Subject preference in Ixcatec relative clauses
(Otomanguean, Mexico)
Evangelia Adamou

To cite this version:

HAL Id: hal-01709511
https://hal.archives-ouvertes.fr/hal-01709511
Submitted on 15 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Subject preference in Ixcatec relative clauses 
(Otomanguean, Mexico)*

Evangelia Adamou
French National Centre for Scientific Research (CNRS), Oral Tradition Languages and Civilizations laboratory (LACITO)

(preprint version)

Abstract

Subject preference in relative clauses (RCs) has been reported in typologically diverse languages, but overall one notes that the number of languages analyzed experimentally remains extremely low. This paper presents experimental and natural evidence from Ixcatec, a critically-endangered Otomanguean language. Ixcatec is relevant to the discussion on universal subject preference for having syntactically and morphologically ambiguous subject and object RCs that can offer an unconfounded result. Study 1, a picture-matching comprehension experiment, shows that 63% of the ambiguous RCs are interpreted as subject RCs. Results from reaction times show that subject RC interpretations are numerically faster than object RC interpretations, but this difference does not reach significance. Analysis of a three-hour, free-speech corpus in Study 2 indicates that transitive subject RCs are only slightly more frequent than object RCs. In conclusion, although the Ixcatec data support universal subject preference, they also show how this preference is weaker than predicted.

Keywords: Relative clauses, subject preference, Otomanguean.

1 Introduction

In a much cited article by Keenan and Comrie (1977) it is shown that, cross-linguistically, subjects are more accessible to relativization than direct objects, indirect objects, and oblique objects; this preference is known as the Accessibility Hierarchy. Since then, there has been a great amount of work confirming the

* Acknowledgments: First, I wish to thank the Ixcatec speakers for their collaboration in this study. Thanks are also due to Judith Aissen and the anonymous reviewers for their valuable feedback, as well as to Denis Costaouec, Niki Costaouec, Xingjia Rachel Shen, Anne Belingard and Rosemary Beam de Azcona for their help with the experiment. For the corpus collection, funding was provided by the Endangered Languages Documentation Programme Lexical Documentation of Ixcatec, a highly endangered Otomanguean language of Oaxaca awarded to Denis Costaouec (ELDP, HREL, MDP 0214 held at the University Paris Descartes from 2010 to 2013). The study on relative clauses was conducted as part of the CNRS French-Mexican research project (PICS, 2015-2017), Mesoamerica and the structure of the relative clause, co-directed by Enrique L. Palancar (CNRS) and Roberto Zavala (CIESAS).
original explanation that the Accessibility Hierarchy “directly reflects the psychological ease of comprehension” (Keenan & Comrie 1977: 88). In particular, experimental studies based on behavioural tasks and measures of brain activity offer ample evidence for the fact that subject relative clauses (SRCs) are easier to process than object relative clauses (ORCs). However, counter-evidence has also been reported for some languages and types of relative clauses (see Kwon et al. 2010 for an overview and Section 2 for more detailed references). Knowing that overall the number of languages analyzed experimentally remains extremely low, it seems important to confront processing universals to a greater number of typologically diverse languages. Indeed, roughly 57 languages are reported in a survey conducted by Anand, Chung & Wagers (2011).

In this paper I present both experimental and natural evidence relevant to this discussion from Ixcatec, a critically-endangered Otomanguean language of Mexico that has ambiguous relative clauses in the third person. This is illustrated in (1).

Ixcatec, ambiguous relative clause

(1) ndi²ra² kiʔi² sa¹ kʷa²-ʔi¹ [la² teⁿgiʔe² sa¹ miʔ-tja²]

where LOC.PRED DEF CLF.F-little COMP follow DEF CLS-woman

SRC: ‘Where is the girl that follows the woman?’ or

ORC: ‘Where is the girl that the woman follows?’

In the absence of pragmatic or semantic cues, as is the case in example (1), there are no syntactic or morphological cues that would allow the Ixcatec speakers to disambiguate between a SRC and an ORC interpretation. More specifically, Ixcatec SRCs and ORCs are formed with the gap strategy and an uninflected complementizer la². Subjects and objects both appear in the postverbal position within the relative clause. Moreover, Ixcatec is a head-marking language with a single series of subject verb suffixes restricted to first and second person, and no indexing of object arguments. In sum, the syntactic and morphological specificities of Ixcatec make it an excellent candidate for testing universal subject preference.

Indeed, following Polinsky et al. (2012: 275–276), I hypothesize that the study of fully ambiguous relative clauses should offer an unconfounded result for subject preference since comprehenders do not expect any disambiguating syntactic or morphological cue. In consequence, if disambiguating semantic and pragmatic cues are neutralized (see Warren & Gibson 2002), as in an experimental setting, comprehenders should interpret the great majority of clauses as SRCs in accordance with universal subject preference (Keenan & Comrie 1977).

The study of relative-subject preference in lesser-known languages, however, raises a number of methodological hurdles. It has been developed in well-described languages with written traditions, through methods that cannot be directly implemented for the study of under-described, unwritten, and critically-endangered languages such as Ixcatec. For the study of the subject processing advantage, the experimental method in Study 1 was inspired by the sentence-picture matching comprehension tasks employed in Polinsky (2008) for Russian and Clemens et al. (2015) for the study of two Mayan languages, Ch’ol and Q’anjob’al. Due to the age and cultural background of the Ixcatec speakers, the task was adapted to the specificities of this particular field. For the study of subject preference in natural speech (Study 2), a frequency study is based on a free-speech
corpus collected during the Ixcatec language documentation programme (2010–2013).¹

The present paper is organized as follows: Section 2 lays out the theoretical background. Section 3 offers some background on Ixcatec and presents relevant typological features. Section 4 introduces the relative clauses in Ixcatec. Section 5 presents the sentence-picture matching experiment (Study 1) and Section 6 the results of the analysis of the free-speech corpus (Study 2). Finally, Section 7 offers some conclusions and discusses theoretical implications of the Ixcatec data.

2 Theoretical background

In 2.1 I introduce the theoretical background for relative clauses and relevant relativization strategies for the study of Ixcatec. In 2.2 I present an overview of the literature on universal subject preference as expressed in the Accessibility Hierarchy and in relative clause processing. In 2.3 I focus on the specificities of ambiguous relative clauses and their importance for the discussion on subject preference.

2.1 Relative clauses

Relative clauses (RCs) can be broadly defined as follows: “A relative clause (RC) is a subordinate clause which delimits the reference of an NP by specifying the role of the referent of that NP in the situation described by the RC” (Andrews 2007: 206). For example, subordination of the clause ‘The cat catches the rat’, in (2b), to the clause ‘I see the cat’, in (2a), results in the complex clause ‘I see the cat that catches the rat’, in (2c). More specifically, (2c) is composed of a matrix clause and a relative clause (noted in brackets).

(2)  a. I see the cat.
     b. The cat catches the rat.
     c. I see the cat [that Ø catches the rat].

The head of the RC, ‘cat’ (underscored), is the object of the matrix clause. When a coreferential NP is not overtly expressed in the RC, it is dubbed gap and is noted as Ø; see (2c). When the gap has the grammatical function of a subject (S), the RC is referred to as a subject relative clause (SRC), as in (2c), and when it has the function of an object, it is termed object relative clause (ORC). The RC in (2) is postnominal, for following the head NP, as opposed to prenominal/RCs that precede the head NP. Moreover, the RC in (2) can be described as an externally headed relative clause (EHRC) as opposed to internally headed relative clauses (IHRCs) and free relatives.

An alternative relativization strategy involves the so-called resumptive pronouns. This strategy has been thoroughly discussed for languages such as Modern Hebrew (Borer 1984, Shlonsky 1992), Levantine Arabic (Shlonsky 1992), Irish (McCloskey 1990) and Modern Greek (Alexopoulou 2006). See an example

¹ Endangered Languages Documentation Programme (ELDP), Lexical Documentation of Ixcatec, a highly endangered Otomanguean language of Oaxaca (HRELP, MDP 0214), 2010–2013, PI. D. Costaouec.
from Modern Hebrew in (3) illustrating optional resumption for an ORC. The resumptive strategy in Modern Hebrew is obligatory for obliques, and it is not possible for subjects unless they are used in embedded clauses.

Modern Hebrew
(3) raʔiʔet haʔyeled [sheʔiʔasher rinaʔohevet (ʔoto)]
saw-I ACC the-boy REL Rina loves him
‘I saw the boy that Rina loves.’ (Andrews 2007: 220)

Resumptive pronouns were at first considered as alternative realizations of gaps but more research led to their analysis as ordinary pronouns (see an overview of the literature in Asudeh 2012). This discussion, however, is beyond the scope of this paper.

2.2 Relative subject preference
In a typological perspective, Keenan & Comrie (1977) have shown that subjects (both intransitive and transitive) are more accessible to relativization than direct objects, indirect objects, oblique objects, possessors, and objects of comparison. To render this preference the authors have proposed the Accessibility Hierarchy, illustrated in (4):

(4) subject > direct object > indirect object > oblique object > possessor > object of comparison

This means that if a language can relativize a position in the Accessibility Hierarchy, it can relativize any higher position, either directly or via syntactic promotion to a higher position. We note that subject preference can also be treated as a “subject constraint” in some languages, such as Malagasy, which can only relativize subjects. As a result Malagasy speakers need to promote major NPs to the subject position through a voice system (Keenan & Comrie 1977: 69–70).

The Accessibility Hierarchy has since been refined by several scholars. For example, following Lehmann (1986), subsequent publications no longer integrate the possessor and the object of comparison in this hierarchy due to counter-evidence. Hawkins (1999) further developed the Accessibility Hierarchy by relating it to relativization strategies. The proposal is that if the gap strategy is grammatical in a given position in the hierarchy, then it will be grammatical for all the higher positions (Hawkings 1999: 256). But if the resumptive pronoun strategy is grammatical in a given position in the hierarchy, then it will be grammatical for all the lower positions (Hawkings 1999: 258).

Although the Accessibility Hierarchy has been largely corroborated by cross-linguistic evidence, there is an ongoing discussion as to what drives subject preference in the languages of the world. A great amount of experimental work has confirmed the original explanation that the Accessibility Hierarchy “directly reflects the psychological ease of comprehension” (Keenan & Comrie 1977: 88). Indeed, a subject processing advantage is reported for English RCs with different methodologies, e.g., Event Related Potentials (ERP) in King & Kutas 1995, functional Magnetic Resonance Imaging (fMRI) in Just el al. 1996, Caplan et al.

Similar results are obtained for a number of accusative languages, e.g., German (Mecklinger et al. 1995), Dutch (Frazier 1987), Hungarian (MacWhinney & Pleh 1988), Russian (Polinsky 2008), Hebrew (Friedmann & Novogrodsky 2004). For ergative languages, Carreiras et al. (2010) have questioned the subject processing advantage in Basque, whereas subject preference has been demonstrated in Avar (Polinsky et al. 2012), a Northeast Caucasian language, and in two Mayan languages, Ch'ol and Q'anjob'al (Clemens et al. 2015). Also, although most studies from languages with prenominal RCs confirm subject preference in processing, e.g., Japanese (Miyamoto & Nakamura 2003; Ueno & Garnsey 2008), and Korean (Kwon et al. 2010; Kwon et al. 2013), research on Chamorro indicates that subject preference is weaker in prenominal RCs than in postnominal RCs (Borja, Chung & Wagers 2016). Moreover, there is conflicting evidence for Chinese for which Hsiao & Gibson (2003) report object preference, while more recent studies indicate subject preference (Vasisht et al. 2013).

The reasons behind the cognitive ease of SRCs are also under debate. Memory-based accounts suggest that ORCs are more difficult to process due to working memory capacity (King & Just 1991). Gibson (1998) argues that the greater memory load in ORC processing results from greater integration costs related to the distance in filler-gap dependencies. Warren & Gibson (2002) further show that the cognitive burden related to syntactic complexity can be attenuated by semantic and pragmatic factors.

Moreover, it is generally admitted that frequency and experience facilitate processing. Wells et al. (2009) demonstrated through statistical learning that corpus frequencies affect processing ease. It is thus argued that the relative subject processing advantage in English is related to the fact that SRCs are more frequent in written and spoken corpora (65% in Reali & Christiansen 2007; 67% in Duffield & Michaelis 2011). MacDonald & Christiansen (2002), Wells et al. (2009), and MacDonald (2013) suggest that the processing difficulty for ORCs results from a mismatch between the linguistic regularities that favour SRCs and the actual linguistic input. More specifically, MacDonald (2013) hypothesizes that individual speakers opt for the least costly means in language production, i.e., SRCs; these individual choices turn into statistical regularities affecting larger populations, i.e., statistic prevalence of SRCs; and finally, individual comprehenders largely rely on these regularities when processing linguistic input, i.e., processing advantage for SRCs. However, Desmet & Gibson (2003) highlighted the need to carry out more fine-grained corpus analyses when discussing relations between frequency effects and processing. Indeed, a more fine-grained analysis in Gordon & Hendrick (2005) shows that transitive SRCs are less frequent than ORCs.

In sum, despite a lively and active debate, it has been shown that processing is easier for SRCs than for ORCs, reflecting a universal subject preference, and that this cognitive ease is influenced by pragmatic, semantic, syntactic, and frequency-related factors.
2.3 Ambiguous relative clauses

To obtain an unconfounded result for universal subject preference, Polinsky et al. (2012: 275) suggest that researchers need to take a closer look at languages with ambiguous RCs. The rationale behind this statement is that the study of ambiguous RCs excludes the morphological and syntactic bias which is otherwise present and interacts with the universal subject preference. For example, nominal case influences the comprehension of RCs, e.g., in accusative languages nominative case generally codes subjects and therefore the use of a head NP with nominative case favours a SRC interpretation. Also, congruent word order between RCs and main clauses is likely to facilitate the processing of RCs and vice-versa (MacDonald & Christiansen 2002).

In the absence of syntactic and morphological cues, speakers of languages with ambiguous RCs are likely to rely more heavily on semantic and pragmatic criteria. It follows that if one reduces the semantic and pragmatic cues in an experimental setting, using agents equal in animacy and visual stimuli depicting reversible actions as in Polinsky (2008) and Clemens et al. (2015), one also greatly reduces the confounds to universal subject preference. The experimental results could then be compared to frequency studies of natural speech, as frequency is known to affect syntactic ambiguity resolution more generally (Desmet & Gibson 2003).

Several experimental studies have explored processing in temporary ambiguous RCs in which ambiguity is lifted in the last word. This is illustrated for German, a verb-final language. The RCs in (5) are ambiguous for the feminine noun phrases and the feminine relative pronoun die and are only disambiguated through the auxiliary, hat or haben, which appears in final position.

German, temporary ambiguous RCs

(5) a. Das sind die Professorinnen, die die Studentin gesucht hat.
   (These are the professors that the student sought has.)

   b. Das sind die Studentinnen, die die Professorin gesucht haben.
   (These are the students that the professor sought have.) (Mecklinger et al. 1995: 479)

The studies of temporary ambiguous RCs in German (Mecklinger et al. 1995) and in Dutch (Frazier 1987) strongly support subject preference. In these studies, participants seem to opt first for the subject interpretation, and when necessary proceed to syntactic reanalysis.

In contrast, the Basque data from temporary ambiguous RCs show a clear object advantage, associated to absolutive gaps (Carreiras et al. 2010). Carreiras et al. (2010) observe that this advantage does not follow statistical frequencies of RCs in corpora, but Clemens et al. (2015) argue that frequency of nominal case is a potential explanatory factor. More specifically, they note that the Basque stimuli start with an ambiguous NP in the RC that can be interpreted either as an absolutive plural (associated to an ergative singular gap), as in (6b), or as an ergative singular (associated to an absolutive plural gap), as in (6c). They further suggest that an ergative singular interpretation of the noun in the RC (associated to an absolutive gap), as in (6c), could have been favoured because ergative singular case is more frequent than absolutive plural case in speech corpora.
Basque
(6) a. Ambiguous RC
Irakasleak aipatu dituen ikasleak.
SRC: ‘The student who mentioned the teachers...’ or
ORC: ‘The students whom the teacher mentioned...’

SRC
b. [Ø Irakasle-ak aipatu ditu-en] ikasle-a-k
   teacher-ABS.PL mentioned has-REL student-SG-ERG
   ‘The student who mentioned the teachers...’ (adapted from Carreiras et al. 2010: 82 and Clemens et al. 2015)

ORC
c. [Irakasle-a-k Ø aipatu ditu-en] ikasle-a-k
   teacher-SG-ERG mentioned has-REL student-ABS.PL
   ‘The students whom the teacher mentioned...’ (adapted from Carreiras et al. 2010: 82 and Clemens et al. 2015)

More recently, a study on Chamorro, an Austronesian language, tested subject preference in ambiguous RCs for both postnominal and prenominal RCs. Interestingly, the results show object preference in prenominal RCs, but subject preference for postnominal RCs as well as a subject processing advantage in both prenominal and postnominal RCs (Borja, Chung & Wagers 2016).

Another study tested subject preference for two ergative Mayan languages, Ch’ol and Q’anjob’al (Clemens et al. 2015). In (7) I present an example of an ambiguous RC in Ch’ol.

Ch’ol
(7) a. Ambiguous RC
Ta’ juli jiñi x’ixik ta’bä itsäk’ä jiñi wiñik.
SRC: ‘The woman that cured the man arrived’ or
ORC: ‘The woman that the man cured arrived.’ (Clemens et al. 2015)

SRC
b. Ta’ jul-i jiñi x'ixik [ta’-bä i-tsäk’-ä jiñi wiñik Ø]
   PRFV arrive-ITV DET woman PRFV-REL 3.ERG-cure-TV DET man
   ‘The woman that cured the man arrived.’ (Clemens et al. 2015)²

ORC
c. Ta’ jul-i jiñi x’ixik [ta’-bä i-tsäk’-ä Ø jiñi wiñik.]
   PRFV arrive-ITV DET woman PRFV-REL 3.ERG-cure-TV DET man
   ‘The woman that the man cured arrived.’ (Clemens et al. 2015)

² Authors’ glosses: DET—Determiner; ERG—Ergative; ITV—Intransitive; PRFV—Perfective; REL—Relative; TV—Transitive.
It can be seen that Ch’ol RCs are constructed with the gap strategy and that subjects and objects both appear in the postverbal position within the RC. Moreover, there is no case on the nouns that would allow disambiguation between subject and object interpretation. The results of processing for the ambiguous RCs in Ch’ol confirm universal subject preference with 68% of the responses favouring an ergative-subject interpretation (Clemens et al. 2015). The study also shows that participants were faster in the subject responses than in the object ones.

The remainder of the paper presents evidence relevant to this discussion from Ixcatec, an under-studied language of Mexico with fully ambiguous RCs.

3 Background on Ixcatec

In 3.1 I provide some sociolinguistic information about Ixcatec. In 3.2 I present a typological overview of Ixcatec and introduce relevant constructions for the study of RCs, such as information structure-related phenomena and interrogative clauses.

3.1 Sociolinguistic overview

\[\text{\textit{\textit{\textit{Wa}n\textit{ni}}} is an Otomanguean language better known in the literature under its name in Nahuatl, \textit{Ixcatec} (ISO code: \textit{ixc}). Ixcatec is a “critically endangered language” (cf. endangerment scale in Krauss 2006), spoken by less than ten speakers of whom only four are fluent. This situation is the result of a century-long shift to Spanish, beginning with Spanish colonization in 1519. Contact intensity with Spanish was most likely low in the past centuries, characterized by scarce contact with Spanish speakers involving only a small part of the community members (Adamou 2016). A rapid process of shift to Spanish took place during the early twentieth century and is nowadays completed. The last Ixcatec speakers live in the municipality of Santa Marí Ixcatlán in the State of Oaxaca, in Mexico (see Map 1). Even though nowadays Santa Marí Ixcatlán has roughly 400 inhabitants, at the time of the arrival of the Spaniards, in 1522, it was an important centre for the Mixteca zone with an estimated population of 10,000 to 30,000 people (Hironymous 2007).

Map 1. Santa Marí Ixcatlán, State of Oaxaca, Mexico
Ixcatec belongs to the Popolocan branch of the Otomanguean stock together with Chocho, Popoloc, and Mazatec (Fernández de Miranda 1951, 1956; Hamp 1958; Gudschinsky 1959; Swadesh 1960: 83). The published work on Ixcatec consists of the phonology and dictionary by Fernández de Miranda (1959, 1961) as well as her work on possessives (Fernández de Miranda 1953). Veerman-Leichsenring has published on Ixcatec nominal phrases (Veerman-Leichsenring 2001a) and discussed the Ixcatec data in several other papers (e.g., Veerman-Leichsenring 2000). More recently, work has been done on Ixcatec phonetics and phonology (Alarcon Montero 2010), on syntax and morphology (Adamou & Costaouec 2013; Adamou 2014; Costaouec & Swanton 2015), and on spatial language and cognition (Adamou 2017).

### 3.2 Overview of relevant features in Ixcatec

Ixcatec has three lexically contrastive tones, a high tone transcribed with a superscripted \(^{1}\), a mid tone, transcribed with \(^{2}\), and a low tone, transcribed with \(^{3}\). It is a head-marking language, i.e., grammatical relations are marked on the verb (Nichols 1986), and core arguments are not coded through case morphemes or adpositions. Ixcatec has accusative alignment in indexing (A = S ≠ P) (Malchukov et al. 2010), i.e., only the single argument of intransitive verbs (S) and the agent-like argument of transitive verbs (A) are indexed on the verb through a series of suffixes, with third person being zero; see Table 1. We also note that a dozen experience predicates take a different coding for person, namely through possessive suffixes; see Table 1.

Table 1. Person marking in Ixcatec through verbal suffixes and possessive suffixes for a class of experience predicates

<table>
<thead>
<tr>
<th></th>
<th>Intransitive ‘enter’ (perfective)</th>
<th>Transitive ‘cut’ (perfective)</th>
<th>Experience predicate ‘be sick’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>kwi(^{2})tẽ(^{2})ʔẽ(^{2})(\text{ma}^{3})</td>
<td>kwi(^{2})t’u(^{2})-(\text{na}^{3})</td>
<td>mĩ(^{2})ʔi(^{1})-(\text{na}^{4})</td>
</tr>
<tr>
<td>2SG</td>
<td>kwi(^{2})tẽ(^{2})ʔ-a(^{2})</td>
<td>kwi(^{2})t’w-a(^{2})</td>
<td>mĩ(^{2})ʔi(^{1})-a(^{2})</td>
</tr>
<tr>
<td>3SG</td>
<td>kwi(^{2})tẽ(^{2})ʔẽ(^{2})</td>
<td>kwi(^{2})t’u(^{2})</td>
<td>mĩ(^{2})ʔi(^{1})-e(^{2})</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>kwi(^{2})tẽ(^{2})ʔ-i(^{2})</td>
<td>kwi(^{2})t’w-i(^{2})</td>
<td>mĩ(^{2})ʔi(^{1})-i(^{2})</td>
</tr>
<tr>
<td>1PL.EXCL/HON</td>
<td>kwi(^{2})tẽ(^{2})ʔẽ(^{-2})-ri(^{2})</td>
<td>kwi(^{2})t’u(^{-2})-ri(^{2})</td>
<td>mĩ(^{2})ʔi(^{1})-na(^{2})ri(^{2})</td>
</tr>
<tr>
<td>2PL</td>
<td></td>
<td></td>
<td>mĩ(^{2})ʔi(^{1})-ri(^{2})</td>
</tr>
<tr>
<td>3PL</td>
<td>kwi(^{2})tẽ(^{2})ʔẽ(^{-2})-ma(^{2})</td>
<td>kwi(^{2})t’u(^{-2})-ma(^{2})</td>
<td>mĩ(^{2})ʔi(^{1})-e(^{2})ma(^{2})</td>
</tr>
</tbody>
</table>

Ixcatec is a pro-drop language, i.e., free pronouns are optionally used for all functions, and NPs are generally omitted. Ixcatec has a VS unmarked order, as exemplified in (8a) and (8b). In main clauses, S arguments are focused in the preverbal position, as shown in (8c), but in adverbial clauses S arguments are focused \textit{in situ}, as shown in (8d) (the focused noun ‘young man’ is marked by an optional focus suffix ‘-\(\text{na}^{2}\)’ which has no extra pragmatic meaning).
(8) a. \( kw\-a^2ni^1h\=\=nw^1 \)  
   PFV-fall  CLS-little  
   ‘A boy falls.’ (RRM-pear story)\(^3\)

b. \( fu^2wo^2\-ku^1 \) \( b\=\=ngu^1 \) \( kw^a^2 \)  
   come\-ANT one\ woman
   ‘One woman has come.’ (RRM-pear story)

c. \( sa^1 [kw^a^2\-ni^1ki^2]_F \)  
   DEF CLF.F\( \cdot \)NP \( \) PROG.3SG\( \cdot \)dance\\( \cdot \)CO.3SG.F
   ‘NIKI is dancing.’ (JSB-elic)

d. \( ka^2ndi^2 \) \( ba^2tu^1\=\=phi^2 \) \( kw^a^2 \)  
   when\ IPFV.PL\( \cdot \)go\ all\ CLS\-man\-FOC
   ‘When all the YOUNG MEN are going...’ (JSB-conv)

Comparison between (8b) and (8c) shows that when an S argument is moved to the preverbal position, a cross-reference morpheme is suffixed on the verb, functioning as a resumptive pronoun. Ixcatec has three cross-reference morphemes (\( \cdot da^2 \) ‘male’, \( \cdot kw^a^2 \) ‘female’, and \( \cdot ba^2 \) ‘animal’), suffixed on verbs, possessives, and predicate adjectives. The Ixcatec cross-reference morphemes corefer to nouns formed with the noun classifiers \( di^2 \) ‘man’, \( kw^a^2 \) ‘woman’, \( ?u^2 \) ‘animal’ (glossed CLF), to some animate nouns even though they have no classifier, and to the masculine and feminine third singular pronouns which bear the same suffixes as those used for the cross-reference morphemes, i.e., \( su^wa^1\-da^2\he \) and \( su^wa^1\-kw^a^2 \) ‘she’; see Table 2. Ixcatec noun classifiers were grammaticalized from lexical material; this is also the case for the cross-reference morphemes to which was added a deictic suffix –\( a \) (Veerman-Leichsenring 2001c). Following Grinevald (2000), I distinguish noun classifiers in Ixcatec from so-called class terms (glossed CLS): the latter partake in word formation for inanimates but are not associated with any cross-reference morphemes.

<table>
<thead>
<tr>
<th>Cross-reference morpheme</th>
<th>NP</th>
<th>Examples</th>
</tr>
</thead>
</table>
| \( V\cdot da^2 \) | CLF.M\( \cdot \)N | \( di^2\so\=\iba\-du\ibu \) ‘soldier’, \( di^2\-ta^1\-ta^1\-ni^2 \) ‘our father’, \( di^2\-pe\=\ibru\=\= Pedro \)  
   | PRO\( \cdot \)da^2 | \( su^wa^1\-da^2\=\=3SG.M \) |
| \( V\cdot kw^a^2 \) | CLF.F\( \cdot \)N | \( k\=\=\aa^2\-g\=\=a\ib\=\= \) ‘girl’, \( k\=\=\aa^2\-h\ib\=\=a^2\-ni\ib^1\ibu \) ‘Juanita’ \( su^wa^1\-k\=\=\aa^2\=\=3SG.F \)  
   | PRO\( \cdot \)kw^a^2 | \( ne\ib\=\=e\=\=e\; \) ‘his mother’, \( k\=\=\aa^2\ib \) ‘woman’ |
Ixcatec monotransitive main clauses have a rigid SVO order. Compare the examples in (9) from responses to the Questionnaire on Information Structure (Skopeteas et al. 2006). The transitive arguments (A) are preverbal whether they are discourse-given, as in (9a), or discourse-new, as in (9b).

(9) a. \( sa^1 \) \( k^w^a^2 \) \( ki^1=?u^te^2ka^1-k^w^a^2 \) \( sa^1 \) \( mi^2-nda^2wa^2 \)
   \( DEF \) \( woman \) \( PROG.3SG\cdot push\cdot \) \( DEF \) \( CLS\cdot man \)
   \( CO.3SG.F \)
   ‘The woman is pushing the boy.’ (QUIS tasks)

   b. \( sa^1 \) \( k^w^a^2-\tilde{I}^i \) \( ki^1=???u^te^2ka^1-k^w^a^2 \) \( sa^1 \) \( li^2-\tilde{I}^i \)
   \( DEF \) \( CLF\cdot \) \( little \) \( PROG.3SG\cdot push\cdot \) \( DEF \) \( CLS\cdot little \)
   \( CO.3SG.F \)
   ‘The girl is pushing the boy.’ (QUIS tasks)

In the contemporary data, the A arguments in the preverbal position generally trigger the cross-reference morphemes on the verb, but this is variable in the corpus of the 1950s. The optionality of resumption via the cross-reference morphemes attested in transitive clauses seems to be related to a change in word order from a VSO order as noted by Veerman-Leichsenring (2001b: 311) and Adamou (2016: 152–153). There are two arguments in favour of such an analysis. First, the two most closely related languages of the Popolocan branch, Popoloc and Chocho, exhibit an unmarked VSO order and movement of the transitive S to the preverbal focus position triggers the use of a cross-reference morpheme on the verb (Veerman-Leichsenring 2001b: 94). Second, in a typological perspective, the unmarked VS word order noted in Ixcatec is consistent with a verb-initial order in transitive clauses, as the position of subjects across the languages of the world tends to be the same in both intransitive and transitive clauses (Dryer 2011).

Objects are generally focused in situ, i.e., postverbally; see (10a) for a main clause and (10b) for an adverbial clause. In these examples both nouns receive the optional focus suffix ‘-na\(^\tilde{}\).

(10) a. \( k^w^a^2-?\tilde{a}^2?u^1\cdot ku^1\cdot na^3 \) \( sa^1 \) \( [t^j^1][u^2]\cdot F\cdot na^2 \)
   \( PFV\cdot throw\_in\cdot ANT\cdot 1SG \) \( DEF \) \( tempesquisles\cdot FOC \)
   ‘I have thrown in the \( tempesquisles \) (kind of vegetables).’ (JSB\_conv)

   b. \( k\tilde{a}^2ndi^2 \) \( k^w^a^2hi^2-ku^1 \) \( sa^1 \) \( [n^i^2nu^2]\cdot F\cdot na^2 \)
   when \( PFV\cdot come\_ANT \) \( DEF \) \( tortilla\cdot FOC \)
   ‘Once one has got the tortillas...’ (JSB\_conv)

Ixcatec yes/no questions are introduced by an optional interrogative particle ‘?a\(^\tilde{}\), see (11a) and (11b). Example (11b) also illustrates that the interrogative particle precedes a focused NP; note that the numeral ‘five’ receives the optional focus particle ‘-na\(^\tilde{}\) and is realized in the unmarked order NUM\_N.

\( V\cdot ba^4 \) | \( CLF.\ AN\_N \) | \( ?u^4\cdot ku^2\cdot ra^2\cdot fi^2 \) | ‘rooster’, \( ?u^4\cdot a^1\cdot hi^2\cdot a^2 \) | ‘spider’
(11) a. ʔa² kw-i² hi² ki²tʃu²hu¹
    QUEST PFV-arrive brother.POSS.3SG
    ‘Did his brother arrive?’ (RRM-conv)

b. tu¹hu² ka² nū¹hū¹-ra²
    PROG.PL.be all four·DEM.DIST

ʔa² [ʃũ²]F-na² mi²-nda²wa² tu¹hu²
    QUEST five·FOC CLS·man PROG.PL.be
    ‘There are all four there. Is it FIVE men there?’ (RRM-pear story)

Ixcatec wh words, such as ja’ra² ‘who’, nda’ra² ‘what’, ndi’ra² ‘where’, always precede verbs in direct questions, in (12a), and indirect questions, in (12b). In terms of word order, (12a) shows that in an interrogative clause an A argument follows the verb as opposed to the rigid preverbal order observed in main clauses, exemplified in (9); the wh-object is preverbal. Unlike what we have seen for intransitive main clauses in which a focused S argument in the preverbal position triggers a cross-reference morpheme, notice that a fronted wh-subject does not trigger a cross-reference morpheme on the verb; see (12b).

(12) a. he² nda’ra² tse² sa¹ ta’ʔtʃi²-ra²
    now what do DEF old_man·DEM.DIST
    ‘Now, what is that old man doing?’ (RRM-pear story)

b. tʃi²kʷi² na² ja’ra² be²tu²ma²
    who_knows who order
    ‘Who knows who commands?’ (JSB-conv)

Ixcatec shows inversion of nouns in a nominal possession structure, a phenomenon described for other Mesoamerican languages (see Aissen 1996: 457 for Tzotzil, Mayan). Compare the [possessor_noun₁ possessed_noun₂-POSS] reverse order in (13a), with the ordinary structure [possessed_noun₁-POSS possessor_noun₂] shown in (13b).

(13) a. tʃi²kʷi² na² ja’ra² ndi¹-ʔe²
    who_knows who house-POSS.3SG
    ‘Who knows whose house.’ (JSB-conv)

b. ndi¹-ʔe² sa¹ di²-ra² pa²dri² no²
    house-POSS.3SG DEF sir·DEM.DIST godfather
    ‘At godfather’s house.’ (RRM-conv)

Example (14a) illustrates the syntactic phenomenon known as pied piping involving one of the few Ixcatec prepositions, the comitative ku² ‘with’ which accompanies the noun ‘boy’ in the preverbal focus position (the noun is also marked by the optional focus marker -na⁴). (14b) shows that fronting of the wh-word in an interrogative clause triggers an applicative-instrumental suffix on the verb which
is otherwise not used (see 14c). Compare (14a) and (14b) with (14c) in which the noun and the comitative ku² ‘with’ (used as an instrumental) both follow the verb.

(14) a. ku² ka² [li²-ʔi¹]ₐ-na² ti¹=ϕi²-na³
    with all CLS-little-FOC PROG.1/2SG-go-1SG
    ‘I’m going with all the BOYS.’ (JSB-conv)

b. nda¹ra² kʷ-a¹ne²-fi²
    what FUT-eat-APPL.INS
    ‘What will one eat it with?’ (JSB-conv)

c. ka²ndi² kʷ-a²ne² ku² nda²⁷na³
    when PFV-eat with sauce
    ‘When one eats with sauce...’ (JSB-conv)

Lastly, (15) shows an example of a free relative in which the wh-word (with no interrogative semantics) is used preverbally.

(15) ja¹ra² tse²-ʔa²na² ja² be²-ʔa²na²
    who do-NEG work IPFV.eat-NEG
    ‘Whoever does not work, does not eat.’ (JSB-elic)

To summarize, Ixcatec interrogative and focused words are moved to a preverbal, syntactically marked position, but only focus movement is associated with the use of resumptive, cross-reference morphemes. As we will see, these syntactic features are also relevant for the Ixcatec RCs which are presented in 4.

4 Relative clauses in Ixcatec

4.1 presents the Accessibility Hierarchy in Ixcatec for the gap strategy based on elicited data. 4.2 is dedicated to the ambiguous Ixcatec RCs and is based on a translation task using visual stimuli.

4.1 The gap strategy and the Accessibility Hierarchy
In Ixcatec, RCs are generally externally headed⁴ and are always postnominal. They are formed using the gap strategy and an optional invariable complementizer la²; see an example in (16c).

(16) a. ʔi²na¹na³ kʷ-i²ʃkũ²-na³ sa¹ ʔu²-tʃi¹tu¹
    1SG PFV-see-1SG DEF CLF-AN-cat
    ‘I see the cat.’ (JSB-elic)

b. sa¹ ʔu²-tʃi¹tu¹ kʷ-i¹ka¹ sa¹ ʔu²-ra¹ʔi³
    DEF CLF-AN-cat PFV-catch DEF CLF-AN-rat
    ‘The cat catches the rat.’ (JSB-elic)

⁴ Some rare examples of internally headed RCs have been identified in the spontaneous corpus but more research is required to account for this type of RCs.
c. ʔi²na¹na³ kʷ-i²j[kũ²-na³ sa¹ ?u²-tʃi¹tu¹
1SG PFV-see-1SG DEF CLF.AN-cat
[la² kʷ-i¹ka¹ Øs sa¹ ?u²-ra'tʔi³]
COMP PFV-catch DEF CLF.AN-rat
‘I see the cat that catches the rat.’ (JSB-elic)\(^5\)

The complementizer \(la^2\) also forms complement and adverbial clauses; see (17a) and (17b) respectively.

(17) a. sa¹ kʷa²-eⁿφe²rmė'ra ku²-tʃe²-ku¹-na³
DEF CLF.F-nurse PFV-tell-ANT-1SG
la² nda² jṭa¹ si¹
COMP how ugly EXS.PRED
‘I told the nurse, how ugly it is!’ (Adamou & Costaouec 2013: 193)

b. kʷa²tu'=bì²hi²-ʔa²na²-na² la² má¹hũ¹
PFV.PL-arrive-NEG-FOC SUB sweep
‘Won’t they come to sweep?’ (Adamou & Costaouec 2013: 193)

The gap strategy in Ixcatec is used for the relativization of subjects, as in (18). The gap strategy is also used to relativize direct objects, in (19), indirect objects, in (20), oblique objects, in (21), and possessors, in (22).

Subject
(18) sa¹ a³lu'mnu² [la² ki¹=tʃa²ʃi² Øs]
DEF student COMP PROG.3SG-read
‘The student that is studying.’ (JSB-elic)

Direct object
(19) ni²nu² [la² ku¹=tʃe²ʔe²-na³ Ø₀ sa¹ kʷa²-ʃu²ŋgu²]
tortilla COMP PFV-give-1SG DEF CLF.F-young_woman
‘The tortillas that I give to the young woman.’ (JSB-elic)

Indirect object
(20) sa¹ li²-ʔi¹ [la² sa¹ kʷa²-ru²fi¹na² ku²=tʃe²-kʷa²
DEF CLS-little COMP DEF CLF.F-NP PFV-tell-CO.3SG
hŋgu² i²sto'ria² Ø₀]
one story
‘The kid that Rufina tells a story to.’ (JSB-elic)

\(^5\) The subject gap in this example is postverbal because the unmarked position of subjects in the Ixcatec RCs is postverbal. However, the position of the subject gap before the object is rather arbitrary; the unmarked word order in the Ixcatec transitive main clauses is SVO while the two arguments are never expressed simultaneously in the RCs.
In conclusion, and as illustrated in Table 3, Ixcatec allows relativization of any position on the Accessibility Hierarchy (Keenan & Comrie 1977) through the gap strategy and the use of the complementizer la².

<table>
<thead>
<tr>
<th></th>
<th>Subject</th>
<th>Direct object</th>
<th>Indirect object</th>
<th>Oblique object</th>
<th>Possessor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COMP + gap</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

### 4.2 Ambiguous relative clauses in Ixcatec

In this section I focus on third person subject and object RCs which are potentially morphologically and syntactically ambiguous. Consider again the Ixcatec RC in (16c), repeated in (23). Even though the semantics and the pragmatics favour the subject interpretation, ‘I see the cat that catches the rat’, the Ixcatec RC is morphologically and syntactically ambiguous: there is no case marking, no agreement and no word order that could indicate whether this is a SRC, as in (23a), or an ORC, as in (23b).

**SRC**

(23) a. ?i³na¹na³ kʰ-i²ʃkʰ²-na³ sa¹ ?u²-tʃi²tu¹
    1SG kʰ-see-1SG DEF CLF.AN-cat

    [la² kʰ-i¹ka¹ Øs sa¹ ?u²-ra⁴tʔi³]
    COMP PFV-catch DEF CLF.AN-rat

    ‘I see the cat that catches the rat.’

**ORC**

b. ?i³na¹na³ kʰ-i²ʃkʰ²-na³ sa¹ ?u²-tʃi²tu¹
    1SG kʰ-see-1SG DEF CLF.AN-cat

    [la² kʰ-i¹ka¹ sa¹ ?u²-ra⁴tʔi³ Øo]
    COMP PFV-catch DEF CLF.AN-rat

    ‘I see the cat that the rat catches.’

In order to examine third person RCs in detail, a translation task supported by visual stimuli was conducted. Prior experience with questionnaires via translation revealed that it was not possible to control for the responses, especially when these were ambiguous RCs. The visual stimuli thus aimed at ensuring that the speakers
correctly identified the situation that they were asked to describe in translation from Spanish.

4.2.1 Materials
Thirty-six pictures depicting actions between participants of equal animacy were produced for the translation task. 30 pictures involved human participants that the speakers were familiar with in order to minimize recognition problems that were encountered in previous studies conducted with standardized stimuli (e.g., QUIS tasks). Six pairs of inanimates were also depicted in this task, using Mexican handcrafted animal figures and other everyday objects. The target words for the human and inanimate participants were selected in order to trigger a cross-reference morpheme in Ixcatec.

4.2.2 Participants
The participants in this study were the four most fluent speakers of Ixcatec, two male and two female, all in their 80s. Unlike the majority of the Ixcatecs born in the first part of the twentieth century, the last Ixcatec speakers learned Ixcatec from having attended less schooling than their peers and from having spent more time with their monolingual Ixcatec grandparents. Three of the four speakers learned Ixcatec from their parents in a bilingual context, using Ixcatec and Spanish at home, and Spanish at the community and at school. One of the speakers was raised monolingual in Ixcatec and learned Spanish from school. All four speakers have not been using Ixcatec in their everyday lives for more than 50 years. They have participated at the Ixcatec language documentation programme (2010-2013) and used Ixcatec for the recording sessions. During the same period, three of the four speakers also gave classes of Ixcatec at the kindergarten, primary, and secondary school, once per week, with the collaboration and assistance of the Spanish-speaking teachers. Participants were paid for their participation in this study.

4.2.3 Procedure
The task took place in the houses of the participants at Santa María Ixcatlán in September 2015. Each pair of pictures was placed on a table and the researcher asked a question in Spanish, once for a SRC and once for an ORC. The participants were instructed to translate from Spanish the sentences that corresponded to the pictures by using an interrogative clause: “Where is X that is doing Y to Z?”.

The elder male speaker was presented with the visual stimuli of the translation task before conducting the comprehension experiment (Study 1 presented in Section 5). The other three speakers conducted the translation task once the comprehension experiment (Study 1) was completed. The translation task lasted 30 to 45 minutes. The sessions were recorded with a Marantz solid-state recorder and an external microphone.

4.2.4 Coding
In total, the task provided 288 transitive RCs. Responses were annotated and coded as 1) “ambiguous”, when there were no morphological or syntactic cues; 2) “word order”, when in the RC the argument was preverbal; 3) “cross-reference morpheme”, when the marked word order in the RC had triggered a cross-reference
morpheme on the verb that could differentiate the two participants based on
gender; and 4) “non-applicable”, when the response was not a RC or was disfluent.

4.2.5 Results

The translation task provided four types of RCs: 1) morphologically and
syntactically ambiguous RCs; 2) ORCs with an SV order in the RC; 3) ORCs with
an SV order in the RC associated to a cross-reference morpheme on the verb; and
4) SRCs with cross-reference morphemes on the verb functioning as resumptive
pronouns for the head of the RC.

A quantitative look at the results of the translation task shows that the Ixcatec
speakers in most cases translated the Spanish SRCs and ORCs in an ambiguous
manner (N=204 or 79% of the correct responses). The second most frequent
strategy relied on marked word order, SV, within the RC; this strategy was only
used for the ORCs (N=43 or 17% of the correct responses) and was particularly
favoured by the two male speakers. In the ORCs with marked word order, speakers
PSG and RRM used the cross-reference morphemes on the verb such that for some
examples it was possible to distinguish between agent and patient (N=10 or 4% of
the correct responses). This was not the case for the ORCs with marked word order
produced by speaker CRG. Table 4 summarizes the various strategies per speaker.

Table 4. Strategies used by the Ixcatec speakers in the translation task for SRCs
and ORCs

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Ambiguous</th>
<th>Word-order change</th>
<th>Cross-reference morphemes</th>
<th>N.A.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JSB</td>
<td>68</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>72</td>
</tr>
<tr>
<td>PSG</td>
<td>48</td>
<td>14</td>
<td>4</td>
<td>6</td>
<td>72</td>
</tr>
<tr>
<td>CRG</td>
<td>45</td>
<td>25</td>
<td>0</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td>RRM</td>
<td>43</td>
<td>3</td>
<td>6</td>
<td>20</td>
<td>72</td>
</tr>
<tr>
<td>Total</td>
<td>204</td>
<td>43</td>
<td>10</td>
<td>31</td>
<td>288</td>
</tr>
</tbody>
</table>

In what follows I present each type in detail.

1) *Ambiguous RCs*

The translation task confirmed that SRCs and ORCs in Ixcatec may be
ambiguous. In (24) it can be seen that the word order is the same in both the SRCs
and the ORCs; that *la*² is uninflected and thus conveys no information with respect
to the syntactic role of the RC; and that there is no verb morphology that would
allow the disambiguation of the RC. Visual inspection of waveforms in Praat does
not indicate the presence of any disambiguating prosodic cues.
Translation of a SRC
(24) a. ndi’ra² ki’i² sa¹ mi²-t[a²]
    where LOC.PRED DEF CLS-woman

    [la² ki¹=he²nga³ Øs sa¹ mi²-nda²wa²]
    COMP PROG.3SG=scratch DEF CLS-man
    ‘Where is the woman that is scratching the man?’

Translation of an ORC
b. ndi’ra² ki’i² sa¹ mi²-nda²wa²
    where LOC.PRED DEF CLS-man

    [la² ki¹=he²nga³ sa¹ mi²-t[a² Øo]
    COMP PROG.3SG=scratch DEF CLS-woman
    ‘Where is the man that the woman is scratching?’

One notes that in (24) the object and the transitive subject in the RCs are both post-verbal. This means that, although the word order in the transitive main clauses is SVO, transitive RCs are more conservative and have retained a VS(O) word order.

2) Word order change in the ORCs
The translation task revealed that speakers may use a word-order change within the RC that allows them to disambiguate between the two types of clauses. Indeed, a number of ORCs were realized with a marked word order SV in the RC. Compare the VS order in (25a) and (26a) with the SV order in (25b) and (26b).

Translation of a SRC
(25) a. ndi’ra² ki’i² sa¹ mi²-t[a²
    where LOC.PRED DEF CLS-woman

    [la² te²ngi²e² Øs sa¹ k’wa²-ʔi¹]
    COMP follow DEF CLF.F-little
    ‘Where is the woman that follows the girl?’

Translation of an ORC
b. ndi’ra² ki’i² sa¹ k’wa²-ʔi¹
    where LOC.PRED DEF CLF.F-little

    [la² sa¹ mi²-t[a² te²ngi²e² Øo]
    COMP DEF CLS-woman follow
    ‘Where is the girl that the woman follows?’

Translation of a SRC
(26) a. ndi’ra² ki’i² sa¹ mi²-nda²wa²
    where LOC.PRED DEF CLS-man
Translation of an ORC
b. ndi’ra² ki’i² sa¹ li²-ʔĩ¹
where LOC.PRED DEF CLS-little

[la² sa¹ mi²-nda²wa² ki¹=βa²tha¹ Øo]
COMP DEF CLS-man PROG.3SG=cover
‘Where is the girl that the man is covering?’

3) Cross-reference morphemes in the ORCs

As shown in (27b), marked word order in the ORCs is accompanied by the use of a cross-reference morpheme on the verb when the noun phrase allows for it. For example, the noun ‘man’ triggers the use of the cross-reference morpheme ·da² on the verb ‘follow’ because the noun is in the preverbal, marked position within the RC. It is clear that the cross-reference morpheme ·da² is coreferential to the agent of the RC, ‘man’, and not to the head of the RC ‘girl’, in which case the female cross-reference morpheme ·kwa² would have been used.

Translation of a SRC
(27) a. ndi’ra² ki’i² sa¹ li²-nda²wa²
where LOC.PRED DEF CLS-man

[la² te²ngi²ʔe² Øs sa¹ kwa²-ʔĩ¹]
COMP follow DEF CLS-little
‘Where is the man that follows the girl?’

Translation of an ORC
b. ndi’ra² ki’i² sa¹ kwa²-ʔĩ¹
where LOC.PRED DEF CLS-little

[la² sa¹ li²-nda²wa² te²ngi²ʔe²-da² Øo]
COMP DEF CLS-man fol follow·CO.3SG.M
‘Where is the girl that the man follows?’

4) Cross-reference morphemes in SRCs as resumptive pronouns

In ten responses, cross-reference morphemes were used as resumptive pronouns, agreeing with the head of the RC. This is apparent in the intransitive clauses, as illustrated in (28), where the female cross-reference morpheme ·kwa² corefers to the head of the RC ‘girl’.

(28) ndi’ra² ki’i² sa¹ kwa²-ʔĩ¹ [la² ki¹=ϕe¹-kwa²]
where LOC.PRED DEF CLS-little COMP PROG.3SG=sleep·CO.3SG.F
‘Where is the girl that is sleeping?’
In the transitive clauses, speakers PSG and RRM used the cross-reference morphemes as resumptive pronouns solely in SRCs, as in (29a), but not in ORCs, as in (29b).

Translation of a SRC
(29) a. ndi’ra² ki’i²i² sa¹ kwa² mo²re’na²
where LOC.PRED DEF woman brunette

\[[la² ki¹=tsu²te’n ga²-kwa² sa¹ mi²-ʃa² sk-?e² ru’wa¹]
COMP PROG.3SG=kick DEF CLS- head- white
CO.3SG.F woman POSS.3SG

‘Where is the brunette that is kicking the blonde?’

Translation of an ORC
b. ndi’ra² ki’i²i² sa¹ kwa² Sk-?e² ru’wa¹
where LOC.PRED DEF woman head-POSS.3SG white

\[[la² ki¹=tsu²te’n ga² sa¹ kwa² mo²re’na² Ī₀]
COMP PROG.3SG=kick DEF woman brunette

‘Where is the blonde that the brunette is kicking?’

The cross-reference morphemes used as resumptive pronouns could thus serve as a cue to disambiguate between SRCs and ORCs. However, this distinction is not systematic as speaker PSG used the cross-reference morpheme -kwa² as a resumptive pronoun for both a SRC, in (30a), and an ORC, in (30b), despite absence of marked word order in the latter that could be responsible for the use of the cross-reference morpheme.

Translation of a SRC
(30) a. ndi’ra² ki’i²i² sa¹ kwa²-ʃi¹
where LOC.PRED DEF CLF.F-little

\[[la² ki¹=ka²de’ngi²-kwa² sa¹ mi²-tʃ[ai²]
COMP PROG.3SG=pull-CO.3SG.F DEF CLS-woman
‘Where is the girl that is pulling the woman?’

Translation of an ORC
b. ndi’ra² ki’i²i² sa¹ mi²-tʃ[ai²
where LOC.PRED DEF CLS-woman

\[[la² ki¹=ka²de’ngi²-kwa² sa¹ kwa²-ʃi¹]
COMP PROG.3SG=pull-CO.3SG.F DEF CLF.F-little
‘Where is the woman that the girl is pulling?’

Finally, one notes that the two other speakers, JSB and CRG, did not use the cross-reference morphemes as resumptive pronouns for either SRCs or ORCs.
In conclusion, despite variability in the responses of the last Ixcatec speakers, it appears that Ixcatec cross-reference morphemes are preferred in SRCs and can be used as resumptive pronouns coreferring to the head NP of the RC.

4.2.6 Discussion

The results of the translation task show that ambiguous RCs were the default option. This result goes against the prediction expressed in MacDonald (2013) that producers try to minimize comprehension burden by using the following three biases: easy first (e.g., given information), plan reuse (i.e., syntactic priming), and interference reduction (e.g., choosing different syntactic structures in order to disambiguate between similar participants). In keeping with the last principle, the Ixcatec speakers should have favored unambiguous syntactic constructions in production as this would have distinguished between participants equal in animacy, e.g., a man and a woman. More specifically, a switch from the basic VS order to an unambiguous SV order should have been favored in the RCs as to allow unambiguous ORC interpretation but it is noteworthy that this was not the case.

Let us now examine the possible influence of Spanish on Ixcatec RCs. Spanish RCs are formed with the gap strategy and the complementizer que for subjects, see (31a), and for direct objects, see (31b). In terms of word order, notice that the object in the SRC follows the verb, whereas the transitive subject in the ORC can follow the verb or precede it like in example (31b).

<table>
<thead>
<tr>
<th>Spanish</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(31) SRC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. ¿Dónde está la niña [que Øs está jalando a la mujer]?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>where is the girl that is pulling at the woman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Where is the girl that is pulling the woman?’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ¿Dónde está la mujer [que la niña está jalando Øo]?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>where is the woman that the girl is pulling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Where is the woman that the girl is pulling?’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is important to note the similarity between the Ixcatec and the Spanish ORCs: both languages have externally-headed, postnominal RCs, both use the gap strategy and a complementizer. Moreover, the use of the preverbal position for a transitive subject, as observed in some Ixcatec ORCs, parallels the word order in Spanish ORCs. The results in the translation task could have therefore been influenced by cross-language transfer and more specifically syntactic priming, leading the Ixcatec speakers to use an increased number of ORCs with a marked word order.

Spanish also uses relative pronouns to relativize direct objects, as shown in (32). This relativization strategy applies to all the lower positions of the Accessibility Hierarchy, such as indirect objects, obliques, and adjuncts. The use of relative pronouns, however, is not encountered in the Ixcatec data.
Spanish, ORC
(32) ¿Dónde está la mujer [a quien la niña está jalando]?
where is the woman at whom the girl is pulling
‘Where is the woman that the girl is pulling?’

Lastly, Suñer (1998) describes a relativization strategy typical of informal Spanish that relies on the complementizer que and a resumptive pronoun. It is used to relativize subjects and all the lower positions; see (33) for an ORC.

Spanish, ORC
(33) Una cierta senadora [que Luis la llamó].
a certain senator that Luis her called
‘A certain senator whom/that Luis called.’ (Suñer 1998: 337)

Despite the fact that both Spanish and Ixcatec use resumptive pronouns, it is difficult to draw a strong parallel between the Spanish and the Ixcatec resumptive strategies as the Ixcatec resumptive strategy involves cross-reference morphemes which have different syntactic constraints than the Spanish pronouns.

5 Study 1: Comprehension experiment on subject and object relative clauses

5.1 Goals and predictions
The translation task presented in Section 4 demonstrates the importance of ambiguous RCs in Ixcatec. A picture-matching comprehension experiment was designed in order to examine how these ambiguous RCs would be interpreted by the participants when semantic and pragmatic factors that help disambiguate RCs in natural speech are minimized. In accordance with universal subject preference, Ixcatec participants were expected to strongly favour subject interpretation and to respond faster for SRC interpretations than for ORC interpretations.

The translation task, presented in Section 4, revealed that word-order change could be mobilized in ORCs resulting in an AV order. The change in word order sometimes also triggered the use of cross-reference morphemes on the verb. The picture-matching experiment aimed at testing the efficiency of this strategy in disambiguating between a transitive SRC and an ORC. For the unambiguous ORCs with an AV order Ixcatecs were expected to correctly interpret them as ORCs for two reasons:

1) The Ixcatec RCs with an AV order parallel the Ixcatec main clauses in which transitive subjects are restricted to the preverbal position (AV).
2) The Ixcatec RCs with an AV order parallel the Spanish ORCs.

5.2 Materials
The visual stimuli in this experiment were the same as the stimuli used for the translation task, presented in Section 4. In particular, 36 visual stimuli were created, each consisting of a pair of pictures. The stimuli were printed in colour, in A4 size, and were presented next to each other. Following Polinsky (2008) and Clemens et al. (2015), each pair of pictures depicted reversible actions by agents equal in animacy and only one of them was the target picture. Figure 1 shows a
pair of pictures with inanimate objects: in the picture on the left, the fish is behind/following the rooster; in the picture on the right, the rooster is behind/following the fish.

72 audio files were associated to each pair of pictures, 36 for SRCs and 36 for ORCs as parts of indirect commands involving nominal heads (Polinsky 2008). For example, two audio files were associated to the pair of pictures shown in Figure 1: one for an intended SRC, as in (34a), and one for an intended ORC, as in (34b). In this example, both the SRC and the ORC correspond to the picture on the left.

**Intended meaning: SRC**

\[(34)\] a. ndi\textsuperscript{a}ra\textsuperscript{2} ki\textsuperscript{2}i\textsuperscript{2} sa\textsuperscript{1} ?u\textsuperscript{2}-t[e]e\textsuperscript{3}a\textsuperscript{3}

where LOC.PRED DEF CLF.AN-fish

[la\textsuperscript{2} te\textsuperscript{2}ngi\textsuperscript{2}e\textsuperscript{2} O\textsuperscript{8} sa\textsuperscript{1} ?u\textsuperscript{1}-ku\textsuperscript{2}ra\textsuperscript{2}t[\text{\ti}]\textsuperscript{2}]

COMP follow DEF CLF.AN-rooster

‘Where is the fish that is following the rooster?’

**Intended meaning: ORC**

b. ndi\textsuperscript{a}ra\textsuperscript{2} ki\textsuperscript{2}i\textsuperscript{2} sa\textsuperscript{1} ?u\textsuperscript{1}-ku\textsuperscript{2}ra\textsuperscript{2}t[\text{\ti}]\textsuperscript{2}

where LOC.PRED DEF CLF.AN-rooster

[la\textsuperscript{2} te\textsuperscript{2}ngi\textsuperscript{2}e\textsuperscript{2} sa\textsuperscript{1} ?u\textsuperscript{2}-t[e]e\textsuperscript{3}a\textsuperscript{3} O\textsuperscript{0}]

COMP follow DEF CLF.AN-fish

‘Where is the rooster that the fish is following?’

Figure 1. Visual stimuli used in the sentence-picture matching task with corresponding audio stimuli: ‘Where is the fish that is following the rooster?’ and ‘Where is the rooster that the fish is following?’

The acoustic stimuli were based on the responses in the translation task given by the eldest male speaker of Ixcatec, PSG, age 89; see Section 4. In total, 48 RCs were morphologically and syntactically ambiguous, 14 ORCs had a marked SV order in the RC, and 4 RCs had cross-reference morphemes that clearly differentiated between the agent and patient in the RC, whereas 6 files were discarded for having disfluencies (total N=72).

In addition, 14 pairs of fillers were used, a number that was kept voluntarily low for feasibility reasons. In each pair the fillers depicted a single agent who realized different actions in each picture. The visual stimuli were associated to intransitive SRCs, e.g., “where is the girl that is eating?” vs. “where is the girl that is drinking?”.
5.3 Participants
The four most fluent speakers of Ixcatec participated in this study, similar to the translation task presented in Section 4. The comprehension study was also conducted with three semi-speakers, two female and one male, all in their 70s, i.e., ten years younger than the fluent speakers. All of them learned Ixcatec from their families, but more likely had less exposure to it than the fluent Ixcatec speakers who were born ten years earlier. When contacted within the language documentation programme (2010-2013), it appeared that although the semi-speakers could understand the Ixcatec fluent speakers, they were not themselves fluent and had difficulties in participating in Ixcatec conversations or producing narratives. The comprehension experiment offered the perfect setting to include the semi-speakers and thus practically double the sample of the Ixcatec participants. All participants received financial compensation for their collaboration in this study.

5.4 Procedure
The experiment was conducted with each participant individually and took place at the participants’ homes at Santa María Ixcatlán in September 2015. There were four sessions for each participant: two sessions were conducted during the same day, with a small break of approximately five minutes between each session; the other two sessions were organized on a different day depending on the participants’ availability, generally two days later. During one session, each participant completed 50 trials corresponding to the comprehension of the 36 pairs of pictures associated to the audio stimuli of either a SRC or an ORC, and the 14 pairs of fillers associated to intransitive SRCs. In total, all participants saw all 36 pairs of pictures twice and listened to the corresponding SRCs and ORCs. Six participants provided responses to 144 transitive RCs, but, for practical reasons, one of the semi-speakers could not carry out the repeat session and therefore provided only 72 responses. In total, the experiment provided 936 responses.

For each trial, participants saw two pictures that were placed one next to the other on a table or a bench. The pairs of pictures were presented in a randomized order, generated through a computer for the four sessions. This computer-generated order was different for each participant by alternating the order of the four sessions and by manipulating the order of appearance of the pictures within each session (e.g., from last to first, from first to last).

The participants were instructed in Spanish⁶ to point at the picture best answering the question heard in Ixcatec; see Figure 2. They were first familiarized with the visual stimuli to make sure they encountered no difficulties. Participants listened to the audio stimuli either through headphones or through audio players, depending on which medium was best for them given hearing difficulties due to their age. The presentation of the audio stimuli followed the visual stimuli by approximately 10 seconds. Each session lasted from 10 to 20 minutes. The sessions were filmed.

---

⁶Ixcatec is not a language of everyday communication and the last Ixcatec speakers interact with each other and with the researchers in Spanish. It was therefore established that instructions would be given in Spanish.
Note that the fluent speaker PSG was familiar with the visual stimuli from the translation task which had been conducted a few days earlier. As a result, PSG listened to his own voice during the comprehension experiment. The three other fluent speakers conducted the comprehension experiment a few days before responding to the translation task presented in Section 4. The semi-speakers did not respond to the translation task. Participants declared having no difficulties with the pictures. They also said that PSG’s Ixcatec was very good and that they had no difficulties in understanding the audio stimuli.

Figure 2. Participants point at the appropriate picture

4.5 Coding and analysis
The answers were coded in an Excel file for each participant, indicating type of RC (ambiguous, word-order change, cross-reference morpheme), intended meaning (subject or object), preference for the ambiguous RCs (subject or object), accuracy for the unambiguous RCs and with respect to the intended meaning for the ambiguous RCs (correct or wrong), and response times in milliseconds from the beginning of the sound file up until the moment the participant had touched the selected picture (considered as the equivalent of a button press based on the video recording); see Figure 2.

In the analysis, three pairs of clauses from the audio stimuli were considered as non-applicable for including disfluent segments and the responses to this stimuli were not taken into account; i.e., from the 936 responses, 78 responses were removed. The comprehension experiment thus provided responses for 858 transitive RCs in total.

638 audio stimuli were coded as “ambiguous”, i.e., they had no syntactic or morphological marking (417 ambiguous SRCs and ORCs, and 221 ambiguous SRCs, but not their ORC counterparts which had a word-order change or a cross-reference morpheme). The ambiguous clauses were also coded with respect to the intended meaning in the translation task. 170 ORCs were coded for “word-order change” and 50 for “cross-reference morphemes”.

To analyse the reaction times for the ambiguous RCs, linear mixed models (lmer) were constructed using the “lme4” package (Bates et al. 2014) in R (R Core Team 2013). The dependent variable was the “reaction time”, and the independent variable was “type” (SRC or ORC). “Subject” and “audio file length” were coded as random factors.
5.5 Results
In accordance with our predictions for the syntactically and morphologically ambiguous RCs, the SRC readings (N=401 or 63%) were significantly more frequent than the ORC readings (N=237 or 37%); \( p<0.0001 \). See Figure 3.

Figure 3. Preferred interpretation for the ambiguous RCs in the sentence-picture matching task (total N=638)

Figure 4 presents the results for the ambiguous RCs by taking into consideration the intended meaning during production (in the translation task). This count is a methodological precaution as Ixcatec is an under-studied language and it is important to make sure that there was no marking that would have gone unnoticed. Similar to the results for preferred interpretation presented in Figure 3, this analysis shows that participants correctly identified 283 SRCs (out of 430 SRCs or 66% correct responses) and 90 ORCs (out of 208 ORCs or 43% correct responses); \( p<0.0001 \).
The analysis of the ORCs with an SV, marked word order (N=220) indicates that the Ixcatecs interpreted these clauses at chance level ($p=0.088$): 55% of the clauses (N=120) were correctly interpreted and 45% of the clauses (N=100) were interpreted incorrectly. More specifically, as can be seen in Figure 5, the marked SV order in the RC is not a sufficient cue as only 52% of the ORCs were correctly interpreted (N=88). When, however, the word-order change triggered the use of a cross-reference morpheme and there was a gender difference, the recognition of the ORCs improved, with a 65% success rate (N=32).

Figure 4. Accuracy of comprehension of the ambiguous RCs in the sentence-picture matching task when taking into consideration the intended meaning (total N=638)

Figure 5. Accuracy of comprehension of subject and object RCs in the sentence-picture matching task: Results for object RCs with word-order change and cross-reference morphemes (total N=220)
Let us now turn to the accuracy rates per speaker, illustrated in Figure 6. I expected that the semi-speakers would show lower levels of accuracy than the fluent speakers. Nevertheless it can be seen that semi-speakers GHG and MSG scored similarly to the other three fluent speakers, while speaker RRM scored below 50% of correct answers. This indicates that fluency did not affect the comprehension experiment although semi-speaker CRS had the lowest accuracy.

![Figure 6. Accuracy of comprehension of the non-ambiguous ORCs in the sentence-picture matching task (total N=220): Results per speaker (identified by gender, age, and initials)]](image)

Reaction times were counted for all participants but semi-speaker CRS, who was excluded because of the low accuracy rate. Statistical analysis shows that there are no significant differences between reaction times for a SRC interpretation and an ORC interpretation (t<1), although SRCs are numerically faster than ORCs, i.e., 6799 ms vs. 6874 ms (see Table 5).

<table>
<thead>
<tr>
<th>Ambiguous RCs</th>
<th>Mean RTs (sd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject interpretation</td>
<td>6799 ms (2019)</td>
</tr>
<tr>
<td>Object interpretation</td>
<td>6874 ms (4103)</td>
</tr>
</tbody>
</table>

5.6 Discussion
The first research question was how morphologically and syntactically ambiguous RCs are interpreted and processed by the last Ixtec speakers. The second research question was whether the syntactic and morphological cues in ORCs help comprehenders to resolve ambiguity.

With respect to the first question, analysis of the Ixtec data shows that when the RCs are ambiguous, participants tend to interpret them as SRCs. Similarly, when taking into consideration the intended meaning, it was noted that ambiguous SRCs were more frequently identified correctly than ORCs. These findings are
consistent with universal subject preference (Keenan & Comrie 1977). Results for reaction times indicate that SRC interpretations are not significantly faster than ORC interpretations although SRC interpretations are numerically faster than ORC interpretations. This result should be viewed with caution as the sample is small. Absence of a subject processing advantage contrasts with results from processing of fully ambiguous RCs in Mayan languages (Clemens et al. 2015). An interpretation of this result could be that when confronted to ambiguous RCs, the Ixcatec comprehenders engage with either a SRC or an ORC interpretation. In the absence of disambiguating cues, they do not re-evaluate their initial interpretation, something which is known to cause a difference in RC-processing in languages with temporary ambiguous RCs. Frequency of use in a free-speech corpus should help shed some light on the preferred interpretations and the reactions times; see Study 2 in Section 6.

With respect to the second research question, the results show that the use of marked SV word order in itself did not help participants correctly interpret ORCs despite congruence between word order in the ORCs and word order in the transitive main clauses. This may indicate a recent innovation in the word order of the ORCs following the change in the word order of the transitive main clauses. The failure to comprehend ORCs with an SV order could also support the hypothesis that cross-language transfer from Spanish played a role during the translation task. Indeed, when going back to the answers during the translation task in section 4.2, it can be seen that only the two male speakers relied on word order changes. When the change in word order in the RC triggered a cross-reference morpheme distinguishing participants by gender, then the ORCs were interpreted with greater success.

Finally, the low accuracy in RC comprehension for ambiguous and non-ambiguous RCs can be understood as confirming absence of prosodic cues. It is indeed noteworthy that speaker PSG who produced the audio stimuli did not interpret his own productions any better than the other participants. More generally, low recognition scores have been reported for the study of RCs in two Mayan languages by Clemens et al. (2015), especially among the monolingual participants. Low accuracy and slow reaction times are linked by Clemens et al. (2015) to the characteristics of the less-educated, monolingual speakers who have little experience with tests and computers.

6 Study 2: Ambiguity resolution and relative-subject preference in the free-speech corpus

6.1 Goals and predictions
Study 2 presents the analysis of a free-speech corpus of Ixcatec. First, a qualitative look at the morphologically and syntactically ambiguous RCs sheds light on the disambiguating role of semantic and pragmatic cues. Second, a quantitative analysis of frequency of SRCs and ORCs allows to examine whether a frequency effect underlies the SRC interpretation that was favoured in the comprehension experiment in Study 1 presented in Section 5. In agreement with other studies on RCs in spoken and written corpora, we expect SRCs to be more frequent. However, Gordon & Hendrick (2005) suggest that even though transitive and intransitive subjects share the same high structural position, it is important to distinguish
between them when discussing corpus frequencies. Such fine-grained metrics of RCs show that transitive SRCs are not always more frequent than ORCs in English corpora (Gordon & Hendrick 2005).

6.2 Corpus
From 2010 to 2013, a free-speech Ixcatec corpus was collected within the framework of the Ixcatec Language Documentation programme. 50 hours of video recordings are now available on the Endangered Languages Archive (ELAR) http://www.elar-archive.org/index.php. Of these, three hours were quantitatively analyzed for this study.

6.3 Participants
The four most fluent speakers of Ixcatec participated in this study. The two 80-year-old female speakers were brought together to the yard of one of the speakers during several sessions. The female speakers, who are close friends, vividly discussed their everyday life in Ixcatec while being filmed by the researchers. Their conversations indicate fluency as they show frequent overlaps and back-channelling as well as an average 220 ms turn-taking, consistent with cross-linguistic patterns of dyadic informal conversations (Adamou 2016: 10). The two 80-year-old men, who are family members, were brought together for the Ixcatec sessions. They would generally discuss local traditions and recollections. Contrary to the female-to-female conversations, the male-to-male conversations showed few overlaps and back-channelling and slow turn-taking, i.e., 1000 ms in average (Adamou 2016: 10). However, these characteristics are not interpreted as signs of lesser fluency, rather of more formal communication.

The Ixcatec speakers agreed individually to participating in this study by signing formal authorizations after each working session and were financially compensated for their time. More significantly, the Ixcatec documentation research programme was approved by the community’s general assembly.

6.4 Corpus annotation
For the corpus annotation I used the ELAN format from the Max Planck Institute for Psycholinguistics (Nijmigen, Netherlands). First, the texts were annotated in prosodic units. Then using the Elan CorpA version, a tx tier was created for the broad phonetic transcription using the IPA and an ft tier was used for broad translation. Moreover, a tier mot was used for words and was automatically tokenized into morphemes (tier mb), which were glossed in the ge tier following the Leipzig glossing rules. The annotated texts were synchronized with the audio and video files. Following Gordon & Hendrick (2005), RCs were identified and coded as intransitive SRCs (SRCs), transitive SRCs (ARCs), and object RCs (ORCs). Examples that were unclear were discarded from the analysis.

6.5 Results
The analysis of the free-speech Ixcatec corpus revealed a high number of RCs in which the nominal head is not overtly expressed (Lehmann 1984). See an example of an agent in (35), a direct object in (36), and a locative in (37).
Let us now examine how ambiguity is lifted in natural conversations. For example, although the RC in (35) is syntactically and morphologically ambiguous, interpreting it as a SRC is more likely for semantic and pragmatic reasons linking the verb and the NP in the RC. Indeed, the RC includes the animate participants ‘children’ following the verb ‘to teach’. Two interpretations are possible: in a SRC ‘children’ would be the object, in an ORC ‘children’ would be the transitive subject. A SRC interpretation is favoured by the pragmatic context and possibly the higher frequency of a construction where the verb ‘to teach’ is followed by an object ‘children’ than a construction where ‘children’ is the subject of the verb ‘to teach’.

Example (36) is also straightforward as the verb of the matrix clause, ‘buy’, has the first person agent suffix (verbs take only S/A indexing) and requires an object. Although the predicate of the RC, ‘have’, has no person marking, the noun, ‘hand’, has a first person possessive suffix and thus favours the interpretation of a first person agent for the predicate ‘have’ with a gap in lieu of object. This type of ORCs involving an inanimate head NP and an animate NP in the RC are known to be easier to understand than ORCs involving an animate head NP and an inanimate NP in the RC (see Traxler et al. 2002).

Finally, in (37), the verb of the RC, ‘enter’, has a first person inclusive suffix for the agent, and requires a locative complement, understood as being the gap in the RC. RCs involving first and second persons are also known to be easier to process when referring to accessible, unambiguous participants.

For the quantitative analysis I considered all the RCs that are formed with the gap strategy and the complementizer $la^2$, including RCs with a nominal head and RCs with a null nominal head. The quantitative analysis of three hours of the Ixcaté free-speech corpus shows 64% intransitive SRCs (N=112), 19% transitive SRCs (N=34), and 17% ORCs (N=30); see Table 6.
Table 6. Frequency of relativized intransitive subjects, transitive subjects, and objects in a three-hour, free-speech corpus of Ixcatec

<table>
<thead>
<tr>
<th></th>
<th>Intransitive SRCs</th>
<th>Transitive SRCs</th>
<th>ORCs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>112 (64%)</td>
<td>34 (19%)</td>
<td>30 (17%)</td>
<td>176 (100%)</td>
</tr>
</tbody>
</table>

6.6 Discussion

A qualitative analysis of the Ixcatec corpus illustrates how ambiguity is resolved in natural conversations for RCs involving first or second person participants, participants that are not equal in animacy, and when specific constructions favour one interpretation over the other. A quantitative analysis of the corpus shows that RCs with intransitive and transitive subjects are the most frequent, confirming universal subject preference. However, a finer grained analysis of relativized transitive subjects and objects does not indicate any clear subject preference.

These results can be compared to the results from a smaller corpus analysis from Ixcatec (Adamou 2016: 168–169). Analysis of main clauses shows that overt NPs are rare (178 NPs for 648 verbs), of which 46% are intransitive subjects (S), 10% transitive subjects (A), and 44% direct objects (O), see Table 7. Comparison with the results from the RC-uses presented in this paper confirms the prevalence of intransitive subjects (S) in both main clauses and RCs, respectively 46% and 64%. Relative subject preference is also apparent when we compare relativized objects, which are the least frequent type (17%), and overtly expressed objects in main clauses (44%), which are as frequent as intransitive subjects (46%) and more frequent than transitive subjects (10%). We can conclude that objects in Ixcatec are frequently expressed overtly in spontaneous speech but are rarely relativized.

Table 7. Frequency of core arguments in main verbal clauses (adapted from Adamou 2016: 168–169)

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>A</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single prosodic unit</td>
<td>75 (51%)</td>
<td>14 (9%)</td>
<td>59 (40%)</td>
</tr>
<tr>
<td>Dislocations</td>
<td>7 (23%)</td>
<td>4 (13%)</td>
<td>19 (64%)</td>
</tr>
<tr>
<td>Total</td>
<td>82 (46%)</td>
<td>18 (10%)</td>
<td>78 (44%)</td>
</tr>
</tbody>
</table>

7 General discussion and conclusion

This paper presents the first analysis of RCs combining natural and experimental data for Ixcatec, an under-described and critically endangered language of Mexico. It represents, to the best of my knowledge, the only experimental study of RCs in any Otomanguean language. Besides the descriptive value of this study, I suggested that the syntactic and morphological characteristics of Ixcatec allow one to tease apart confounding factors to universal subject preference.

Taken together, the findings from the experimental and the corpus study provide mixed results regarding universal subject preference. In accordance with subject preference, the comprehension experiment (Study 1) showed that ambiguous RCs are more frequently interpreted as SRCs than as ORCs. The same result was obtained when the intended meaning was taken into consideration: SRCs were more accurately identified than ORCs. However, although reaction times were numerically faster for SRCs, there was no statistical significance for
the subject processing advantage. This lack of significance may be due to the small size of the sample. Alternatively, this could suggest that Ixcatec comprehenders do not go through a first stage of SRC interpretation before proceeding to an ORC interpretation as comprehenders in languages with temporary ambiguous clauses do. Rather, Ixcatecs may opt for an interpretation which, in the absence of syntactic or morphological cues, they do not need to re-evaluate. In order to better understand the results from the reaction times and the distribution of the RC interpretations in Study 1, a corpus analysis was carried out (Study 2). The analysis of natural speech collected during a language documentation programme showed that relativized intransitive and transitive subjects are more frequent than relativized objects, in accordance with universal subject preference, but that relativized transitive subjects are only slightly more frequent than relativized objects. The frequency of RCs in production therefore supports the assumption that, in comprehension, Ixcatecs prepare for both RC interpretations.

To conclude, this study illustrates how lesser-known, oral-tradition languages can contribute to the theoretical discussions that have mainly been addressed for well-described, written languages. In sum, although the Ixcatec data confirm universal subject preference, they also show how this preference is weaker than expected. I suggest that in order to grasp the processing mechanisms at play in universal subject preference, we should pursue our investigations by broadening our typological database.

**Abbreviations**

Glosses follow the Leipzig Glossing Rules.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Gloss</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 3</td>
<td>first, second, third person</td>
<td></td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive</td>
<td></td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>animal</td>
<td></td>
</tr>
<tr>
<td>ANT</td>
<td>anterior</td>
<td></td>
</tr>
<tr>
<td>APPL</td>
<td>applicative</td>
<td></td>
</tr>
<tr>
<td>CLF</td>
<td>classifier</td>
<td></td>
</tr>
<tr>
<td>CLS</td>
<td>class</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>cross-reference</td>
<td></td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
<td></td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
<td></td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
<td></td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
<td></td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
<td></td>
</tr>
<tr>
<td>EXS</td>
<td>existential</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>feminine/female</td>
<td></td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
<td></td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
<td></td>
</tr>
<tr>
<td>HON</td>
<td>honorific</td>
<td></td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
<td></td>
</tr>
<tr>
<td>QUEST</td>
<td>interrogative particle</td>
<td></td>
</tr>
<tr>
<td>INS</td>
<td>instrumental</td>
<td></td>
</tr>
<tr>
<td>IPFV</td>
<td>imperfective</td>
<td></td>
</tr>
<tr>
<td>ITER</td>
<td>iterative</td>
<td></td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>masculine/male</td>
<td></td>
</tr>
<tr>
<td>NEG</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>proper noun</td>
<td></td>
</tr>
<tr>
<td>PFV</td>
<td>perective</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
<td></td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
<td></td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
<td></td>
</tr>
<tr>
<td>PRST</td>
<td>presentative</td>
<td></td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
<td></td>
</tr>
<tr>
<td>REL</td>
<td>relativizer</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
<td></td>
</tr>
<tr>
<td>SUB</td>
<td>subordinator</td>
<td></td>
</tr>
</tbody>
</table>
References


Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz-Pavía & Itziar Laka. 2010. Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115(1). 79–92.


Hironymous, Michael. 2007. *Santa Maria Ixcatlan, Oaxaca: From colonial Cacicazgo to modern Municipio*. Austin, TX: University of Texas PhD dissertation.


