [*ndyí’i vā nā*] *isyi’i*
 all FOC 3.HUM die.COMPL
 ‘Everyone died.’

¹¹Example (19) features a focus-sensitive particle *vā* which occurs in many other examples throughout the data. It serves many roles such as emphasizer, restrictive/exclusive particle, and aspectual particle, similar to English *just*. Its role in (19) is uncertain, though speakers note that it is optional in this case.

- b. # *isyi'i* [ndyī'i vā nā]
 die.COMPL all FOC 3.HUM
 ('Everyone died.')

VSO word order becomes preferred with the addition of adverbial modifiers. Temporal adverbs like *ikū* 'yesterday' allow for non-clitic subjects to occur postverbally without a discourse context.

- (20) Context: ∅
 ✓ *ikū isyīn [tyà Juáàn] [īn kárró]*
 yesterday buy.COMPL the.SG.M Juan one car
 'Yesterday, Juan bought a car.'

Wh-questions also confirm the basic word order to be VSO. Cuevas Mixtec features obligatorily preposed wh-words in wh-questions. Even if the wh-word is not a verbal argument, verbal arguments are unable to occur between the verb and the wh-word.

- (21) a. ✓ *ndyí isyīn [tyà juáàn] [īn kárró]*
 where buy.COMPL the.SG.M Juan one car
 'Where did Juan buy a car?'
 b. * *ndyí [tyà juáàn] isyīn [īn kárró]*
 where the.SG.M Juan buy.COMPL one car
 ('Where did Juan buy a car?')

There are some instances of a clitic pronoun co-occurring with a preverbal nominal as a coreferential item, similar to a resumptive pronoun or overt trace.

- (22) *[tyà Juáàn] isyīn rā [īn kárró]*
 the.SG.M Juan buy.COMPL 3SG.M one car
 'Juan bought a car.'

This seems to indicate a sort of topicalization strategy, where the preverbal nominal occurs in a topic position while the pronoun serves as the true verb argument. There are three reasons for suggesting this proposal. First, conjunction of sentences shows that preverbal subject arguments are restricted in their distribution when these resumptive pronouns occur. A preverbal subject argument cannot occur for each conjunct sentence when the resumptive pronoun occurs in each. If each conjunct sentence has a resumptive pronoun, the preverbal subject

argument occurs once for the entire utterance, taking scope above the conjunction itself. Preverbal subjects may occur within each conjunct sentence as long as there is no resumptive pronoun present.

- (23) a. * [ndyí'í vā nā] syítā nā tyā [ndyí'í vā nā]
 all FOC 3.HUM sing.IPFV 3.HUM and all FOC 3.HUM
 syítāsyà'á nā
 dance.IPFV 3.HUM
 ('Everyone sings, and everyone dances.')
- b. ✓ [ndyí'í vā nā] [syítā nā tyā syítāsyà'á nā]
 all FOC 3.HUM sing.IPFV 3.HUM and dance.IPFV 3.HUM
 'As for everyone, they sing and dance.'
- c. ✓ [ndyí'í vā nā] syítā tyā [ndyí'í vā nā] syítāsyà'á
 all FOC 3.HUM sing.IPFV and all FOC 3.HUM dance.IPFV
 'Everyone sings and dances.'

Secondly, there are certain types of modified nominals which would be barred from serving as topics as they are non-referential, such as negated nominals. If a negated nominal occurs in the preverbal position, and it is interpreted as the sentence subject, a clitic subject cannot occur in the subject position after the verb(s).

- (24) [nì īn nà tyāā] kúnì (*nā) kūsū (*nā)
 not.even one the.HUM man want.IPFV 3.HUM sleep.IRR 3.HUM
 'No men wanted to sleep.'
- (25) [cháá nà tyāā] syítā (*nā)
 less the.HUM man sing.IPFV 3.HUM
 'Few men are singing.'

Thirdly, preverbal nominals are crucial for the expression of generic statements. Postverbal subject arguments force a progressive aspectual interpretation of the sentence below, while preverbal subject arguments allow for a topic-comment reading of the same material. It is not crucial that the resumptive pronoun occur for triggering the topic-comment reading.

- (26) a. syéí [tyí chítú★] tyiún★
 eat.IPFV the.AML cat mouse
 'The cat is eating mice.'

- b. [tyí chítú★] syéí (rí) tyiín★
 the.AML cat eat.IPFV 3SG.AML mouse
 ‘The cat eats mice.’

Without the resumptive pronoun, the sentence is interpreted as being an answer to a question. This might suggest that preverbal arguments without co-occurring resumptive pronouns are focalized.

3.3.2 Basic noun phrase structure

Nouns in Cuevas Mixtec do not require modification in order to occur as verb arguments. They frequently occur in bare forms and are often interpreted as indefinites in such cases.

- (27) chítú★ syéí rí tyiín★
 cat eat.IPFV 3SG.AML mouse
 ‘Cats eat mice.’

Different classes of nominal modifiers occur before or after the nominal. There are at least four classes of items which may occur prenominally: quantifiers, numerals, definite articles, and a specifier. The specifier *mīi* serves as a reflexive when modifying a pronoun, as in the case of *mīi rā* ‘himself’.

- (28) [tyà Juáàn] kúni rā [ná kūsū mīi rā]
 the.SG.M Juan want.IPFV 3SG.M COMP sleep.IRR SPEC 3SG.M
 ‘Juan wants that he himself sleep.’

While modifying a nominal, the function of the specifier seems to be that of encoding focus, or the presentation of the modified nominal as new information.

- (29) [mīi tyà Juáàn] sátátá
 SPEC the.SG.M Juan heal.IPFV
 ‘Juan is the one healing others.’

- (30) ✓ [mīi tyinā] ndé’i
 SPEC dog bark.IPFV
 ‘It is the dog barking.’

Quantifiers occur in a prenominal position. The examples below include the quantifiers *ndyī’i* ‘all’ and *cháá* ‘few’.

- (31) [**ndyī'ī** tyinā] ndé'ī
 all dog cry.IPFV
 'All dogs bark.'
- (32) [**cháá** tyinā] kúmí ndā'à
 few dog have.IPFV hand
 'Few dogs have hands.'
- (33) [**ndyikuì** nìi] và'à ā
 all salt good 3.INA
 'All salt is good.'

Quantifiers do not seem to co-occur with the specifier. This might suggest that both quantifiers and the specifier form a grammatical class.

- (34) * [**mī** **ndyī'ī** tyinā] ndé'ī
 SPEC all dog cry.IPFV
 ('It is all dogs that bark.')

Numerals also occur prenominaly, though they differ from quantifiers in being able to co-occur with the specifier. The examples below feature the numeral *ù'un* 'five'.

- (35) ✓ *isyīnī* *rā* [**ù'un** *nāndyū*]
 see.COMPL 3SG.M five sun
 'He saw five suns.'
- (36) ✓ [**mī** **ù'un** tyinā] ndé'ī
 SPEC five dog cry.IPFV
 'It is five dogs that are barking.'

Quantifiers differ amongst themselves in their capacity to co-occur with numerals. The quantifier *ndyī'ī* 'all' seems to be able to co-occur with numerals, but *sāvā* 'half' cannot.

- (37) **ndyī'ī** **kùmì** tyinā yó'ō
 all four dog here
 'the four dogs here'

- (38) * [sāvā ùsyì tyìnā] ndé'í
half ten dog cry.IPFV
(‘Half of the ten dogs are barking.’)

The quantifier *ndyí'í* has the ability to syllabically reduce in cases where numerals co-occur, while reduction is not possible before a bare noun. This suggests that the item is not identical to other instances of the quantifier *ndyí'í*.

- (39) *ndyí kùmi tyìnā yó'ō*
all four dog here
(‘the four dogs here’)

- (40) [*ndyí ùnà tyàā*] isyí'í
all eight man die.COMPL
(‘The eight men died.’)

- (41) [*ndyí ùvi tyìnā*] ndé'í
all two dog cry.IPFV
(‘The two dogs bark.’)

- (42) * *ndyí isū*
all deer
(‘all deer’)

A large number of items may occur postnominally, including demonstratives and relative clauses. The following example features a demonstrative *káā* ‘over there’, which follows the nominal within the noun phrase.

- (43) [*tyà Juààn*] isyā'àn rā [ñūū★ *káā*]
the.SG.M Juan go.COMPL 3SG.M village over.there
(‘Juan went to the village over there.’)

3.3.3 Noun classifiers and their functions

Cuevas Mixtec features a robust grammatical gender system which is exhibited through both its pronoun and noun classifier inventories. Noun classifiers in Cuevas Mixtec are semi-pronominal items which explicate, and are sensitive to,

the underlying system of grammatical gender in the language. They are semi-pronominal because, unlike pronouns, they are typically not interchangeable with nominals. They exhibit several grammatical functions in the grammar of the language, including at least their uses as definite articles and relative pronouns, as is shown in this subsection. Table 2 provides the inventory of noun classifiers in the language. They often phonotactically resemble their cliticized pronominal counterparts, but not in all cases. These noun classifiers are not unique to Cuevas Mixtec among Mixtec languages, and one may find their analogues across the family. They are called *prestressed pronouns* in Bradley & Hollenbach's (1988) *Studies in the syntax of Mixtecan languages* series, where they are described for several very different varieties of Mixtec. The Mixtec languages differ widely in the exact inventory of genders that are recognized grammatically. Macri (1983) observes the gender systems of six different Mixtec varieties. All of these varieties had masculine, feminine, and animal genders, though they differed in recognizing inanimate, youth, liquid, and sacred genders.

Table 2: Classifiers vs. pronouns

Gender	Classifier	Pronoun
M	<i>tyà</i>	<i>rā</i>
F	<i>ñá</i>	<i>ñá</i>
YTH	<i>tā</i>	<i>syī</i>
HUM	<i>nà</i>	<i>nā</i>
AML	<i>tyí</i>	<i>rí</i>
STR	<i>tú</i>	<i>dú</i>
LIQ	<i>ndrá</i>	<i>rá</i>
INA	<i>ñà</i>	<i>ñā</i>

The grammatical function of these classifiers that is of primary interest for this paper is their occurrences as definite articles, although their uses expand beyond these cases. When occurring in the pronominal position, these items contribute a meaning of a familiar individual which satisfies the nominal description. Since they encode gender, they show agreement constraints which bar a noun classifier from modifying a noun with a conflicting inherent gender.

- (44) *tyà* *tyàā*
 the.SG.M man
 'the man'

- (45) * *ndrá tyāā*
the.LIQ man
(‘the man’)

Among the prenominal modifiers, definite articles are the most adjacent to the noun. They seem to be able to co-occur with all other prenominal items. While co-occurring with quantifiers or numerals, they explicate the formation of partitive constructions.

- (46) [*sāvā tyí tyinā*] *ndé’i*
half the.AML dog cry.IPFV
‘Half of the dogs bark/are barking.’

- (47) [*ùvi nà tyāā*] *kuànū’ù*
two the.HUM man go.home.IPFV
‘Two of the men went home.’

- (48) [*ùvi tyí tyinā*] *ndé’i*
two the.AML dog cry.IPFV
‘Two of the dogs bark.’

They even co-occur with whatever combinations of quantifier and numeral are possible in the language.

- (49) [*ndyī ùvi nà tyāā*] *isyi’i*
all two the.HUM man die.COMPL
‘The two of the men died.’

- (50) [*ndyī ùvi tyí tyinā*] *ndé’i*
all two the.AML dog cry.IPFV
‘The two of the dogs bark.’

Noun classifiers prescriptively occur with proper names to denote individuals, although they may be dropped from names in very casual speech. Proper names without classifiers also refer to names themselves.

- (51) [(*tyà*) *Kōrnélió*] *kúú* [*tyà káā*]
the.SG.M Cornelio be.IPFV the.SG.M over.there
‘That guy over there is Cornelio.’

- (52) [tyà yó'ò] nāñí rā **Kōrnélió**
 the.SG.M here be.called.IPFV 3SG.M Cornelio
 'This guy is called Cornelio.'

In addition to nouns, noun classifiers also modify adjectives, functioning as nominalizers while encoding definiteness.

- (53) **ñà** yó'ò
 the.INA here
 'this one here'

- (54) **ñà** kuíi
 the.INA green
 'the green one'

This strategy of nominalization extends to verb phrases, forming what appear to be light-headed relative clauses.

- (55) **ñà** isyā'ā [ñá Mária] [nùù★ rā]
 the.INA give.COMPL the.F Maria face 3SG.M
 'the one that Maria gave to him'

They also serve as relative pronouns in the sense that they introduce a relative clause which bears a full nominal head. Agreement in gender between the relative pronoun and the relative clause head remains just as important as between nominals and definite articles.

- (56) tūtū **ñà** isyā'ā [ñá Mária] [nùù★ rā]
 book the.INA give.COMPL the.F Maria face 3SG.M
 'the book that Maria gave to him'

- (57) [tyà Juáàn] isyī'ī rā [tyìkuíi **ndrá** ìì★ vā'ā]
 the.SG.M Juan drink.COMPL 3SG.M water the.LIQ NEG good.IPFV
 'Juan drank the water that was not good.'

Full nominal heads in relative clause structures may themselves take on definite articles while a relative pronoun occurs at the same time. This shows that the two usages of noun classifiers as either definite articles and relative pronouns are grammatically distinct.

- (58) [ñà tūtū] ñà isyā'ā [ñá Mária] [nùù★ rā]
the.INA book the.INA give.COMPL the.F Maria face 3SG.M
'the book that Maria gave to him'

Lastly, relative clauses may even occur on proper names to serve as appositive relative clauses, as in the example below.

- (59) [tyà Juààn tyà kútóó k̄'vī] kuà'à và'ā ká'vī rā
the.SG.M Juan the.SG.M like.IPFV read.IRR much good read.IPFV 3SG.M
'Juan, who likes reading, reads a lot.'

4 Regular nominals and definiteness encoding

This section presents the semantic evidence for one of two claims. Cuevas Mixtec very much patterns with other languages displaying multiple strategies for encoding definiteness by associating those strategies with different notions of definiteness. For many nominal items of the language, definite bare nominals are interpreted as unique with respect to some domain or situation, while overt definite articles contribute an anaphoric element to the interpretation of the nominal. This is shown by observing patterns in the choice of definiteness marking strategy within the semantic environments of both immediate and larger situation uses, anaphoric uses, and bridging. Thus, Cuevas Mixtec displays the correspondence of bare form with uniqueness interpretation, and overt marking with familiarity interpretation, that has been noted for other languages by Schwarz (2013) and Jenks (2015). Most nominals of the language pattern this way, encompassing predicates without clear semantic associations among themselves, such as creatures and buildings. For this reason, it is assumed that these nominals represent a default in definiteness encoding, owing them the label of *regular* nominals.

4.1 Uniqueness with regular nominals

Regular nominals in their bare forms may be interpreted as uniqueness definites, and this is evidenced by the use of bare forms for various non-anaphoric purposes explained by Hawkins (1978). Bare forms are the natural form of regular nominals for the expression of larger situation definiteness, meaning that they are able to encode definiteness as characterized by reference to an entity uniquely identified within general world knowledge. The word *yòò* 'moon' refers to an entity uniquely identified as a moon in most real world interactions, and it displays resistance to modification by a noun classifier.

- (60) ✓ [tyà juáàn] ndé'é rā (#ñà) yòò
 the.SG.M Juan look.IPFV 3SG.M the.INA moon
 'Juan is looking at the face of the moon.'

Bare nouns are also used for immediate situation definiteness. They encode reference to an entity whose description is unique with respect to contextual knowledge that is shared between interlocutors. This is similar to larger situation definiteness in that uniqueness is anchored to a domain of shared knowledge, but it differs in that this domain is quite small, non-global, or situational. Any particular dog is not unique on a global scale, but dogs can be unique relative to their owners, as in the case of a family dog. Thus, the word *tyinā* 'dog' rejects modification by a noun classifier in the following example.

- (61) Context: A family's dog has gone missing for a week. A relative enters their house one day to find them cheerful and then proceeds to ask why they are suddenly happy.
indyikókōō (#tyí) tyinā
 return.COMPL the.AML dog
 'The dog came home!'

The results are fairly replicable for many examples of localized uniqueness. Churches are often unique to many villages in the Mixtec region of Mexico. The word *vēñù'ū* 'church' may not take a definite article assuming the context provided below.

- (62) Context: A man is visiting a Mixtec village, many of which have one church.
isyīnī ì (#ñà) vēñù'ū
 see.COMPL 1SG the.INA church
 'I saw the church.'

There is another usage not identified by Hawkins, though it is observed in more recent studies of definiteness. So-called *weak definites* (Carlson 2006) are actually neither unique nor anaphoric. They are nominals which appear to take on definiteness marking without referring to specific individuals. Below, the weak definite *the hospital* seems to refer more to the situation of being in a hospital rather than being in a particular one.

- (63) *Every accident victim was taken to the hospital.* (John to Mercy Hospital, Bill to Pennsylvania Hospital, and Sue to HUP) (Schwarz 2014: 3)

I have been able to identify at least one word, *yà'vī* 'market', which constitutes a case of a weak definite. In the example below, the word takes on a bare form, without definite articles.

(64) Context: Five people leave a room and return, having each gone shopping at a different market.

- ✓ [*ndyī'i vā nā*] *isyā'àn yà'vī*
all FOC 3.HUM go.COMPL market
'All of them went to the market.'

A final note concerning the encoding of uniqueness for regular nominals is that these nominals may not necessarily occur in bare forms for a uniqueness interpretation. Modified nominals may also have uniqueness interpretations at least in the case of partitive constructions. The example below demonstrates that a larger situation definite like *yòò* 'moon' may retain its uniqueness interpretation while modified by the quantifier *sāvā* 'half'. The resulting partitive construction is not interpreted as quantification over a group of moons, but quantification over portions of the unique moon with respect to Earth.

- (65) [*tyà Juààn*] *isyīnī rā* [*sāvā yòò*]
the.SG.M Juan see.IPFV 3SG.M half moon
'Juan saw half of the moon.'

This might suggest that uniqueness is interpretable within the complement of a quantifier. If that is true, it would entail that uniqueness interpretations of bare nominals syntactically correspond to an embeddable phrasal projection of some sort, such as a determiner phrase. However, the exact structural relationship between quantifiers and nominals in Cuevas Mixtec remains to be explained.

4.2 Familiarity with regular nominals

Besides bare forms of nominals, definite descriptions are also formed with overt marking by means of the language's definite articles, but the occurrence of definite articles comes with an alternative set of functions. Definite articles are somewhat awkward when occurring in semantic environments that suggest the uniqueness of the definite description's referent, as shown previously. Definite articles are much more preferred when used to indicate an anaphoric relationship with an antecedent nominal, corresponding to the interpretation of the def-

inite description as familiar.¹² This would follow Schwarz and Jenks's findings that languages which feature bare nominals as definite descriptions in addition to overt definiteness marking tend to reserve the overt marking for the expression of familiarity. In the narrative below, the first sentence presents a character, Juan, who is visiting a library to obtain a book that he is searching for. The two follow-up sentences are near identical in form, though they differ in their anaphoric properties due to the presence or absence of the definite article *ñà*. The first follow-up sentence has a bare nominal *líbrú* 'book' which is interpreted existentially under negation. This follow-up then claims that there are no books at all at the library. In contrast, the presence of the definite article in the second follow-up allows for continued comment on the book Juan was looking for, saying that it was absent from the library without comment on other books.

- (66) *isyā'àn* [*tyà* *Juáàn*] *bīblīōtéká táàn ná nī'ì rā* [*īn*
 go.COMPL the.SG.M Juan library so COMP obtain.IRR 3SG.M one
líbrú]
 book

'Juan went to the library in order that he get a book.'

- a. *sūū kòó líbrú índāà káā*
 but NEG book located.IPFV over.there
 'But there was no book there.'
- b. ✓ *sūū kòó [ñà líbrú] índāà káā*
 but NEG the.INA book located.IPFV over.there
 'But the book was not there.'

Because only the second follow-up is a continued comment on the book in the first sentence, it is the definite article which creates the crucial anaphoric link.

It is not crucial for the speaker to be familiar with the identity of the individual denoted by the definite nominal. The speaker may invoke a definite article for creating anaphoric links between coreferential nominals if their referent is learned about from hearsay. The narrative below introduces an unspecified turkey that

¹²It is worth noting here that there is a bit of variation in judgment across generational lines about the use of definite articles. The data here better reflects younger generational speech, which features broader usage of definite articles for creating anaphora. Older speakers seem to dislike definite articles on regular nominals, or are at least much pickier about when they are used. This might indicate a diachronic shift in the use of definite articles from something other than familiarity, which might also coincide with the development of definite articles in Cuevas Mixtec. To my knowledge, definite articles are rarely described for other Mixtec variants.

can only be referred back to with the occurrence of the definite article *tyí*. The two follow-up sentences are both declarations of hearsay that a turkey was sick, though only the second followup sentence is felicitous because of the presence of the definite article.

- (67) *isyā'ní* [*tyà juáàn kólō*
 kill.COMPL the.SG.M Juan male.turkey
 'Juan killed a turkey.'
- a. # *káchí nā ñà kú'vì vā kólō*
 say.IPFV 3.HUM COMP sick.IPFV FOC male.turkey
 'They say that a turkey was sick.'
- b. ✓ *káchí nā ñà kú'vì vā [tyí kólō]*
 say.IPFV 3.HUM COMP sick.IPFV FOC the.AML male.turkey
 'They say that the turkey was sick.'

Again, the first follow-up sentence is bizarre, but this time because it is interpreted as an assertion about a different turkey. The second follow-up sentence is interpreted as being about the same turkey, thanks to the presence of the definite article.

Definite articles may even occur on mass nouns, where their presence similarly encourages the formation of an anaphoric link between the definite description and a coreferential antecedent. The presence of the definite article allows a nominal to refer back to a particular collection of mass that was introduced before. The following example introduces a patch of salt that is later commented further upon as being brown.

- (68) *isyīkā* [*tyà Juáàn*] *nùù* nìì*
 walk.COMPL the.SG.M Juan on salt
 'Juan walked on salt.'
- a. *yā'ā vā [#(ñà) nìì]*
 brown FOC the.INA salt
 'The salt was brown.'

An interesting effect is observed with overt definiteness marking on mass nouns. The occurrence of the article encourages the interpretation of the nominal referent as being unitized in some manner, so as to distinguish a particular body of mass. In the following example with introductory and followup sentences, the occurrence of the definite article turns out to be optional, but with distinct

effects in the interpretation of the referent of *tyikuií* ‘water’. The lack of the article forces the interpretation of the nominal’s referent to be a greater collection of water that is salient within a situation and which may not altogether be a participant in the drinking event. The occurrence of the article encourages an interpretation of the nominal’s referent to be a delimited amount of water which is participating in the drinking event, such as from a bottle.

(69) [tyà Juáàn] isyĩĩ tyikuií
 the.SG.M Juan drink.COMPL water
 ‘Juan drank water (of unspecified source).’

a. ìì★ và’ā tyikuií
 NEG good.IPFV water

‘The water (from the river or lake) was not good.’

b. ìì★ và’ā [ndrá tyikuií]
 NEG good.IPFV the.LIQ water

‘The water (from a bottle) was not good.’

This seems to be in line with the findings on definiteness in Cuevas Mixtec for count nouns. The first followup sentence appears to represent a case of immediate situation definiteness, whereby the bare nominal indicates a unique individual relative to a situation. The bare nominal must then refer to the maximal amount of water given a situation, which is not identical to the amount of water that Juan drank. The article allows the nominal to refer back to the unitized water that Juan had drank and can be further commented on.

4.3 Bridging

The last usage of definite descriptions to be addressed in this paper are cases of bridging. For many examples of bridging, both bare nominals and nominals with definite articles may serve as anaphora for a non-coreferential antecedent, but there are also some cases which demonstrate a clear preference for one strategy of definiteness encoding over the other. It turns out that these special cases include those relationships between definite descriptions and their antecedents that were first outlined by Schwarz (2009), and Cuevas Mixtec patterns with other languages by invoking its strategies for marking uniqueness or definiteness for the same cases.¹³ In this language, when there is a part-whole relationship

¹³Schwarz has reported on encountering some variation among speakers’ intuitions regarding bridging examples, and I have found similar variation for these examples in Cuevas Mixtec.

between the definite description and the antecedent, the definiteness marking strategy of choice is the bare nominal, which indicates uniqueness. The following example has the definite description *tú yé'é* ‘the door’ in a part-whole relationship with the indefinite nominal *īn vē'é* ‘a house’. The occurrence of the definite article *tú* is dispreferred.

- (70) *isyīn* [tyà *Juáàn*] [*īn vē'é*]
 buy.COMPL the.SG.M Juan one house
 ‘Juan bought a house.’
- a. *syàà tá'vi vā* [(#*tú*) *yé'é*]
 already broken.IPFV FOC the.STR door
 ‘The door was broken.’

Also in line with observations from other languages, since the antecedent is not coreferential with the definite description, it need not even be an individual. The antecedent can also be a situation that is introduced and which may naturally entail conditions such as the existence of certain kinds of entities. The example below has a definite description *kárró* ‘car’ with an antecedent adverbial phrase *tá'ān sákākā* which introduces a situation of driving. Note that the inclusion of the definite article sounds awkward to speakers in this case.

- (71) Context: Juan has a strange hearing problem which causes him to go deaf or have selective hearing in special circumstances.
 [*tá'ān sákākā tyà Juáàn*] *ì kūvī tāsò'ō rā*
 every.time drive.IPFV the.SG.M Juan NEG can hear.IPFV 3SG.M
 [(#*tú*) *kárró*]
 the.STR car
 ‘Every time Juan drives, he cannot hear the car.’

The scenario presented in the adverbial phrase entails the existence of some vehicle to be driven within the event. The car is interpreted as being part of the driving event, which perhaps induces the choice of the bare form for the definite description.

Finally, when there is a producer-product relationship between the definite description and the antecedent, the preference of definiteness marking strategy is

However, speakers’ intuitions regarding bridging examples vary specifically in the strength of preference for one definiteness marking strategy over optionality between the two. They do not vary with respect to which strategy is preferred, and when intuitions are strongest, the findings in the data align with Schwarz’s and Jenks’ findings in other languages.

the one of overt marking. The example below presents a scenario of a purchased book which necessarily has an author. Authors are not in part-whole relationships with books but producer-product relationships with them, so the nominal *āūtòòr* ‘author’ preferably takes on a definite article for association between the nominal’s referent and an aforementioned book.

(72) [tyà juáàn] isyīn rā [īn tūtū]
 the.SG.M Juan buy.COMPL 3SG.M one book
 ‘Juan bought a book.’

a. [#(tyà) āūtòòr] kúú rā [tyà ñūū★ nùù★ yūkù]
 the.SG.M author be.IPFV 3SG.M the.SG.M village face mountain
 ‘The author was (one) from San Miguel Cuevas.’

As far as the data regarding regular nominals, definiteness marking strategy, and interpretation are concerned, there are few surprises, if any. The next section discusses cases of nominals which do stray from the patterns noted above, particularly by either overextending the usage of definite articles for expressing uniqueness or completely barring modification by definite articles.

5 Internal variation in definiteness marking

There are at least two other classes of nominal which do not display a pattern akin to that of the regular nominals described before. These other classes of nominal are small when compared to regular nominals which display the overt correspondence between encoding strategy and notion of definiteness. The *irregular* nominals require the presence of an overt definite article for both uniqueness and familiarity interpretations. The *complex* nominals do not occur with definite articles, perhaps because they seem to have one already morphologically built in, and so their bare forms serve for both uniqueness and familiarity interpretations. Table 3 summarizes the general correspondences between definiteness marking strategy and interpretation for all three classes.

Table 3: Presence of overt definite article according to usage

Noun type	Uniqueness	Familiarity
Regular	*	✓
Irregular	✓	✓
Complex	*	*

The differences displayed by these other classes of nominal are shown by comparison with regular nominals in their distinct grammatical behavior with respect to some of Hawkins's (1978) usages of definite articles. These nominals display differences in the obligatoriness of the absence or presence of definite articles while undergoing immediate situation uses, larger situation uses, and anaphoric uses. Therefore, they reflect distinct styles of encoding either uniqueness or familiarity. There are only a handful of examples of these classes of nominal that this paper is able to provide, and the exact size of each class is yet to be determined.

5.1 Irregular nominals and definiteness encoding

While regular nominals display a predictable pattern of associating distinct notions of definiteness with distinct definiteness marking strategies, the distinction is not recognized in the morpho-syntax of irregular nominals. Irregular nominals are so called because they differ from regular nominals in not exhibiting bare forms as definite descriptions, or rather, they do not feature bare forms as uniqueness definites as with regular nominals. While the strategy for the encoding of anaphoricity remains identical among these two classes of nominals, irregular nominals extend the use of definite articles to also encode uniqueness. For example, the word *yīvī* 'people' does not permit bare forms to serve as definite descriptions where other nominals can. The sentences below present a context where a man named Juan is visiting a village and is surprised by the disappearance of its inhabitants. In this case, *yīvī* cannot take a bare form and must take a definite article in order for the sentence to be acceptable.

(73) [*tyà Juáàn*] *isyā'àn rā* [*ñūū★ káā*]
 the.SG.M Juan go.COMPL 3SG.M village over.there
 'Juan went to the village over there.'

a. *sūū kòó nī indānī'ì rā* [**(nà) yīvī*]
 but NEG even find.COMPL 3SG.M the.HUM people
 'But he did not find the people.'

A different result is reached if the irregular nominal is switched out for a regular nominal such as *vēñū'ū* 'church'. In a typical Mixtec village, there is one church dedicated to the local patron saint, whom the village also tends to be named after. In this case, the nominal takes a bare form because there is no previous mention of a church to serve as an antecedent for the definite article.

- (74) [tyà Juáàn] isyā'àn rā [ñūū★ káā]
 the.SG.M Juan go.COMPL 3SG.M village over.there
 'Juan went to the village over there.'
- a. sūū kòó nī indānī'i rā [(#ñà) vēñù'ū]
 but NEG even find.COMPL 3SG.M the.INA church
 'But he did not find the/a church.'

The inventory of irregular nominals is not very large at all, and they all seem to have an interesting semantic similarity. Table 4 provides a list of the irregular nominals I have been able to recognize so far.

Table 4: Irregular nominals

Cuevas Mixtec	English
<i>yīvī</i>	<i>people</i>
<i>tyàā</i>	<i>man</i>
<i>ñā'à</i>	<i>woman</i>

Notice that each nominal is a human predicate, such that the selection of predicates represented among the irregular nominals seems to be indicative of an animacy hierarchy. It would appear that the most animate predicates, human predicates in this case, form a special class that exploits the definite article for further uses beyond what is typical within the language. The influence of animacy hierarchies in grammar has been well documented (Dahl & Fraurud 1996), and there are clear examples of its interaction with definiteness in languages as common as Spanish. For Cuevas Mixtec, there seems to be a particular relationship between animacy and uniqueness in particular, which has been grammaticized in a way that treats unique members of the highest rank in animacy as if they were familiar.

It is important to note that, despite the seeming obligatoriness of the definite article in the presence of irregular nominals, the definite article is only obligatory for the expression of definiteness. These same nominals may occur with some other types of determiners without the definite article, such as with certain quantifiers. Therefore, cases of definite articles on irregular nominals are not cases of prefixes.

- (75) ✓ [tā'an ūn tyāā tyà kúmí ūn búrrú] kání rā rí
 every one man the.SG.M have.IPFV one donkey hit.IPFV 3SG.M 3SG.AML
 'Every man that has a donkey hits it.'

There are even some environments where the irregular nominal may shed off any prenominal material. The only environment where I have noticed this is that of the preverbal position while the nominal also takes a relative clause. The example below shows that the same nominal has an optional definite article in the preverbal position, but an obligatory article in the postverbal position. Both the preverbal position and the relative clause seem to be important for optionality of the definite article, and it is mysterious why this should be the case.

- (76) a. [(tyà) tyāā tyà kútóó ká'vī] kuà'à và'ā ká'vī rā
 the.SG.M man the.SG.M like.IPFV read.IRR much good read.IPFV 3SG.M
 'The man who likes reading reads a lot.
 b. ikū kuà'à và'ā ikā'vī [* (tyà) tyāā tyà kútóó
 yesterday much good read.COMPL the.SG.M man the.SG.M like.IPFV
 ká'vī]
 read.IRR
 'Yesterday, the man who likes reading read a lot.'

Generally, however, irregular nominals must take on definite articles if they do not co-occur with numerals or when they occur in preverbal position. They must even take on definite articles if they are modified by quantifiers. Many nominals are able to occur in partitive constructions without any modifying material besides the quantifier. In such constructions, nominals actually tend to have generic interpretations, at least if the modified nominal occurs in preverbal position.

- (77) ✓ [sāvā tyìnā] ndé'í
 half dog cry.IPFV
 'Half of dogs bark.'
- (78) ✓ [sāvā tyikuí] isyī'í [tyà Juààn]
 half water drink.COMPL the.SG.M Juan
 'Juan drank half the water.'

Unlike regular nominals, irregular nominals are incapable of occurring without the definite article in the same environment.

- (79) [sāvā *(nà) tyàā] kúmí ndā'à
 half the.HUM man have.IPFV hand
 ('Half of men have hands.')

It may be worth noting other cases of definite articles occurring on uniqueness definites beyond the class of irregular nominals, so as to demonstrate the semantic complexity of interactions between definite articles and nominals more broadly. There are some very special cases of definite articles occurring on regular nominals in order to make their referents more precise. The example below presents a case where a definite article is used to specify a member of a pair of unique individuals, rather than encode familiarity. The word *mārtóòn* 'administrator' has only two possible referents in the village festival context, the male and female administrators. The occurrence of definite articles allows for precision as to which of them is being referred to. As a regular nominal, the word *mārtóòn* has the capacity to occur in a bare form as a uniqueness definite, but it also takes on the definite article only to make precise the gender of the referent.

- (80) *isyā'àn* *ì* *nùù* *ndútútú* [*nà* *ndyáá* *chūūn*]
 go.COMPL 1SG where meet.IPFV the.HUM mind.IPFV work
 'I went to where the festival organizers were meeting.'
- a. *tyā* *kòò* *mārtóòn* *nī* *ìsyōō*
 and NEG administrator even there.be.COMPL
 'But the administrators were not there.'
- b. *tyā* *kòò* [*tyà* *mārtóòn*] *nī* *ìsyōō*
 and NEG]the.SG.M administrator even there.be.COMPL.
 'But the male administrator was not there.'

The example shows that definite articles serve many purposes beyond the formation of definite descriptions in the language. Since they also encode gender, it seems possible for some regions of the grammar of Cuevas Mixtec to exploit this aspect of their meaning while ignoring their tendency to also encode familiarity. Altogether, the data present a picture of overt definiteness marking in this language that complicates the narrow pattern observed for regular nominals.

5.2 Complex nominals and definiteness encoding

The last category of nominals to be discussed here are what will be called complex nominals. Complex nominals differ from both previously mentioned classes of

nominal in that they are barred from taking on definite articles. This may be due to the fact that, as compounds, they already feature a sort of built-in noun classifier. The only case of a complex nominal that this paper discusses is that of *tyàxini* ‘mayor’, as in the example below, which shows the unacceptability of a definite article. The word is a compound of a noun classifier *tyà* and the word *xini* ‘head’, such that they are inseparable in order to retain the meaning of ‘mayor’. If the complex nominal is switched out for a regular nominal like *māéstró* ‘teacher’ in the same example, the option of attaching a definite article becomes available.

(81) Context: There is a competition between the mayors of each village, followed by a separate competition between the teachers.

a. [*(*tyà)* ***tyàxini*** *tyà* *isákāná’á*] *kuàñū’ù* *rā*
 the.SG.M mayor the.SG.M win.COMPL go.home.IPFV 3SG.M

‘The mayor that won went home.’

b. [*(tyà)* ***māéstró*** *tyà* *isákāná’á*] *kuàñū’ù* *rā*
 the.SG.M teacher the.SG.M win.COMPL go.home.IPFV 3SG.M

‘The teacher that won went home.’

As a nominal, the word may be modified with numerals and indefinite articles despite the noun classifier. Even with the constraint against the occurrence of definite articles, complex nominals may still occur as familiarity definites. In the second sentence below, the nominal *tyàxini* ‘mayor’ has the interpretation of referring to the same mayor that was previously mentioned.

(82) [*tyà* *juáàn*] *indātú’ún* *rā* *syí’in* [*ūn* ***tyàxini***]
 the.SG.M Juan chat.COMPL 3SG.M with one mayor

‘Juan chatted with a mayor...’

a. *tyā* ***tyàxini*** *ikūsì* *īnī* *rā*
 and mayor cheerful.COMPL inside 3SG.M

‘and the mayor was happy.’

The rejection of definite articles for these items may have an explanation in the occurrence of a derivationally built-in noun classifier *tyà*. Compounding with classifiers occurs quite commonly across Mixtec and other Otomanguan languages, though it is better understood as a diachronic phenomenon which has resulted in fossilized forms of classifiers (Macri 1983). Classifiers in compounds, or so-called *lexical classifiers*, are distinct from the grammatically active noun classifiers for all Mixtec languages. They constitute a much larger inventory with

meaning contributions that have been lost over time, and not all words of the language feature them. Lexical classifiers may co-occur with definite articles, substantiating the claim that they are a distinct class of fossilized forms.

- (83) *tyí tyì-xú'ù*
 the.AML CLF-money
 'the goat'

In addition, the noun classifier is actually interchangeable with other noun classifiers, in particular the plural human classifier *nà*. This allows the nominal to take on a plural number meaning in what seems to be the only case of nominal inflection in this language. Likewise, this item is still unable to co-occur with a definite article.

- (84) Context: There is a gathering of villages.
*[sāvā (*nà) nàxini] kú'vi nā*
 half the.HUM mayors sick.IPFV 3.HUM
 'Half the mayors were sick.'

Therefore, the classifier that occurs in the complex nominal is not quite the same as the lexical classifiers that have been more widely described for Mixtec languages.

Beyond etymological considerations, the rejection of definite articles could also be explained from a semantic point of view. Mayors are of course relational nouns, or designations dependent on an individual's relationship with something else. For someone or something to be a mayor, there must be a town for that individual to be a mayor of, perhaps automatically inducing a bridging environment with a part-whole relationship. Further investigation on other relational nouns would be necessary to substantiate this. It does seem to be the case that true relational nouns such as body parts also reject modification by definite articles. In contrast, body parts differ from *tyàxini* in that they seem to be averse to occurrences as bare nominals and require at least some other form of modification.

- (85) Context: A teacher is overseeing a boy make a drawing of a man. The teacher takes a look at the boy's progress, and notices that the head of the man is drawn disproportionately large.
*ká'nū ndyá'ā (*ñà) xini #(rā)*
 big very the.INA head 3SG.M
 'The head is too big.'

The complex nominal therefore presents another challenge to the development of a homogenous account of definiteness in Cuevas Mixtec. There seems to be an active grammatical role that the noun classifier plays in the construction of a relational noun such as ‘mayor’ while avoiding the typical usage of noun classifiers as encoding familiarity as definite articles. The data in this section also presented the case of obligatory definite articles on irregular nominals for cases where regular nominals would be bare, demonstrating the apparent influence of an animacy hierarchy on the distribution of definite articles. The existence of both classes of nominal complicates an account of definiteness encoding strategy as uniformly corresponding to the expression of either uniqueness or familiarity for Cuevas Mixtec.

6 Conclusion

This paper served as an presentation of the internal variation exhibited within Cuevas Mixtec with respect to strategies of definiteness marking, and what that variation may be the result of. The data support the findings of Schwarz (2009; 2013) and Jenks (2015) that languages which feature distinct strategies for definiteness marking will often associate those strategies with distinct notions of definiteness. One strategy will correspond to the expression of uniqueness, or the function of referring to an individual that uniquely fulfills the description provided by the noun. Another strategy will correspond to (strong) familiarity, or the function of creating an anaphor to previous linguistic expression in a discourse. Schwarz (2013) and Jenks (2015) found that in many languages which feature bare nominal definite descriptions in addition to overt definiteness marking, bare definite nominals will be interpreted as unique, while familiarity requires the overt marking. The pattern is replicated in Cuevas Mixtec, which has bare nominals serve as uniqueness definites in many contexts, and requires the occurrence of overt definite articles for the expression of familiarity. This was shown by observing the grammatical constraints on definite descriptions within different semantic environments listed by Hawkins (1978). Bare definites are preferred in cases of larger situation and immediate situation uses of definite descriptions, environments which reinforce the uniqueness of the definite description’s referent. Nominals with overt definite articles were preferred in cases where the definite description was used as an anaphor, corresponding to the familiarity characterization of definite descriptions in the literature. Even the case of bridging demonstrated the predicted correspondences between relationship type and preferred strategy of definiteness marking. Where the relationship between the

definite description and its antecedent was a part-whole relationship, the bare form seemed to be preferred. Where the relationship between the definite description and its antecedent was a producer-product relationship, modification by overt definite articles seemed to be preferred.

In contrast, the pattern explained above is only reserved for a large subset of the nominal inventory of Cuevas Mixtec. There are smaller classes of nominal which either lack the capacity to occur in bare forms for most contexts, or lack the capacity to take on definite articles. Those nominals that cannot shed the definite article were called *irregular nominals*, and they appear to retain the capacity to express uniqueness despite the presence of the article. Irregular nominals were shown to retain the definite article in immediate situation uses of definite descriptions, and unable to shed them even in the presence of other modifiers such as quantifiers. The definite article was shown not to be a prefix because there are environments where it may disappear, such as when a numeral occurs in its place. All of the irregular nominals seem to be predicates of humanity of some sort, meaning ‘man’, ‘woman’, or ‘people’. The data therefore suggest an interaction between overt definiteness marking, especially uniqueness marking, and an animacy hierarchy. Irregular nominals contrast with complex nominals, which seem to not take definite articles at all. Complex nominals included the relational noun ‘mayor’, which more frequently undergoes uses as a uniqueness definite. Examples of complex nominals are difficult to encounter, so a much more thorough study of this class is necessary to determine all the semantic properties involved in the inventory. Ultimately, the data show that if we are to assume an account of the semantics of definiteness along the lines of Schwarz and Jenks, there must also be some account for the nominal contribution in how definiteness marking preferences are determined.

Acknowledgements

I would like to thank Miguel “Una Isu” Villegas, Leoncio Vásquez-Santos, and Rufino Dominguez-Santos for their time and patience during elicitation and their excellent assistance and insight on aspects of Cuevas Mixtec grammar. I would also like to thank the following people for their guidance or recommendations concerning research on the topic of definiteness: Florian Schwarz, Anastasia Giannakidou, Itamar Francez. I would lastly like to thank those who have helped me put this paper together, by proofreading it or otherwise: George Borawski, the editors of this volume, and anonymous reviewers. Finally, kudos to Ana Aguilar-Guevara, Julia Pozas, and Violeta Vásquez-Rojas for putting together a terrific

conference and making this volume possible. This material is based on work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1144082. I am responsible for any errors that appear in this work.

Abbreviations

HUM	plural human	SPEC	specifier
INA	inanimate	STR	plant/structure
LIQ	liquid	YTH	youth

References

- Bradley, C. Henry & Barbara E. Hollenbach (eds.). 1988. *Studies in the syntax of Mixtecan languages*. Vol. 1 (Summer Institute of Linguistics Publications in Linguistics 83). Dallas: Summer Institute of Linguistics, University of Texas at Arlington.
- Carlson, Greg N. 2006. Generic reference. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 14–18. Amsterdam: Elsevier.
- Carroll, Lucien. 2015. *Ixpantepec Nieves Mixtec word prosody*. San Diego: University of California. (Doctoral dissertation).
- Clark, Herbert H. 1975. Bridging. In *TINLAP '75 Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing*, 169–174. Stroudsburg: Association for Computing Machinery.
- Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47–64. Amsterdam: John Benjamins.
- Ebert, Karen. 1971. *Referenz, Sprechsituation und die bestimmten Artikel in einem Nordfriesischen Dialekt (Fering)*. Kiel: Christian-Albrechts-Universität zu Kiel. (Doctoral dissertation).
- Egland, Steven T. 1978. *La inteligibilidad interdialectal en México: Resultados de algunos sonderos*. Mexico City: Instituto Lingüístico de Verano.
- Hawkins, John A. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croon Helm.
- INEGI (Instituto Nacional de Estadística y Geografía). 2010. *Censo general de población y vivienda*. Mexico City: INEGI.
- Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. *Proceedings of SALT 25*. 103–124.

- Josserand, J. Kathryn. 1983. *Mixtec dialect history*. New Orleans: Tulane University. (Doctoral dissertation).
- Macri, Martha J. 1983. Two noun class systems in Mixtec. In Alice Schlichter, Wallace L. Chafe & Leanne Hinton (eds.), *Studies in Mesoamerican linguistics* (Reports from the Survey of California and Other Indian Languages 4), 291–306. Berkeley: Department of Linguistics, University of California.
- Rensch, Calvin R. 1976. *Comparative Otomanguean phonology* (Language Science Monographs 14). Bloomington: Indiana University Publications.
- Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3). 287–350.
- Schwarz, Florian. 2009. *Two types of definites in natural language*. Amherst: University of Massachusetts. (Doctoral dissertation).
- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559.
- Schwarz, Florian. 2014. How weak and how definite are weak definites? In Ana Aguilar-Guevara, Bert Le Bruyn & Joost Zwarts (eds.), *Weak referentiality* (Linguistik Aktuell/Linguistics Today 219), 213–235. Amsterdam: John Benjamins.
- Spores, Ronald & Andrew K. Balkansky. 2013. *The Mixtecs of Oaxaca: Ancient times to the present* (The Civilization of the American Indian Series 237). Norman: University of Oklahoma Press.
- Velasco Ortiz, Laura. 2005. *Mixtec transnational identity*. Tucson: University of Arizona Press.
- Wespel, Johannes. 2008. *Descriptions and their domains: The patterns of definiteness marking in French-related creoles*. Stuttgart: University of Stuttgart. (Doctoral dissertation).

