

## THE DEFINITE ARTICLE IN YUCATEC MAYA: THE CASE OF *LE... O*<sup>1</sup>

Violeta Vázquez-Rojas Maldonado, Josefina García Fajardo,

Rodrigo Gutiérrez Bravo, and Julia Pozas Loyo

EL COLEGIO DE MÉXICO

### 1. INTRODUCTION

The aim of this paper is to provide a semantic characterization of a particular determiner in Yucatec Maya (ISO code: yua, henceforth Yucatec), a Mayan language spoken by approximately 824,670 speakers (INEGI 2010) in the Yucatan Peninsula, Mexico, and parts of Belize. In Yucatec, there is a paradigm of determiners formed by a prenominal element *le* plus a clitic added at the end of the noun phrase, as pointed out by Andrade (1955 [1940]), Hanks (1990), Briceño Chel (1996) and Lehmann (1998, 2012), *inter alia*.

---

<sup>1</sup> First and foremost, we would like to thank Rosa María Couoh Pool, Alfredo Hau Caamal, José Miguel Kanxoc Kumul, José Asunción Nahuat Canul, Irma Pomol Cahum, Gloria Nayeli Tun Tuz, and Dalia Tuz Ek for their judgements, translations and comments on the Yucatec data reported here, which are the core of research presented in this paper. We would also like to thank three anonymous reviewers, as well as Lisa Matthewson and David Beck for their suggestions, which greatly helped to improve the structure and content of this paper. Thanks also to Barbara Partee and Donka Farkas for their invaluable observations and for having directed us to useful references. We would also like to thank the audiences at the Conference on Indigenous Languages of Latin America (University of Texas at Austin, October 2013) and at the XII Congreso Nacional de Lingüística (Universidad Autónoma de Querétaro, October 2013) for their helpful feedback and suggestions. Lastly, many thanks also to Jorge Arroyo for his assistance with the final typographical revision of this paper. All remaining errors are our own. This research was partially funded by a grant from PRODEP (formerly PROMEP) for the project *Estudio sintáctico y semántico de los términos que generan la referencialidad nominal y sus inferencias: Un estudio translingüístico*, COLMEX-CA-47.

The resulting forms are shown in the following examples:<sup>2</sup>

- (1) a. **Le** kajtalil way=**a**'  
Le hamlet here=CL  
'**This** hamlet here' (Monforte *et al.* 2010:23-16)
- b. **Le** mejen iik'=**o**'  
le little air=CL  
'**Those** tiny spirits' (Monforte *et al.* 2010:49-5)
- c. **Le** janal=**e**'  
le food=CL  
'**That** food' (Briceño Chel 1992:20)

The examples in (1) show the three determiners formed with prenominal *le* plus a clitic: *le...a*', *le...e*' and *le...o*', which have been regarded as demonstratives in the relevant literature. However, a closer look at *le...o*', based on data from controlled elicitation and narrative texts, reveals that this determiner has also the referential functions expected of a

---

<sup>2</sup> The abbreviations used in the examples are the following; 1: first person, 2: second person, 3: third person, ABS: absolutive, AF: agent focus, CAUS: causative, CIT: reportative, CL: clitic, CLAS: classifier, CP: completive, CPS: compulsive, DIM: diminutive, DUR: durative, EP: epenthesis, ERG: ergative, EX: existential, FEM: feminine, FOC: focus (adverb), HAB: habitual, INCH: inchoative, IND: indicative, IRR: irrealis, LOC: locative, NEG: negation, PASS: passive, PL: plural, PREP: preposition, PRF: perfect, PRSV: presentative, RDP: reduplication, SG: singular, TOP: topic, TRM: terminative, TRNS: transitive, V: vowel. The glosses ERG and ABS correspond to what is known in traditional Mayan linguistics as the *A* and *B* pronominal series. It should be noted that ergativity in Yucatec is split on the basis of aspect, and hence not every instantiation of a pronominal element labeled as ERG and ABS necessarily bears an ergative or absolutive grammatical relation. All examples are presented according to the 1984 phonemic orthographic conventions of the *Academia de la Lengua Maya de Yucatán* (CDAM 1984) and so they do not necessarily reflect their phonetic form accurately. In this orthographical system, symbols have their expected values except for *ch*=[tʃ], *j*=[h], *x*=[ʃ], and '=[ʔ]. We do not provide a morpheme gloss for each of the elements that constitute the *le* + clitic complex. Since the purpose of our paper is precisely to find the specific semantic value for these elements, we will simply gloss *le* as *le* and for the terminal clitics we use the gloss "CL".

definite article. The case in favor of this claim has not been systematically argued in previous analyses of the determiner system of Yucatec. Given the growing interest in recent literature on the cross-linguistic expression of definiteness (Arkoh & Matthewson 2012; Schwarz 2013), it is of the utmost importance in the grammatical description of a language to determine the formal resources available in that language to make definite reference. The results of our investigation allow us to conclude that Yucatec has an explicit mechanism to signal definite reference and that it consists of the determiner *le...o* which, in its use as a marker of definiteness, can be characterized as a definite article.

We assume that a determiner is a syntactic category that is typically comprised of three sorts of elements: demonstratives, quantifiers, and articles. Determiners are functional elements in the noun phrase that delimit the reference of the noun that accompanies them. We adhere to the following general definition provided by Panagidiotis (2014):

Determiners are a nominal syntactic category distinct both from adjectives and nouns, despite the close affinity among them. They are commonly understood to comprise the word classes of article, demonstrative, and quantifier, as well as some possessives and some nominal agreement markers.

English demonstratives and articles are examples of determiners. A determiner that is specialized solely in marking the feature [+/- definite] is usually called a *definite/indefinite article* (Gillon 2015:176). A definite article is thus a kind of determiner that expresses solely definite reference –in other words, a determiner without any deictic or person component added. In the framework adopted in this paper, noun phrases headed by a definite article can be identified by their ability to find a referent within the Universe of Discourse (the set of all entities whose existence is assumed by the interlocutors of a linguistic exchange). Noun phrases with definite articles, crucially, are able to signal generic reference, unlike noun phrases headed by other determiners (such as demonstratives or possessives), but they must also be able to refer to particular individuals and in general to entities in any subdomain of the Universe of Discourse (UD)

(unlike indefinite articles, which may make generic reference, but cannot have an anaphoric interpretation). In what follows, we show that *le...o*’ has all of these properties, and thus that it should be considered a definite article, aside from the fact that a homophonous item can function as a demonstrative.

This article is organized as follows: in Section 2, we present a brief overview of nominal expressions in Yucatec and a review of the previous analyses of *le...o*’, *le...a*’, and *le...e*’. In Section 3, we provide a detailed definition of what we consider a definite article alongside the empirical criteria needed to identify it. Section 4 presents a description of the methodology used for elicitation and data collection. In Section 5, we present the results of our controlled tests as well as some spontaneous data and we provide the corresponding analyses organized in the four domains in which definite noun phrases can find their referents, according to the framework put forth in Section 3. In addition, in Section 5.5 we discuss some examples in which, as stated in previous analyses, *le...o*’ functions as a demonstrative. The section concludes with a brief discussion of the forms *le...a*’ (1a) and *le...e*’ (1c). This discussion is intended to show that, although these two determiners share a number of contexts of use with *le...o*’, they do not allow the noun phrase they introduce to find its referent in just any subdomain of the Universe of Discourse, and therefore they are not definite articles. Finally, in Section 6 we summarize the main conclusions of our research.

## 2. OVERVIEW OF THE PHENOMENON

In Yucatec, certain noun phrases can be introduced by a determiner *le* and wrapped by a clitic following the noun and its modifiers (if there are any). Examples of these constructions appearing in full sentences are presented below.

(2) **le** ko’olel=**o**’ t-u ts’-aj-∅ u ma’alob nook’.  
 le woman=CL CP-ERG:3 put-PRF-ABS:3SG ERG:3 good clothes  
 ‘and the woman put on her good clothes.’ (Góngora Pacheco 1990:19-17/18)

(3) **Le** x-ko’olel=**a**’, yáan u lu’u-s-ik-∅ teen  
 le FEM-woman=CL CPS ERG:3 leave-CAUS-PRF-ABS:3SG 1SG

in           meyaj!  
 ERG:1SG   work

‘This woman is going to take my job away from me!’

(Ceh Moo 2011:69-11/12)

(4) tumen   **le**   máak=**e**’   ma’   ya’ab   taak’in   t-u           p’aat-aj-∅           ti’.  
 because le   person=CL NEG   much   money   CP-ERG:3   leave-PRF-ABS:3SG PREP  
 ‘because the man didn’t leave him much money.’           (Ceh Moo 2011:45-4/5)

Most descriptions of these forms in Yucatec claim that the clitics are obligatory in the presence of *le*. However, it is not uncommon that in texts and spontaneous speech the clitics are omitted under certain circumstances. Andrade (1955 [1940]) suggests this is due to phonotactic reasons (=o’ is omitted after [o], and =a’ is omitted after [a]), but also to the fact that =o’ is sometimes omitted if the constituent introduced by *le* is especially long and complex. Hanks (1990:57) further observes that the clitics are commonly omitted when the lexical description, gestures, or available knowledge are enough to identify the referent introduced by *le*. There is a further morphosyntactic restriction (which we address in detail in §5.4 below): when two or more nominal expressions with *le* co-occur in the post-verbal field, only one of them can show its clitic. However, in the absence of these conditions the clitics are obligatory. While the bulk of our paper deals with the form *le...o’* and how we analyze it as a definite article, in what follows we will occasionally need to compare it with the two other forms, *le...a’* and *le...e’*. Alongside this comparison we present some facts about *le...a’* and *le...e’* that have not been considered (or have only been mentioned in passing) in the previous literature.

Since the constructions introduced by *le* characteristically correspond to noun phrases, we now provide a brief description of nominal expressions in Yucatec. There are a fair number of detailed descriptions of nominal expressions in this language, such as Briceño Chel (1993, 1996) and Lehmann (1998), and brief comments on the properties of Yucatec nominal expressions can be found in Andrade (1955 [1940]), Lucy (1992), Verhoeven (2007:105–107) and numerous other works. The description we provide

below is merely introductory and makes reference only to those properties of Yucatec nominal expressions that are related to the discussion of definite determiners that follows. For more detailed descriptions, we refer the reader to the works cited above.

In their most basic form, nominal expressions in Yucatec can be made up of a bare noun with no inflectional morphology. Bare nouns are allowed in argument position, with a collective interpretation as in (5) or with a mass interpretation as in (6):

(5) ...utia'al u chan man-ik-∅ **máak** ba'al u  
 for ERG:3 DIM buy-IND-ABS:3SG person thing ERG:3

jaan-t-∅ bey=o'.  
 eat-TRNS-ABS:3SG thus=CL

‘...for people to buy a little thing to eat like that.’ (Monforte *et al.* 2010:101-9)

(6) María-e' t-u man-aj-∅ **sakam.**  
 María-TOP CP-ERG:3 buy-PRF-ABS:3SG tortilla.dough  
 ‘María bought tortilla dough.’

Nominal inflection is fairly poor in Yucatec and possibly the most common inflectional category is plurality. Plural marking is not restricted to animate nouns, as in (7); it is also possible with inanimates, as shown in (8).

(7) a. ko'olel-o'ob  
 woman-PL  
 ‘women’ (Lehmann 1998:19)

b. ba'alche'-ob  
 animal-PL  
 ‘animals’ (Monforte *et al.* 2010:16-3)

(8) a. naj-o'ob

house-PL  
'houses' (Monforte et al. 2010:24-8)

b. ba'al-o'ob  
thing-PL  
'things' (Monforte et al. 2010:14-16)

Plural marking is optional (as can be seen in (9a) below),<sup>3</sup> although high animacy and definiteness strongly favor that a semantically plural noun display plural morphology (Lucy 1992; Lehmann 1998:19). Nominal expressions may include numerals and numeral classifiers (Lucy 1992; Briceño Chel 1993). There are only four Mayan cardinals currently in use in Yucatec: *jun* 'one', *ka'a* 'two', *óox* 'three', and *kan* 'four'. Numerals obligatorily appear with a numeral classifier or a mensurative expression, which are formally (i.e. morphosyntactically) identical in Yucatec. This is illustrated in (9):

(9) a. ka'a túul **peek'**  
two CLAS dog  
'two dogs' (Briceño Chel 1996:98)

b. jun kúul **che'**  
one CLAS wood  
'one tree' (Góngora Pacheco 1990:30-14)

c. jun p'ít **taak'in**  
one little.bit money

---

<sup>3</sup> A reviewer notes that, given that plural marking is optional, it might not be an inflectional category at all. In our view, inflectional categories are not necessarily characterized by being obligatory. The fact that the plural suffix *-(o)'ob* can attach to almost any member of the (count) nominal class, and that it does not involve a categorial or distributional change in the base, speaks in favor of its inflectional nature. In any case, the discussion is not germane to the main point of this paper, and if some later analysis shows that the plural suffix in Yucatec is not an inflectional category, this would not affect our central claims in any way.

‘a little bit of money’

(Monforte *et al.* 2010:308-4)

For numbers five and above, Spanish numerals are used. In these cases, the classifier appears as a possessed element, as can be seen in (10a), where the ergative pronoun preceding a nominal functions as a possessive marker. The classifier is also accompanied by a *-VI* suffix (where *V* copies the vowel of the stem), as in (10a). Example (10b) shows that classifiers are optional with Spanish numerals.

- (10) a. *trece u p'éeel-el waaj*  
thirteen ERG:3 CLAS-RDP tortilla  
‘thirteen tortillas’

(Monforte *et al.* 2010:138-5)

- b. *cinco k'áan*  
five hammock  
‘Five hammocks’

(Monforte *et al.* 2010:30-19)

Consider now the morphosyntactic properties of nominal expressions introduced by *le*. As first shown in examples (1-4), *le* requires the presence of one of three clitics, =*a*’, =*o*’, or =*e*’. These elements are phrasal clitics since they appear at the right edge of nominal expressions irrespective of how much material appears between the head noun and this edge. Consequently, noun phrases bracketed by *le* + clitic sequences can also include embedded verb phrases or clauses. The latter case can be seen in the following examples with relative clauses:

- (11) [*le x-ch'úupal [k-u wen-el=i']*]=*o*'...<sup>4</sup>  
*le* FEM-girl HAB-ERG:3 sleep-IND=LOC=CL  
‘The girl that sleeps there.’

(Góngora Pacheco 1990:30-18)

---

<sup>4</sup> As noted by two anonymous reviewers, example (11) is unusual because only one clitic is ever allowed to appear in any given clitic position in Yucatec (Hanks 1990: 491), yet in this case two different clitics appear on the right edge of this constituent: the clitic =*o*' related to *le*, and a free-standing locative clitic =*i*' semantically related to the embedded verb *wen-el* ‘sleep’. At present we do not have any



- (12) [ **le** máak [ ts'o'on-∅]=o' .  
 le person shoot:PASS-ABS:3SG=CL  
 'The person that was shot.' (Monforte et al. 2010:65-4)

When a clause is embedded in the nominal expression, as in the examples above, the right edge of the noun phrase is aligned with the right edge of the embedded clause, and it is in this position that =a', =o', or =e' are cliticized (Gutiérrez-Bravo 2012). In these examples, the clitic at the right edge of the relative clauses is =o', since the *le...o'* sequence is the primary object of our study: however, similar text examples can readily be found with the clitics =a' and =e'.

*Le* can co-occur with other pre-nominal constituents. In Yucatec, as in other Mayan languages, ergative pronouns also function as possessives, as in (10a). In (13) we show that *le* can co-occur with the ergative proclitics:

- (13) **le** u y-íicham=o'  
 le ERG:3 EP-husband=CL  
 'Her husband' (Monforte et al. 2010:27-14)

Example (14) shows that *le* can also co-occur simultaneously with both ergative proclitics functioning as possessive pronouns and with the numerals previously described.

- (14) **le** in ka'a túul w-íits'in-o'ob=a'  
 le ERG:1SG two CLAS EP-younger.sibling-PL=CL  
 'These two younger siblings of mine' (Briceño Chel 1996:101)

---

explanation for this fact, nor do we have any other examples displaying this behavior. However, since =i' is an anaphoric locative clitic that is not in a paradigmatic relation with the clitics that appear with *le* (see Hanks 1990: 545, fn. 18), analyzing its specific morphosyntactic properties is not directly relevant to the semantic analysis that we develop in what follows. As such, we leave this issue open for future research.

For ease of exposition in this paper we focus exclusively on constructions where *le* is the only prenominal constituent. Lastly, observe that the clitics =*a*', =*o*', and =*e*' are also obligatorily found with other determiners/adverbial elements, most notably the presentative *je*', the locative *te*', and the adverb *bey* 'thus' (which can be found with =*a*' or =*o*', but not with =*e*').<sup>5</sup> This is illustrated for *je*' and *te*' in the following examples from Briceño Chel (1992).

(15) a. Je' janal=a'.  
 PRSV food=CL  
 'Here's the food.'

b. Je' janal=o'.  
 PRSV food=CL  
 'There's the food.'

(16) a. te' kaaj=o'.  
 LOC town=CL  
 'in that town'

b. te' kaaj=e'.  
 LOC town=CL  
 'in that (distant) town'

These examples show that the terminal clitics are separable from the prenominal element *le*, and that they very likely contribute deictic content not only to determiners, but also to adverbs of location, adverbs of manner and presentatives. It makes sense, therefore, to consider that determiners that co-occur with these clitics are always deictic (hence,

---

<sup>5</sup> Observe, however, that the impossibility of combining *bey* with the clitic =*e*' only holds in those cases where there is no intervening material between *bey* and the terminal clitic (see Hanks 1990: 545, fn 17). For a complete description of the properties of =*a*', =*o*', and =*e*' with these other determiners/adverbs, see Hanks (1984, 1990, 2005).

demonstratives). However, as we will argue, this is not necessarily the case with the determiner *le...o'*. Before we begin our analysis of *le...o'* as a definite article in Yucatec, there is an important descriptive and analytical issue to address. While it is unanimously accepted that the clitics =*a'* and =*o'* are in a paradigmatic relation to one another, the status of the clitic =*e'* is considerably more controversial. The clitic =*e'* observed in the *le...e'* construction in (1c) and (4) is homophonous with (at least) two other clitics widely observed in Yucatec: the topic clitic in (17), and the clitic that functions as a clause linker (“continuator” in the terminology of Verhoeven 2007) in (18):<sup>6</sup>

(17) Pero leti'=e' k-u p'áat-al te' jool=o'.  
 but 3SG=TOP HAB-ERG:3 stay-IND LOC door=CL  
 'But her, she stayed at the door.' (Monforte et al. 2010:27-14)

(18) In k'áat-ø=e' káa k'áax-ak-ø ja'.  
 ERG:1SG wish-ABS:3SG=CL SUBJ fall-IRR-ABS:3SG water  
 'I wish that it would rain.' (Verhoeven 2007:126)

In (17), =*e'* attaches to a noun phrase that corresponds to the topic of the sentence. In (18), it cliticizes to a verbal element, rather than to a nominal one.

Since these three clitics (the topic marker, the clause linker and the =*e'* clitic used in determiners) are widely used in Yucatec, it is worth considering whether they are not one and the same clitic, instead of three different homophonous elements, as we claim.<sup>7</sup> There are indeed many cases where, for instance, distributionally it is not possible to distinguish the =*e'* clitic that accompanies *le* and the other two clitics. For instance, in (19) it is not straightforward to claim that =*e'* is the clitic that accompanies *le*, or if it is the same topic clitic as the one observed in (17). Recall that, when two clitics compete for a host, only

---

<sup>6</sup> The clause linker clitic optionally signals the right edge of the clause. Its precise behavior and function are still poorly understood, but the existing evidence indicates that its distribution appears to be conditioned by prosodic, and not syntactic or semantic requirements: see for instance Verhoeven & Skopeteas (2015).

<sup>7</sup> We thank two anonymous IJAL reviewers for bringing this issue to our attention.

one of them can actually occupy the slot, so in (19) it is not possible to know if there is only one clitic =e' corresponding to a determiner, or if there is a clitic =e' marking a topic, or both:

- (19) **Le** nojoch máasewáal máak=e' chéen t-u pak-t-aj-ø...  
 le big Indian person=CLonly CP-ERG:3 look-TRNS-PRF-ABS:3SG  
 'The large Indian man just stared at him...' (Ceh Moo 2011:58-20/21)

In spite of this distributional overlap, syntactically it is actually possible to distinguish between these three clitics. First, the clitic that functions as a clause linker is only observed in segments of discourse or complex constructions where more than one clause or sentence is involved. Second, sentence topics with the =e' clitic always occupy a clause-initial position, as expected of sentence topics in general. Now observe that in examples (20) and (21) neither of these two conditions is met. In these examples, there is no other clause following =e' (in fact, (21) is a monoclausal construction), so =e' cannot be functioning as a clause linker:

- (20) ...leti' túun=e' t-u beet-aj-ø u chéen leti' jun-puul=i'  
 3SG then=TOP CP-ERG:3 do-PRF-ABS:3SG ERG:3 only 3SG one-CLAS=CL  
 máax ts'iib-ø tuláakal ba'al t-u yo'olal **le** kaaj=e'.  
 who write.AF-ABS:3SG every thing PREP-ERG:3 about le town=CL  
 'And that made him the first and only one who wrote everything about the town.'  
 (Ceh Moo 2011:49-12/14)

- (21) T-u tukl-aj-ø bey-o' **le** xi'ipal-e'.  
 CP-ERG:3 think-PRF-ABS:3SG thus=CL le young.man=CL  
 'The young man thought about it thus.' {txt}

Perhaps more importantly, the nominal expressions with *le...e'* in these cases do not occupy the clause-initial position, which eliminates the possibility of =e' being the

sentence topic clitic. For our purposes, what is relevant is that there are both text and elicited examples where it is clear that  $=e'$  is neither a topic marker nor a clause linker. Hence, data like this justify our stance that the clitic  $=e'$  at the right edge of nominal expressions introduced by *le* can be taken to be a different clitic altogether.

Having clarified this important descriptive issue about the clitic sequences that will be the core of our analysis, we now have a general idea of the structure of noun phrases in Yucatec. Next, we provide a survey of the analyses and descriptions of the paradigm of determiners made up by *le* and a terminal clitic that can be found in the literature.

In the grammatical sketch by Bricker *et al.* (1998:382), *le* is described as a definite article, but this work does not mention anything about the relation between *le* and the clitics it appears with, and no description or analysis of nominal expressions introduced by *le* is provided either. The descriptions by Blair (1964) and Andrade (1955 [1940]) of *le* and its relation with the clitics are somewhat more detailed. Blair (1964:122–124) explicitly characterizes *le* as a demonstrative, and he describes the clitics that accompany it as deictic elements. Hence he claims that  $=o'$  “refers to something not so immediate in time or space. It often translates as *that* or *there*,” (Blair 1964:123), in contrast with  $=a'$  which he describes as referring to something near in time or space. The clitic  $=e'$  is actually the one for which Blair provides the most elaborate description, mentioning that it “may refer to something in mind, something already under discussion, or it may mark or set a *topic* to be commented on. Thus it often may be translated as *the* or *as for* though frequently it is not translated at all” (Blair 1964:123). Therefore, in Blair’s description the sequence that is taken to be the closest one in meaning with a definite article is *le... e'*, in contrast with the analysis that we develop in what follows. It should be pointed out, however, that the descriptions of the clitics provided by Blair are not specific to the meaning they convey when they are found in combination with *le*: the deictic contrasts described by this author are taken to be the same irrespective of the element that these clitics appear combined with. Accordingly, in this work there is no specific description of nominal expressions introduced by *le* and then combined with  $=o'$ . Similarly, Briceño Chel (1996:96) assigns to these elements a strictly demonstrative function and explicitly claims that Yucatec does not have a definite article.

The analysis by Andrade (1955 [1940]) shares some characteristics with the analyses of Blair (1964) and Briceño Chel (1996), but it is different in that Andrade does not assign to *le* and the clitics it appears with a uniquely demonstrative function. With respect to the clitics observed in these constructions, Andrade (1955 [1940]: §4.51) points out that:

[t]he terminal component *-a'* is used exclusively in ostensive references to items nearer to the speaker than to the listener, or to a time contemporary with that in which the reference is made; or in reference to what the speaker proceeds to demonstrate, exemplify, or speak of. *-o'* is used both in ostensive and in retrospective references. In ostensive references, the item is nearer to the listener than to the speaker, or roughly at about the same distance from both. For retrospective references, either *-o'* or *-e'* are used, depending mainly on topical distinction.

For Andrade, “ostensive reference” corresponds to the extralinguistic context, while “retrospective” reference points to elements in the discourse. With respect specifically to the combination of *le* with these clitics, Andrade labels them as “contextual definitives”, and suggests that *le...=a'* is used only for ostensive reference, whereas *le...=o'* is used for both ostensive and retrospective reference. Finally, regarding *le...=e'* in particular, Andrade points out that (1955 [1940]:§4.51):

[t]he form *le [...]-e'* is used in retrospective references and in vague references to what may be present or occur in a non-immediate future time. In retrospective references, *le [...]-e'* is not used when the identity of the item specified by [the nominal expression] is the dominant topic. Aside from such instances, it is difficult to disclose what governs the choice of *le [...]-e'* or *le [...]-o'* in retrospective references.

Importantly, Andrade’s distinction between “ostensive” and “retrospective” reference does not correspond to a distinction between definite and demonstrative determiners. A definite article is a determiner that exclusively marks definiteness (to be characterized

more precisely below) while having no deictic content, unlike demonstratives. Given that both demonstratives (which are deictic determiners) and definite articles can have ostensive reference and retrospective reference, this distinction is orthogonal to the categories of demonstrative/definite article. Furthermore, since Andrade's work is not explicit on this matter, there is no reason to assume that he considers *le...o'* to be a definite article on these grounds.

Much more elaborate analyses of *le* and the clitics that it appears with are developed by Hanks (1984, 1990, 2005) and Bohnemeyer (2012). In contrast with the works cited above, Hanks' proposal is that *le* and its clitics are not primarily used for ostensive/deictic reference. It is beyond the scope of this paper to provide a summary of Hanks' analysis that would minimally cover this author's findings with respect to the use of *le* and its clitics in specific spatial and social contexts. Let us just point out that Hanks describes *le...a'* and *le...o'* as nominal deictics, composed of *le*, which he takes to be by itself a definite article (Hanks 1990:18-19, 21), plus the terminal deictics =*a'* and =*o'*. It is not clear if the term "definite article" is used here to refer to a determiner without inherent deictic content, since Hanks describes *le* as an "initial deictic" (where "initial deictic" is the term used to group the initial component of deictic expressions that combine with "terminal deictics" such as =*a'* and =*o'*). Furthermore, it should be noted that the analysis of these forms is only a subpart of a larger analysis of deictic expressions in Yucatec in which the function of initial deictics (i.e., *le* for the case at hand) is to orient "the interactants to the character of a referent and its role in the predication", whereas the terminal deictics provide the "indexical grounding of the reference in the given portion of the interactive framework" (Hanks 1990:27). For our purposes, however, what is crucial is that this author specifically claims that *le* by itself corresponds to a definite article (Hanks 1984:155, Table 1, Hanks 1990:18-19, 21 and 163, Hanks 2005:198, Table 1). As will be discussed in detail in what follows, our proposal is different from this claim.

Lastly, there is the analysis by Bohnemeyer (2012), which focuses on the spatial properties conveyed by *le* + clitic sequences. Based on elicited and spontaneous data, he argues, *contra* Hanks (2005), that the location of the addressee does not determine the choice of the clitic. As such, the discussion of *le...o'* in this work centers around the spatial properties of its referents and how they are different from those of the *le...a'*

sequence. More to the point, in Bohnemeyer's analysis only *le...a* ' is "semantically specified for exophoric reference" (Bohnemeyer 2012:106), while *le...o* ' has a "more general indexical meaning", which includes exophoric uses but is "also used for anaphoric reference and definiteness marking" (2012:116). Although there is no explicit definition of "definiteness" in Bohnemeyer (2012), he seems to identify definite descriptions with those whose referent has been previously mentioned or is "assumed by the speaker to be uniquely identifiable to the addressee" (2012:106).

Summing up, although there are various and considerably different analyses of *le* and the clitics it appears with, there is yet no systematic analysis showing that the *le...o* ' sequence has the properties of a definite article: it is precisely this proposal that we will argue in favor of in this paper. In order to make our point clear, in the following section we present the theoretical framework that we use in our analysis of *le...o* ' as a definite article.

### **3. A CHARACTERIZATION OF THE DEFINITE ARTICLE**

Our analysis of Yucatec determiners is largely based on García Fajardo's (1985 [1984], 1989, 1990, 1994) approach, which considers that the meaning of a definite article consists of instructing the hearer to locate a set of entities that fit the description provided by the noun phrase within any of the components of the Universe of Discourse, a concept that we explain in detail below. We use the term "definite description" to refer to noun phrases that are explicitly headed by a definite article. The term "definite description" helps us set apart phrases with definite articles from noun phrases headed by demonstratives or possessives, which shall be named, respectively, "demonstrative descriptions" and "possessive descriptions". We set aside the problem of whether determinerless noun phrases in some languages can be considered "definite descriptions" simply because they make (some sort of) definite reference.

Historically, formal analyses of definite descriptions emphasized their uniqueness inferences (Russell 1905), their reference to a previously mentioned entity (Christophersen 1939, Heim 1983), and their inclusive character (Hawkins 1978). These notions (uniqueness, familiarity and inclusiveness) are central to the use of definite descriptions; nevertheless, the structure of the domain in which these properties arise still



needed to be defined. This leads to the definition of the “Universe of Discourse” (UD), as stated originally by García Fajardo (1985 [1984]), a crucial concept that serves as the basis for the framework adopted in this paper. The Universe of Discourse amounts to the set of all possible referents at a given communicative exchange, and these referents are organized into multiple sub-domains or “components”, which we will explain below. In this framework, the definite article is given a syncategorematic or “instructional” meaning. Its role is to direct the listener toward the set that has the properties described by the NP and that belongs to UD. When singular definite descriptions (such as *the cat* in English or *el gato* in Spanish) pick out one element (i.e., when they have a particular—non-generic—interpretation), the information ‘one entity’ combines with the meaning that we just assigned to the definite article: “set in the Universe of Discourse with the properties described by NP” (García Fajardo 1985 [1984]). The combination of these two pieces of information leads to the following inference: “the set in UD that fulfills the description provided by NP has only one element”, and this inference is the *uniqueness presupposition* that has been traditionally associated with the definite article. As such, uniqueness is inferred as a presupposition that arises from the meaning provided by the definite article when it combines with the conceptual information provided by the NP. If there is more than one element in some component of UD (e.g., if there are three entities in the speech situation that fulfill the property described by NP), then the singular noun phrase with the definite article picks out one element; however, the inference “the set in UD has only one element with the properties described by NP” makes the construction inappropriate for such a situation. Nevertheless, the phrase with the definite article may target another subdomain of UD that satisfies the presupposition.

The UD is a theoretical construct that represents a domain assumed to be shared by the interlocutors of a linguistic exchange, and it is constituted by all the entities to which definite descriptions may be targeted. The UD is organized into the following four components:

- *Discourse Interpretation Level*: the subdomain of entities that have been previously introduced in the discourse, and the relations among them.

- *Perceptual Representation of the Speech Situation*: the subdomain of entities which are physically present during the utterance act.
- *Representation of Previous Particular States*: the subdomain of entities that belong to the memories shared by the interlocutors.
- *Conceptual Information*: the subdomain of entities that belong to the extension of the conceptual information described by the noun phrase, considering only the noun and its modifiers (i.e. the Noun Phrase, as opposed to the whole Determiner Phrase).

Thus, in this approach uniqueness is generated as a presupposition which is logically inferred from the meaning proposed for the definite article plus the information about the number of referents and the descriptive content of the noun, without adding any further stipulation to the theory or any explicit content of unique reference to the article's meaning. Moreover, the semantic content proposed for the definite article allows us to explain why speakers are able to make generic reference (i.e. they can refer to the total extension of the set named by the NP) using definite descriptions.

As we will see, the generic use of definite descriptions is crucial for distinguishing them from other definite noun phrases, like demonstrative and possessive descriptions, since the last two cannot receive generic interpretations. What characterizes definite descriptions in our framework is the requisite of finding a referent within *any* of the four subdomains of UD: since Conceptual Information is a subdomain of UD, it is possible to obtain the generic reading directly, without any additional rearrangement of the formal system. Definite noun phrases that are not definite descriptions pose other restrictions (such as possession or distal deixis), which do not allow them to pick out all the extension of the noun phrase in any subcomponent of UD. In other words, demonstrative and possessive descriptions may find their referent in *some* of the subdomains of UD, but only phrases with a definite article may have *any of the four* subdomains of UD as their possible domain of reference.

To confirm that *le...o'* in Yucatec is indeed a definite article, and not simply a demonstrative, we must show that noun phrases headed by this determiner can refer to sets within any of the components of UD: the Discourse Representation Level, the

Perceptual Interpretation of the Speech Situation, the Representation of Previous Particular States and the Conceptual Information level. In other words, they must be able to make anaphoric reference, to refer to entities in the physical context surrounding the speech act, to refer to individuals that belong in the memories of the interlocutors, and to make generic reference. In contrast, demonstratives are not capable of making the last kind of reference, although they may overlap with definite articles in picking out entities in the first three subdomains.

#### **4. METHODOLOGY**

The data presented here were collected with questionnaires involving both acceptability judgment tasks and production tasks, although texts were used as a secondary source of data. Oral texts were taken mostly from Monforte *et al.* (2010) and Stolz *et al.* (2012), and some data were also collected from a novel (Ceh Moo 2011). The questionnaires were based on the framework presented in §3. The questions were designed to test whether speakers could locate the referent of noun phrases with each of the three determiners (*le...a'*, *le...e'* and *le...o'*) within each of the four domains of UD. Our purpose is to investigate whether one of these determiners corresponds to a definite article. According to the framework described in §3, if a determiner is a definite article, then its corresponding noun phrase should be felicitously used in *any* of the four following settings: (a) tracing an antecedent that has been previously mentioned (Discourse Representation Level); (b) locating a referent in the speech situation (Perceptual Interpretation of the Speech Situation); (c) recovering a referent from shared memories between interlocutors (Representation of Previous Particular States) and, (d) making generic reference (Conceptual Information). Additionally, we tested the use of each determiner in different deictic situations and in scenarios that either satisfied or did not satisfy uniqueness conditions. Importantly, this does not entail that other forms cannot refer to entities in some or other components.

Two questionnaires were used. The first questionnaire included 94 entries, and the second one included 26. Our consultants were native Yucatec speakers from eastern Yucatan, who were also fluent speakers of Spanish as a second language. The two questionnaires were elicited first with one consultant, and then they were replicated with

other consultants. All participants ranged in the age of 21-30. In the tags used to label the data obtained with these questionnaires, the initials correspond to the initials of each collaborator, the first number after the dash to the number of the questionnaire (1 or 2), and the second number and/or letter to the specific item in the questionnaire.

Following standard practice in semantic fieldwork (Matthewson 2004), the questions were designed to obtain grammaticality, truth-value, and felicity judgments. Each data point consists of a pair <sentence, context>. We collected these judgments basically by two strategies: in one of them, the elicitor presented a context to a speaker and then asked if a particular sentence in Yucatec (which had been constructed by the elicitor and previously checked for grammaticality) was acceptable and true in such a context. This is known as an acceptability judgment task. The other strategy was based on production tasks: a context was described, along with a communicative intention (e.g., “There is one table right in front of us, how would you ask me to clean it?”), and the speaker was asked to give a sentence in Yucatec that conveyed such communicative intention in the context provided (i.e. a free translation task). As stated by Matthewson (2004:386), the outcome of a production task should be considered a grammatical sequence in the object language (assuming, as we do, that the consultant is being cooperative). We further assume that the sentence provided must also be felicitous in the context presented and true (if it is a declarative sentence), otherwise it would have not been volunteered as an adequate expression for the context. This kind of task was particularly useful when eliciting anaphoric links and in prompting generic statements. These sentences were later modified (with the relevant variable being the choice of the determiner) and judged by the consultant compared to the same context in which the original sentence had been provided, resorting to the first strategy described above (i.e. acceptability judgment task).

The contexts for each question were stated verbally in Spanish to avoid priming in the object language. Note that the aim of our inquiry is to determine whether *le...o*’, *le...a*’ and *le...e*’ can be used in different contexts. Therefore, describing the context verbally in Yucatec would have forced us to use of the very determiner whose choice we wanted to test. In contrast, stating the context in Spanish avoided using *le...o*’ (or *le...a*’, or *le...e*’) before asking the relevant question.

When the contexts in the questionnaires were not stated only verbally, they included the physical setting up of “scenes”, which provided the context needed to judge the acceptability of a sentence or to refer to a particular object. When asked to provide a description of a certain situation, or to judge if a certain sentence was acceptable in a given context, the speaker was asked to do so as if he were talking to an addressee, which was often the elicitor—that is, the elicitor avoided asking questions like “How would I say” / “Could I say x”, and instead the speaker was asked “How would you describe x to someone?”, “How would you ask me to do y?”, or “Would you say z?” The speaker was instructed to refer to certain objects in different spatial configurations, or was asked to judge the acceptability of some statements referring to the situation. When the elicitor asked the speaker to translate a sentence from Spanish to Maya, a clear and complete context of the intended situation of use was provided. We did not rely on translations from Maya to Spanish, but we did consider speakers’ comments as potential clues for hypothesis design.

As for textual examples, we looked for occurrences of *le...o*’ in which it was clear that the noun phrase was either referring to an entity that had been previously mentioned (Discourse Interpretation Level) or to an entity that had not been mentioned explicitly but that was part of the shared memories of the interlocutors. Given the nature of the data, it was hard, but not impossible, to find instances of *le...o*’ that referred to entities which were present in the speech situation. Lastly, we also looked for uses of *le...o*’ that made reference to a whole class of objects (generics). When searching in texts, we only focused on *le...o*’, in order to confirm that the results that our questionnaires had yielded were supported by naturally occurring data.

## **5. ANALYSIS AND RESULTS**

Our results show that *le...o*’ can appear felicitously in any of the four conditions described, and hence, it corresponds to a definite article. These findings were also corroborated in texts. We present the relevant data below. In section 5.6, we show that the two other determiners were not accepted in every context in which definite articles are expected to be felicitous, according to the results of our elicitation tasks.

### *5.1 Discourse Interpretation level*

This level refers to the subdomain of UD which consists of all the entities that have been previously introduced in the current discourse, along with the relations among them. At this level, discourse interpretation is constructed step by step, utterance by utterance, with the help of grammatical and conceptual information. Some definite descriptions find their reference in the interpretation of a previous noun phrase, as in (22):

- (22) I remember our grandfather's study. There was **a desk** near the window, there were also a couch and some books in that room. **The desk** was always covered in piles of paper.

In (22), the noun “desk” is first used with an indefinite article at the beginning of the second sentence. In the last sentence, the definite description “the desk” refers back to the same entity referred to by “a desk”. The anaphoric relation in (22) is direct. However, the antecedent of a definite description may also be presented by a different noun that holds a conceptual relation with it, for example a hypernym. In (23) “the fruit” refers to the entity introduced by “apples and grapes”, and the association is possible because “fruit” is a hypernym of the other two nouns.

- (23) Last Saturday I went to the market and bought **some apples and grapes**. I also bought some cheese, meat, and wine. When I arrived home, I realized that **the fruit** was not in the bag.

One more way of establishing anaphoric relations is by associating the referent of a noun with some of its parts.

- (24) I remember our grandfather's **studio**. **The desk** was in front of **the window**.

In (24), “desk” and “window” are associated with the aforementioned “studio” in a meronymic relation: by world knowledge, we associate studios with these kinds of entities as their parts. It is the part-whole relation that licenses the anaphoric definite noun

phrase “the window” and “the desk”, since they refer to entities indirectly introduced by the noun “studio” in the Discourse Representation Level of UD.

Consider now the following elicited data from Yucatec, where the first mention of a piglet involves the cardinal *jun*, ‘one’, while the second mention of this same individual is obtained by means of a noun phrase with *le...o*’:

- (25) *Síib-∅*                      *teen jun túul chan k’ék’en yéetel jun túul t’eel,*  
 give:gift:PASS–ABS:3SG 1SG one CLAS small pig and one CLAS rooster
- ba’ale’ le chan k’éen=∅’, t-in kon-aj-∅.*  
 but le small pig=CL CP–ERG:1SG sell–PRF–ABS:3SG
- Intended reading: ‘I was given a little pig and a rooster, but the pig, I sold it’.
- (AN-2-26b)

The elicitation showed that noun phrases with *le...o*’, like *le chan k’éen=∅*’, ‘the little pig’ in (25), can establish direct anaphoric relations with an antecedent, which is evidence that they can find a referent in what was introduced above as the Discourse Representation Level of UD.

Data extracted from narrative texts corroborate this finding. Example (26a) introduces a new entity (‘a man’) at the beginning of a story. Reference to this same individual is then recovered several sentences later in the text by *le maak=o*’ ‘the man’, in (26b).<sup>8</sup>

- (26) a. *Yaan-chaj-∅ jun túul máak=e’ yaan-∅ u kool*  
 EX–INCH:CP–ABS:3SG one CLAS person=CL EX–ABS:3SG ERG:3 cornfield  
 ‘There was once a man who owned a cornfield...’ (Stolz *et al.* 2012:38)
- b. *K-u y-a’al-ik-∅ le maak=∅’*  
 HAB–ERG:3 EP–say–IND–ABS:3SG le person=CL

---

<sup>8</sup> The orthography in the examples taken from Stolz *et al.* (2012) has been modified according to the conventions we use in this paper. All the glosses and free translations are our own.

‘The man said:’

(Stolz *et al.* 2012:39)

A similar situation is presented in (27), where the antecedent noun phrase in (27a) is *jun p’éel kaan* ‘a snake’, and the anaphoric noun phrase is *le káan=o’* in (27b).

(27) a. K-u                    k’uch-ul=e’    k-u                    y-il-ik-ø=e’                    jun  
          HAB-ERG:3    arrive-IND=CL    HAB-ERG:3    EP-see-IND-ABS:3SG=CL    one

p’éel kaan

CLAS snake

‘When he arrived he saw it was a snake.’

b. Pero le máak=o’    káa t-u                    ch’a’-aj-ø                    óotsilil    ti’  
          but le person=CL    then CP-ERG:3    take-PRF-ABS:3SG    pity                    PREP

le kaan=o’.

le snake=CL

‘But then the man felt pity for the snake.’

(Stolz *et al.* 2012:40)

From these examples we conclude that noun phrases with *le...o’* are capable of referring directly to an entity that was introduced explicitly in a previous passage of discourse.

Discourse-anaphoric reference with *le...o’* is not limited to particular entities introduced by noun phrases; it is also possible with abstract entities, like time intervals, which are accessed indirectly by the description of events, as in (28):

(28) **Le** tiempo k-in                    w-a’al-ik-ø                    teech=o’    temprano  
          le time    HAB-ERG:1SG    EP-say-IND-ABS:3SG 2SG=CL    early

k            líik’-il            meyaj

ERG:1PL rise-IND    work

‘At the time I’m telling about we would rise to go work early.’



In (28), the antecedent of *le tiempo kin wa'alik teecho*, 'the time I'm telling you about', is not the referent of a previous single noun phrase, but a whole time span that comprises the events that have been narrated in the preceding stretch of discourse (not included here for reasons of space).

Lastly, in (29), we witness a particular type of anaphora in which the referent of the noun phrase with *le...o*' is not exactly the previously introduced plural entity. The following examples are taken from a narrative where the speaker has been talking about an old couple.

(29) a. In            w-íicham    teen=e'    k-u            bin=e', k-u            bin  
                  ERG:1SG   EP-husband 1SG=TOP   HAB-ERG:3   go=CL   HAB-ERG:3   go

ch'a'-aj-ø            si'            ti'            le    nukuch    máak-o'ob=o'.  
grab-PRF-ABS:3SG firewood    PREP    DM    great    person-PLUR=CL

'My husband, he used to go, he would go fetch firewood for the old folks.'

(Monforte *et al.* 2010:27-9, 10)

b. **Le** nojoch    máak=o'    leti'=e'    k-u            sáastal    k-u            máan  
                  le    big            person=CL 3SG=TOP    HAB-ERG:3 to.dawn    HAB-ERG:3    pass

día    ti'            yaan-ø            te'    jool    le    corral=o'.  
day    PREP            EX-ABS:3SG    LOC entrance le    corral=CL

'The old lady, from dusk until the (whole) day had passed, she, she was there at the corral's gate.'

(Monforte *et al.* 2010:27-11, 12)

The noun phrase with *le...o'* in (29b) does not refer to the totality of people mentioned (*le nukuch máako'obo'*, 'the old people' in (29a)), but instead only recovers a part of it: *le nukuch máako'* 'the old lady'.<sup>9</sup>

### 5.2. Perceptual Representation of the Speech Situation

This component of UD corresponds to the subdomain of entities which are physically present during the speech act (i.e. entities we perceive in the scene in which the speech occurs). For instance, we may express (30), when sitting at the table for dinner; or (31), at the beach, looking at a boy who plays with sand at a distance of twelve feet; or (32), while watching a ship on the horizon:

(30) Pass **the salt**, please.

(31) **The boy** is building a beautiful sand castle.

(32) **The ship** has a little flag; do you see it?

In these examples, definite descriptions are felicitously used regardless of the distance between the interlocutors and the referent. All that matters is that there is no other entity fitting the same description, and all that is required is that the referent is present and visible to the interlocutors in the scene where the enunciation occurs.

As shown in the following examples, *le...o'* can be used to refer to an entity in the setting of the utterance, at any distance from—but perceivable by—the speaker.<sup>10</sup>

(33) Context: A gourd is placed on the table. The speaker is asked how he/she would request the elicitor to pass him/her the gourd.

Ts'a-∅            teen    **le** luuch=**o'**.

---

<sup>9</sup> The feminine prefix *x-* is absent from this second nominal expression (since it is optional in Yucatec), but from the sentences that follow in the original text it is clear that reference is being made in this case to the wife, and not the husband.

<sup>10</sup> If there are two possible referents at different distances from the speaker, *le...o'* is used to refer to the farther entity: we discuss these cases in §5.5.

give–ABS:3SG 1SG le gourd=CL

‘Give me the gourd.’

(NT-1-3ai)

- (34) Context: A table is near both the speaker and the interviewer. The speaker is asked how the interviewer should ask him to touch the table:

Mach– $\emptyset$  le mayakche’=o’.

touch–ABS:3SG le table=CL

‘Touch the table.’

(DT-1-pr3-b)

- (35) Context: Same as in (34). The speaker is asked how the interviewer should ask him to clean the table:

Cho’– $\emptyset$  le mayakche’=o’.

clean–ABS:3SG le table =CL

‘Clean the table.’

(RC-1-pr3-c)

In (36), a stone had been placed near both the speaker and the interviewer. The speaker is asked to judge the acceptability of the sentence in this context. As in the other examples, no previous mention of this stone had taken place before.

- (36) Context: A stone is placed near both the speaker and the interviewer.

Ts’a– $\emptyset$  teen le tunich=o’.

give–ABS:3SG 1SG le stone=CL

‘Give me the stone.’

(DT-2-1b)

In example (37), the speakers were asked to judge the acceptability of (b) as a response to (a):

- (37) a. Máax lox–ech?  
 who punch–ABS:2SG  
 ‘Who hit you?’
- b. (Leti’) le xi’paal=ŋ’  
 3SG le male.kid=CL  
 ‘(He), the boy’ (RC-1-88b)

Examples (36) and (37) were obtained via acceptability tasks. In the contexts in which the corresponding judgments were elicited, it was also possible to refer to the stone in (36) with an indefinite article, which is what one consultant volunteered instead of just replying that (36) was acceptable; crucially, he did not overtly reject it, so this piece of data does not falsify our claim that the use of *le...o*’ is acceptable in this context. The remaining four speakers consulted agreed that the use of *le...o*’ was acceptable. Of these four, one provided a sentence with *le...o*’ in which he additionally included the adjective ‘small’, therefore producing a new sentence with *le...o*’ instead of just judging the one that was presented to him. With respect to (37), all speakers consulted accepted the form with *le... o*’ as a possible answer to the question ‘who hit you?’ but the results varied as to whether they needed the presence of the pronoun *leti*’ before the noun phrase in order to give a complete answer. Three speakers accepted it as it was elicited (without *leti*’), whereas the other two mentioned that the presence of *leti*’ was required in this context. The relevant point here is that; (i) the form with *le...o*’ was accepted in these contexts, irrespective of the fact that other possible expressions were also volunteered, and; (ii) even though the speakers consulted did not provide exactly the same answer for these particular elicitation tasks, none of the different answers they provided contradicted the judgements of the other speakers, and so they do not constitute evidence against our observation that *le...o*’ can be used felicitously in this context, as expected of a definite article. Lastly, example (38) is taken from an interview (which we regard as a kind of conversation):

- (38) **Le** k’áax=ŋ’ t–a w–il–ik=ŋ’ le nuuk–tak–o’ob

le jungle=CL DUR-ERG:2 EP-see-IND=CL le big-PL-PL

ka'anal-tak-o'ob ya'ab u y-áant-o'on=o'.

tall-PL-PL much ERG:3 EP-help-ABS:1PL=CL

'The jungle that you see, the one which is huge and tall that helps us a lot.'

(Monforte *et al.* 2010:16-19, 20)

This example confirms, in a naturally occurring setting, what is observed in (33-37). In spontaneous data it is difficult to control or even determine whether a certain referred entity is present or not in the speech situation, but in (38) the context itself makes it clear that the jungle ('the jungle that you see') is present, in full view of the interlocutors (additionally, one of the authors of this paper was present during the interview). It should be mentioned, though, that *le k'áax=o'*, 'the jungle' is used in several previous passages during the interview (all of them with generic reference), and it also has previous occurrences with the proximal demonstrative *le...a'*, 'this jungle', with a particular reference. However, (38) makes it explicit (by means of the modifying relative clause "that you see") that the referent is directly accessible and present at the moment of speech. It is thus plausible to consider that this noun phrase does not have an anaphoric interpretation even though it might be accidentally co-referential with a demonstrative noun phrase that has been previously used.

### 5.3 Representation of Previous Particular States

We now address the third subdomain of the Universe of Discourse, Representation of Previous Particular States, which refers to the set of entities that belong in the memories shared by the speaker and addressee. Note that these entities are particular individuals and not general concepts. Consider the following two examples:

(39) Last week, Fred told Ivan that he was reading a book by Sacks. Today they met at the library and Ivan asked Fred:

"Did you finish **the book**?"

(40) Fred knows that Ivan has been working for some time on a paper. When they meet on campus, Fred asks Ivan:

“How is **the paper** going?”

The referents for *the book* in (39) and of *the paper* in (40) are not traced anaphorically, because there have not been any previous mentions of these entities in the current discourse. However, they can be, easily identifiable within the memories that the speaker assumes that he and his interlocutor have in common.

Let us turn now to the use of *le...o*’ in such situations in Yucatec. In order to set the context to one in which the entities referred to were present in the shared memories of the interlocutors, the elicitor asked the speaker to translate a sentence from Spanish to Maya describing the exact context shown in (41) and (42). We were thus able to confirm that *le...o*’ can be used to introduce entities that belong to previous particular states:

(41) Context: Pedro has sent his son to fix the fence of their house. Later, they meet in the town square. His son tells Pedro that he has already fixed the fence.

Sáam in koot-∅ le koot=o’.  
 TRM ERG:1SG build.fence-ABS:3SG le fence=CL  
 ‘I have already fixed the fence.’ (RC-1-53)

(42) Context: Pedro has told you that the last time he saw your sister, she was weaving a hammock. Some other day, you meet with Pedro. How would he ask you in Maya:

“Has your sister finished the hammock?”

A kiik=e’ ts’o’ok u ts’o’ok-s-ik-∅ le k’áan=o’?  
 ERG:2 big.sister=TOP TRM ERG:3 finish-CAUS-IND-ABS:3SG le hammock=CL  
 ‘Has your big sister finished the hammock?’ (RC-1-56)

The context description ensures that the referents of *le koot=o* ‘le fence’ in (41) and *le k’áan=o* ‘le hammock’ in (42) are traceable to the shared memories of speaker and addressee.

Similarly, using *le...o*’ to introduce referents that belong to previous particular states is common in narrative texts. In (43) the noun phrase containing ‘jungle’ refers to a particular stretch of jungle near the speaker’s hometown:

- (43) Ma’ bin kíin-s-ik-∅ le nukuch k’áax=o’, nuk-tak-o’ob=i’!  
 NEG go die-CAUS-IND-ABS:3SG le large jungle=CL big-PL-PL=CL  
 ‘Don’t go kill the high jungle(s), the ones that are big!’ (Monforte *et al.* 2010:15-25)

Thus, in (43), the referent of *le nukuch k’áax=o*’ ‘le large jungle’ is a particular entity whose existence is assumed by speaker and addressee. The same is true in the case of ‘the cenote’ in (44), where the noun phrase with *le...o*’ refers to a particular cenote from which the inhabitants of the town used to get water:

- (44) le túun le ts’ono’ot=o’ ti’ bin k-u y-ok-ol u  
 le then le cenote=CL PREP CIT HAB-ERG:3 EP-enter-IND ERG:3  
 ch’a-∅-’ob ja’ u y-uk’-∅-o’ob=i’.  
 grab-ABS:3SG-PL water ERG:3 EP-drink-ABS:3SG-PL=LOC  
 ‘And so, the cenote, that’s where they used to go to get water to drink.’  
 (Monforte *et al.* 2010:24-13)

This cenote is not present at the moment of speech, nor has it been mentioned explicitly in the discourse. Rather, its referent is found in the shared memories of the interlocutors. Finally, in (45), ‘the festivity of Tahdziu’ refers specifically to the yearly festivity for the patron saint of Tahdziu, a neighboring town.

- (45) je’el=o’ táan le fiesta Ts’iu=o’ to’on=e’ bin-o’on  
 PRSV-CL during le festivities Tahdziu=CL 1PL=TOP go-ABS:1PL

úuch–ik–ø                      u            festejar–t–ik–ø  
 happen–FOC–ABS:3SG    ERG:3    celebrate–TRNS–IND–ABS:3SG

úuch–ik–ø                      u            bautisar–t–a’-l                      in            chan    hijo...  
 happen–FOC– ABS:3SG ERG:3    baptize–TRNS–PASS–IND    ERG:1SG DIM    son

‘And so, during the Tahdziu festivities, it so happened that we went to celebrate, as it happens, to have my little son baptized...’                      (Monforte *et al.* 2010:26-1, 2)

In naturally occurring examples it is not possible to state explicitly that the referents of these noun phrases belong to shared memories (unlike what happens with controlled elicitation). However, that the referents of the noun phrases with *le...o*’ belong to previous particular states is inferable from the context to which sentences (43)-(45) belong, since these entities belong to the interlocutors’ town, since there are no previous mentions of these entities in the text, and because the episodic nature of the predicates rules out a generic interpretation (which we discuss in section §5.4).

Summing up, in this section we have shown how the determiner *le...o*’ is able to introduce noun phrases which refer to specific entities that belong to the memories shared by the interlocutors. It is important to say that, in contrast with the previous two subsets of the Universe of Discourse–Discourse Interpretation Level and Perceptual Representation of the Speech Situation– the noun phrase with *le...o*’ does not need to be anaphoric, and the entity to which it refers does not need to be present in the speech situation. Referring to particular entities in shared memories is therefore a different domain in which noun phrases with *le...o*’ are perfectly acceptable and interpretable. We thus confirm that *le...o*’ is able to locate referents in three out of the four subdomains of UD that we identify in our framework for definite reference.

#### 5.4. Conceptual Information

Conceptual Information refers to the subdomain of entities that belong to the semantic extension of the noun (plus its restrictive modifiers if it happens to have any), independently from its particular context of occurrence. The descriptive content of a noun



is a property that holds of a set of individuals, thereby delimiting a class (e.g. the property “velociraptor” holds of a set of individuals which jointly constitute to the class of velociraptors). It is this descriptive content of the noun and the set of entities thus delimited that we refer to as “Conceptual Information”. When a definite article combines with a noun, it is possible for the resulting noun phrase to refer to the whole extension originally denoted by the noun (and its modifiers), therefore achieving generic reference.

The link between definite articles and generic reference has been long recognized in the literature (Christophersen 1939; Hawkins 1978), and is illustrated in the following examples from Spanish (46a) and in English (46b):

- (46) a. *El perro es el mejor amigo del hombre.*  
‘Dogs are man’s best friend.’  
b. *The red gazelle is now extinct.*

Importantly, one of the key features that distinguish the definite article from other definite determiners (such as demonstratives) is that only the former can be used to refer to whole kinds (“toto-generic” reference), as in “The velociraptor was a carnivore”. Demonstratives, too, can have generic uses, but when this happens, they only can achieve a species, or “parti-generic” reading, as in “This velociraptor was widespread in Mongolia, while this other one was more common in South Asia” (see García Fajardo 1985 [1984], Laca 1990, and Krifka *et al.* 1995). We thus consider the ability to refer to any of the components of the Universe of Discourse, including the possibility of making toto-generic reference, as the signature property that allows us to identify definite articles.

We identify a generic reference by the non-episodic nature of the predicate involved. Characterizing predicates (“The hazel dormouse hibernates for six months”) and kind-level predicates (“The popsicle was invented in California”) are good indicators that the relevant reference is generic (Krifka *et al.* 1995), and this is the criterion we rely on to identify generic references in Yucatec.

In Yucatec, *le...o’* is commonly used to make generic reference, as demonstrated in (47) where *le juuj=o’* refers to the whole kind of iguanas, as opposed to a particular entity.

- (47) **Le juuj=o'** jun túul ba'alche'.  
 le iguana=CL one CLAS animal  
 'Iguanas are animals.' (RC-1-62)

We know that *le juuj=o'* 'le iguana' refers to iguanas in general because the predicate intends to express a characteristic property of these animals, and not just of one particular entity. In the same vein, in (48) being male is taken to be a characteristic property of the kind 'priest'.<sup>11</sup>

- (48) **Le padre=o'** mantats' xiib.  
 le priest=CL constantly male  
 'The priest is always a male.' (RC-1-63)

The general character of the statement in (48) is reinforced by the presence of the adverb *mantats'* 'constantly' (*i.e.* 'always'), therefore supporting the observation that *le padre=o'* refers to priests in general, and not to a single priest.

The generic use of *le...o'* was previously noted by Hanks (1990:22), according to whom '[D]eictics ending with *o'* usually identify unique objects but may be generic (*le' ok'ot o'* 'dancing, that dance')...' <sup>12</sup> Generic *le...o'* is also documented in narrative texts. Consider, for instance, examples (49) and (50). In these examples, the noun phrases *le saandia=o'* 'le watermelon' and *le iik=o'* 'le chili', respectively, have occurred in

---

<sup>11</sup> Examples (47) and (48) were obtained as the first response to a production task, where a Spanish sentence was provided to set up a context like "Suppose that you don't know what iguanas are, and so I tell you 'Iguanas are animals'", thus ensuring a generic interpretation. These first outcomes were then modified into sentences with the other two determiners, *le...a'* and *le...e'*, to elicit acceptability judgments. The results of these tests are described in §5.6.

<sup>12</sup> Hanks' data do not provide a context to guarantee that this interpretation is indeed generic, and the translations offered suggest reference to particular events or entities rather than genericity –mostly because they are translated with the demonstrative "that" – but we were able to confirm through controlled elicitation that his original insight was correct.

previous passages, but the habitual interpretation of the predicates in these sentences is consistent with a generic interpretation of the relevant noun phrases with *le...o'*:

(49) Le sandia túun xan=o', bix k pak'-ik-∅  
 le watermelon then also=CL how ERG:1PL plant-IND-ABS:3SG

**le saandia=o'?**

le watermelon=CL

‘So watermelons then, how do we plant watermelons?’ (Monforte *et al.* 2010:14-24)

(50) Lekan k-u y-a'al-ik-∅ u y-óo-t-ik-∅  
 when HAB-ERG:3 EP-say-IND-ABS:3SG ERG.3 EP-want-TRNS-IND-ABS:3SG

u pak'-o'ob=e', to'on=e' ts'-u káaj-al  
 ERG.3 sow-PL=TOP 1:PL=TOP TRM-ERG:3 begin-IND

cosechar-t-ik-∅ **le iik=o'.**

harvest-TRNS-IND-ABS:3SG le chili=CL

‘As soon as they want to sow, we, we will already have started to harvest the chili peppers.’ (Monforte *et al.* 2010:13-15, 16)

The generic interpretation of *le saandia=o'* ‘*le watermelon*’ in (50) and *le iik=o'* ‘*le chili*’ in (51) stems from the presence of a habitual aspect marker together with an indicative suffix, along with world knowledge (e.g. the fact that a particular watermelon cannot be planted more than once, or that particular chili peppers are not harvested more than once, which is what the habitual meaning of the sentence would entail if the referents were particular individuals).

In example (51), *le tóolok=o'* ‘*le basilisk lizard*’ is mentioned for the first time in the text and has a generic interpretation, as is reflected in the habitual interpretation of the predicate.

(51) K-u ka'a taal yaanal, je'ebix tak le juuj, je'ebix tak  
 HAB-ERG:3 two come different like even le iguana like even

le tóolok=o', k-u bi-s-ik-ø,  
 le basilisk.lizard=CL HAB-ERG:3 go-CAUS-IND-ABS:3SG

k-u jaan-t-ik-ø.  
 HAB-ERG:3 eat-TRNS-IND-ABS:3SG

'Something else comes again, (even) like iguanas, even like basilisks, and they  
 take it and eat it.' (Monforte *et al.* 2010:195-18, 19)

Note that in (51), the nominal expression *le juuj* 'le iguana' presumably also has a generic reference. However, it does not appear with the clitic =o' because of the morphosyntactic restriction mentioned briefly in §2.1, which bans more than one clitic in the post-verbal field.

As noted in Hanks (1990:488-494), clitics in the post-verbal field can only appear in the rightmost boundary of any given clause. This is what prevents the clitic =o' from appearing immediately after *le juuj* in (49), where this noun phrase is not at the right edge of the clause. Furthermore, in this specific context the clitic =o' associated with *le juuj* cannot appear at the rightmost boundary of the clause either. This is because, as Hanks points out, only one clitic can appear at any given clitic boundary, and here the clitic position at the rightmost edge of the clause is already taken by the clitic =o' of *le tóolok=o'*. In spite of this, there is good reason to think that the clitic related with *le juuj* is =o' (thus allowing for a generic interpretation) and not, for instance, =a'. The reason is that when two clitics compete for the same clitic position, the outcome of the competition is determined by a hierarchy in which =a' takes precedence over =o' or any other clitic (Hanks 1990:491). Hence, if the clitic associated with *le juuj* were =a', it would take precedence over any other clitic competing for this position, which would result in the sequence *le tóolok=a'*, contrary to what is observed in this example.

### 5.5 Demonstrative uses of *le...o'*

So far, we have argued that *le...o* ' has all the referential functions expected of a definite article. This, however, does not mean that we do not recognize that in some contexts it behaves in the way expected of a demonstrative, as observed in several previous analyses (§2). Before introducing the relevant examples, it is necessary to tackle the question of what the semantic distinction between demonstratives and definite articles is.

We consider noun phrases with definite articles to refer inclusively—that is, their denotation includes the whole set of entities denoted by the noun and its modifiers. Demonstratives are also considered definite determiners in the sense that they also make inclusive reference (Hawkins 1991, Wolter 2006). The difference between them lies in the fact that, while definite descriptions must locate their reference set within UD, demonstratives may “update” their restriction using a pointing gesture or some other deictic device, that is, they delimit their reference set with a deictic value (‘proximity’, non-proximity’, etc.). It is by way of their inherent deictic component that demonstratives delimit the domain within which the set referred to is to be found. For instance, given a sentence like “This pan is hot”, the hearer must locate the set of pans within the domain that is closer to the speaker and then assume that this set (which by the specifications of number, must contain only one element) has the property of being hot. A speaker may say, aided by a pointing gesture: “This pan is hot but this pan is not hot” without asserting a contradiction, because the inclusive reference of the first utterance of “this pan” is restricted to the domain signaled by a first pointing gesture, while the second utterance of the same noun phrase is restricted to a different domain, delimited also deictically by a subsequent pointing gesture. In contrast, definite articles lack this deictic capacity. Building upon a simplified version of an idea put forth in Wolter (2006), we assume that the deictic component in demonstratives does not serve the purpose of locating a particular *individual* in the physical context of use, but rather delimits a particular *domain*, within which the set described by the noun phrase must be located.

Having established that the difference between demonstratives and definite articles lies in the deictic component of the former, we introduce the data in (52) as instances of a demonstrative use of *le...o* '. Note that, in accordance with Bohnemeyer (2012:106), in both cases *le...o* ' points to non-immediate objects, though not necessarily distant ones.

(52) Context: Speakers were asked to hold a large gourd in their hands and not to let go of it. A small gourd was placed at the other end of the table where the elicitation took place (out of the speakers' reach). Then they were instructed to describe each of the gourds. Example (a) is an example of the speakers' description of the gourd they held in their hands, and (b) of the other gourd.

- a. **Le** luuch=**a**', nojoch, túuts', léelemkil.  
le gourd=CL big long shiny  
'This gourd is big, long and shiny. (AH-1-85)
- b. **Le** luuch=**o**', chichan, náach, wóolis, sak.  
le gourd=CL small distant round white  
'That gourd is small, far away, round and white.' (AH-1-86)

We consider the use of *le...o'* in (52b) as an instance of a demonstrative function because the entity referred to by the corresponding noun phrase fulfills the description provided by the noun enriched with a deictic component of non-proximity. Note that it cannot simply locate the referent that fulfills the description "gourd", because there are two of them in this context. Evidently, in (52a) the form *le...a'* is also a demonstrative which, as noted in most (if not all) of the relevant literature, has the value of proximity to the speaker. Examples like (52) thus support the idea put forward in §2 that *le...a'* and *le...o'* (in its demonstrative use) pertain to a single paradigm of demonstrative determiners, in which *le...a'* is a proximity demonstrative, and *le...o'* conveys a more neutral value in distal terms, similar to what Bohnemeyer (2012:12) refers to as a "more general indexical meaning".

Another illustrative case of the demonstrative use of *le...o'* is provided in (53). Speakers accept this use in which two noun phrases with *le...o'* and identical descriptive content can co-occur in the same sentence, each one referring to a different entity. In this particular scenario, the entities referred to were equidistant from the speaker. It is important to observe that in (53) the use of two noun phrases introduced by *le...o'* is

felicitous only when each of them is accompanied by a pointing gesture which deictically restricts a different domain.

(53) **Le** luuch=**o**', nojoch, **le** luuch=**o**', chichan.

le gourd=CL big      le gourd=CL small

'That gourd is big, that (other) gourd is small.'

(DT-2-27b)

Thus, *le...o*' has the referential properties that we expect from a definite article, but (52) and (53) show that *le...o*' also has demonstrative uses, in accordance with previous descriptions. The situation in which the same form is used as a demonstrative and as a definite article is by no means unusual (see Dryer 2013). In fact, given that demonstratives are a common diachronic source of definite articles cross-linguistically, it is not surprising that, at some stage of the grammaticalization process, both the demonstrative function (i.e., a deictic determiner), and a definite article function are expressed by identical forms.

Although the present work does not intend to make any historical claims, it is likely that a similar situation may be occurring in Yucatec: the form *le...o*', identified originally as a demonstrative, coexists with an identical form *le...o*' that is currently a definite article. In the former case, the terminal clitic adds the deictic feature characteristic of demonstratives, thereby delimiting the domain of reference to a physical or conceptual space that is non-proximal to the speaker. The definite article *le...o*', in contrast, lacks this deictic feature, and this is what triggers the interpretation "set in the universe of discourse" as opposed to "set of entities in a particular domain non-proximal to the speaker". Definite articles and demonstratives, in the view adopted here, differ in the non-deictic nature of the former, as opposed to the deictic nature of the latter, but they still share the property of making inclusive reference. Along with the loss of its deictic character, further research might be able to determine if the definite article *le...o*' has also started to show additional signs of grammaticalization.

### *5.6 Other determiners are not definite articles*

What characterizes definite descriptions in our framework is the fact that they can target a referent within any of the four subdomains of UD. This condition is met by *le...o*’ but not, as we argue in this section, by the other two determiners *le...a*’ and *le...e*’. In order to show that these other determiners do not fulfill the whole range of reference location that *le...o*’ does, we must rely on negative evidence. Specifically, we must present contexts in which a certain sub-domain of UD is targeted and then show that speakers reject the use of the other two determiners in such contexts. Since negative evidence is hard—if not impossible—to find in naturally occurring texts (Matthewson 2004), for this section we rely solely on elicited data. Note that we are not interested in describing all the acceptable uses of *le...e*’ and *le...a*’, some of which might coincide with uses of definite articles in particular subdomains of UD. Rather, we are aiming to show that, unlike the definite article, these determiners are not capable of locating a referent within *just any* subdomain of UD: that is, there is at least one domain within which they fail to refer. Following standard notation, the ‘#’ sign preceding some of the following examples signals that, although grammatical, the sequence is unacceptable in the context provided.

First, in (54) we attest that *le...e*’ is not capable of referring to an entity in the subdomain of Perceptual Representation of the Speech Situation.

- (54) Context: Two gourds are placed, one in the hands of the elicitor, the other one at a certain distance, but out of reach for both interlocutors. The elicitor asks if he can refer to each of them by *le luuch=o*’, *le luuch=a*’ and *le luuch=e*’:

Referring to the gourd in the speaker’s hands:

- |    |            |                  |
|----|------------|------------------|
| a. | <b>#Le</b> | <i>luuch=o</i> ’ |
|    | le         | <i>gourd=CL</i>  |
| b. | <b>Le</b>  | <i>luuch=a</i> ’ |
|    | le         | <i>gourd=CL</i>  |
| c. | <b>#Le</b> | <i>luuch=e</i> ’ |
|    | le         | <i>gourd=CL</i>  |



Referring to the gourd out of the speaker's reach, but within sight:

- d.      **Le**      luuch=**o**'  
         le      gourd=CL
- e.      **#Le**      luuch=**a**'  
         le      gourd=CL
- f.      **#Le**      luuch=**e**'  
         le      gourd=CL

(AN-2-25)

In (54c), we see that *le luuch=e'* cannot make reference to the gourd that is held by the speaker, and in (54f) we also observe that *le...e'* cannot refer to a gourd that is placed at a certain distance, but within his/her sight. When referring to the gourd in the hands of the speaker, consultants reject the use of *le...e'*, as well as that of *le...o'*.<sup>13</sup> The only accepted alternative is *le...a'*. When referring to the second gourd, placed at a certain distance from both speaker and elicitor, all consultants accepted referring to it with *le...o'* (54d), rejected the use of *le...a'* (54e) and, once again, rejected the use of *le...e'* (54f).

The impossibility of *le...e'* being used in referring to either the gourd in the domain of the speaker, and to the one out of his/her reach, may well be due to some additional semantic requirement. Specifically, it may be the case that its referent must not be within the visual field of the speaker and/or hearer—an observation made originally in Andrade (1955 [1940]) when contrasting some constructions that use *le...e'* vs. *le...o'*. However, for the purposes of our argument, what matters is that *le...e'* cannot target the whole range of UD to locate a referent, thus eliminating it as a possible definite article. More importantly, neither *le...e'* or *le...a'* can make generic reference, as shown in the following examples.

---

<sup>13</sup> The unacceptability of *le...o'* in (54a) is due to the fact that there are two objects at different distances from the speech act participants, and so, as a demonstrative, it is used for non-proximal entities, while *le...a'* takes precedence if the referent is in the domain of the speaker. If instead of being used as a demonstrative, *le...o'* were functioning as a definite article, it still would be infelicitous because the context does not fulfill the presupposition that there is only one gourd in the relevant situation. Thanks to Lisa Matthewson for pointing this out to us.

(55) Context: Imagine that someone who is not from here does not know what iguanas are. Can I explain it to him as in (a), (b), or (c)?

- a. Le juuj=**o**'      jun túul      ba'alche'.  
le iguana=CL    one CLAS    animal  
'Iguanas are animals (lit. The iguana is an animal).'
- b. #Le juuj=**a**'      jun túul      ba'alche'.  
le iguana=CL    one CLAS    animal
- c. #Le juuj=**e**'      jun túul      ba'alche'.  
le iguana=CL    one CLAS    animal    (AN-2-11)

The unacceptability of *le...a'* and *le...e'* in making generic reference is expected if they are demonstratives, which, by definition, are not capable of referring to a whole class (García Fajardo 1991). As discussed in §2, the possible demonstrative nature of these two determiners has been scrutinized elsewhere (Andrade 1955 [1940]; Blair 1964; Bohnemeyer 2012, *inter alia*). Our purpose here is simply to show with this negative evidence that neither *le...a'* or *le...e'* are definite articles, but rather some other kind of determiner whose thorough semantic characterization is tangential to the central claim made in this paper.

## 6. CONCLUSIONS

Based on a definition of the definite article as the kind of determiner that allows a noun phrase to find its referent within any subdomain of the Universe of Discourse, our results show that noun phrases with *le...o'* are acceptable in contexts where its referent is found in any of the components of UD: (a) the Discourse Interpretation Level, (b) the Perceptual Representation of the Speech Situation, (c) the Representation of Previous Particular States and (d) Conceptual Information. In other words, these noun phrases can recover referents from previous discourse, they can refer to objects in the speech

situation, they can refer to objects that belong in the memories shared by interlocutors, or they can make generic reference. Thus, according to our criteria, *le...o* ' corresponds to a definite article. In contrast, noun phrases introduced by *le... a* ' and *le...e* ' fail to make generic reference. Additionally, *le...e* ' is not suitable for referring to entities present at the moment of speech. Therefore, these two determiners do not correspond to definite articles, since they fail to meet the corresponding criteria.

The present study confirms previous analyses that have hinted at the possibility that *le...o* ' is a marker of definiteness, based on acceptability judgments and production tasks, as well as through naturally occurring data. At this point, we make no claim about the exact semantic contribution of *le...a* ' and *le...e* '. We can assert, however, that although they are possibly definite determiners, and specifically some particular kind of demonstrative, it is clear that neither of them is a definite article, in contrast with *le...o* '.

We also confirmed the claims made in previous studies that identify the demonstrative uses of *le...o* '. In these cases, *le...o* ' may refer to entities that are not unique in fulfilling the description provided by the noun, but rather stand in contrast with some other entity with the same characteristics. In short, *le...o* ' is a definite article, but in some contexts (where there are two entities in contrast), the reference of a noun phrase with *le...o* ' does not raise a uniqueness presupposition and may be accompanied by a deictic gesture, hence fitting the characterization of a demonstrative. This raises the question of whether we are dealing with two homophonous determiners (a definite article and a demonstrative) or if *le...o* ' belongs to a third category that can fulfill all functions at the same time. For the time being, we adhere to the former explanation, but we deem that the answer to this question must rely not only on more empirical evidence but also on a more extensive theoretical discussion about what exactly distinguishes demonstratives and definite articles: if they are mutually exclusive categories, or if they can be conflated in a super-category that comprises all of their functions. The answer to this specific problem, however, is beyond the scope of this paper and so must be left for future research.

## REFERENCES

- ANDRADE, MANUEL J. 1955 [1940]. *A Grammar of Modern Yucatec*. Chicago: The University of Chicago Microfilm Collection of Manuscripts.  
<<http://www.christianlehmann.eu/ling/sprachen/maya/andrade/index.html>>
- ARKOH, RUBY, AND MATTHEWSON, LISA. 2012. A familiar definite article in Akan. *Lingua* 123:1–30
- BOHNEMEYER, JÜRGEN. 2012. Yucatec Demonstratives in Interaction. *Practical Theories and Empirical Practice*, ed. Andrea C. Schalley, pp. 99–128. Amsterdam/Philadelphia: John Benjamins.
- BLAIR, ROBERT. 1964. *Yucatec Maya Noun and Verb Morpho-Syntax*. Ph.D. dissertation, Indiana University.
- BRICEÑO CHEL, FIDENCIO. 1992. *Curso de Maya Yucateco*. Ms. Escuela Nacional de Antropología e Historia (México City).
- \_\_\_\_\_. 1993. *La cuantificación en maya: El uso de clasificadores numerales y de mensurativos*. Honors thesis, Facultad de Ciencias Antropológicas, Universidad Autónoma de Yucatán.
- \_\_\_\_\_. 1996. La frase nominal en maya yucateco. *Estudios de Lingüística Aplicada* 14: 95–104.
- BRICKER, VICTORIA; ELEUTERIO PO'OT YAH; AND OFELIA DZUL DE PO'OT. 1998. *A Dictionary of the Maya Language as spoken in Hocabá, Yucatán*. Salt Lake City: The University of Utah Press.
- COMISIÓN DE DIFUSIÓN DEL ALFABETO MAYA (CDAM). 1984. *Alfabeto maya acordado en la reunión de agosto de 1984*. Mérida: INEA/Academia de la Lengua Maya de Yucatán.
- CEH MOO, MARISOL. 2011. *T'ambilak men tunk'ulilo'ob (El llamado de los tunk'ules)*. México City: Conaculta.
- CHRISTOPHERSEN, PAUL. 1939. *The articles: a study of their theory and use in English*. Copenhagen: Munksgaard.
- DRYER, MATTHEW S. 2013. *Definite Articles*. *The World Atlas of Language Structures Online*, eds. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. <<http://wals.info/chapter/37>>.
- GARCÍA FAJARDO, JOSEFINA. 1985 [1984]. *El sentido de los sintagmas nominales y los*

- tipos de predicación. Mexico City: Instituto Nacional de Antropología e Historia.
- \_\_\_\_\_. 1989. El sentido en la función referencial de frases sustantivas del español. *Nueva Revista de Filología Hispánica* 37:19–26.
- \_\_\_\_\_. 1990. El sentido de conjunto y un tipo de presuposición. Homenaje a Jorge A. Suárez: *Lingüística indoamericana e hispánica*, eds. Beatriz Garza Cuarón and Paulette Levy, pp. 223–27. Mexico City: El Colegio de México.
- \_\_\_\_\_. 1991. El significado de los determinantes españoles. *Nueva Revista de Filología Hispánica* 39:737–52.
- \_\_\_\_\_. 1994. Hacia el universo del discurso desde la semántica formal: El artículo definido. *Segundo Encuentro de Lingüistas y Filólogos de España y México*, eds. A. Alegría, B. Garza and J. A. Pascual, pp. 221–29. Salamanca: Universidad de Salamanca.
- GILLON, CARRIE. 2015. Investigating D in Languages With and Without Articles. *Methodologies in Semantic Fieldwork*, eds. Ryan Bochnak and Lisa Matthewson, pp. 175–204. Oxford: Oxford University Press.
- GÓNGORA PACHECO, MARÍA LUISA. 1990. *U Tzikbalilo'ob Oxkutzcaab Yeetel Mani*. Mérida: Dirección General de Culturas Populares.
- GUTIÉRREZ-BRAVO, RODRIGO. 2012. Relative clauses in Yucatec Maya: Light heads vs. null domain. *Relative Clauses in Languages of the Americas: A Typological Overview*, eds. Bernard Comrie and Zarina Estrada, pp. 253–68. Amsterdam/Philadelphia: John Benjamins.
- HANKS, WILLIAM. 1984. The evidential core of deixis in Yucatec Maya. *Papers from the Twentieth Regional Meeting*, pp. 154–72. Chicago: Chicago Linguistics Society.
- \_\_\_\_\_. 1990. *Referential Practice: Language and Lived Space among the Maya*. Chicago: The University of Chicago Press.
- \_\_\_\_\_. 2005. Explorations in the deictic field. *Current Anthropology* 46:191–220.
- HAWKINS, JOHN A. 1978. *Definiteness and indefiniteness*. London: Croom Helm.
- \_\_\_\_\_. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27:405–42.

- HEIM, IRENE. 1983. File Change Semantics and the Familiarity Theory of Definiteness. *Meaning, Use, and Interpretation of Language*, eds. Rainer Bauerle, Christoph Schwarze, and Arnim v. Stechow, pp. 164–89. Berlin: de Gruyter.
- INSTITUTO NACIONAL DE ESTADÍSTICA, GEOGRAFÍA E INFORMÁTICA (INEGI). 2010. *Censo de Población y Vivienda 2010*. Mexico City: INEGI.  
<<http://www.beta.inegi.org.mx/proyectos/ccpv/2010/>>
- KRIFKA, MANFRED; FRANCIS J. PELLETIER; GREG CARLSON; ALICE TER MEULEN; GENNARO CHERCHIA; AND GODEHARD LINK. 1995. Genericity: An introduction. *The Generic Book*, eds. Gregory N. Carlson and Francis J. Pelletier, pp. 1–124. Chicago: University of Chicago Press.
- LACA, BRENDA. 1990. Generic objects: Some more pieces of the puzzle. *Lingua* 81:25–46.
- LEHMANN, CHRISTIAN. 1998. *Possession in Yucatec Maya*. Munich: Lincom Europa.  
\_\_\_\_\_. 2012. *Gramática Maya*. Erfurt: Erfurt University.  
<<http://www.christianlehmann.eu/ling/sprachen/maya/index.php>>, retrieved 08/10/2012.
- LUCY, JOHN. 1992. *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
- MATTHEWSON, LISA. 2004. On the Methodology of Semantic Fieldwork. *International Journal of American Linguistics* 70:369–415.
- MONFORTE, JORGE; LÁZARO DZUL; AND RODRIGO GUTIÉRREZ-BRAVO. 2010. *Narraciones Mayas*. Mexico City: Instituto Nacional de Lenguas Indígenas.
- PANAGIDIOTIS, E. PHOEVS. 2014. *Determiners*. Oxford Bibliographies in Linguistics.  
<<http://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0187.xml?rskey=iOyZ0V&result=46>>.
- RUSSELL, BERTRAND. 1905. On Denoting. *Mind* 14:479–93.
- SCHWARZ, FLORIAN. 2013. Two Kinds of Definites Cross-linguistically. *Language and Linguistics Compass* 7:534–59.
- STOLZ, CHRISTEL; THOMAS STOLZ; AND ELISABETH VERHOEVEN. 2012. *Maya yucateco de X-hazil Sur, Quintana Roo*. Mexico City: El Colegio de México.

- VERHOEVEN, ELISABETH. 2007. *Experiential constructions in Yucatec Maya*. Amsterdam/Philadelphia: John Benjamins.
- VERHOEVEN, ELISABETH, AND STAVROS SKOPETEAS. 2015. Licensing focus constructions in Yucatec Maya. *International Journal of American Linguistics* 81:1–40.
- WOLTER, LYNSEY KAY. 2006. *That's that: The Semantics and Pragmatics of Demonstrative Noun Phrases*. Ph.D. dissertation, University of California, Santa Cruz.