UNIVERSIDAD DE SONORA

División de Humanidades y Bellas Artes

Maestría en Lingüística

Referential Density in Tarahumara

(Densidad Referencial en Tarahumara)

TESIS

Que para optar por el grado de

Maestro en Lingüística

presenta

Ana Elia Tona Messina

2015

ABBREVIATIONS

1	first person	IRR	irrealis
2	second person	LOC	locative
3	third person	MN	manner
A	agent-like argument of canonical transitive verb	NMLZ	nominalizer/nominalization
		OBL	oblique
APPL	applicative	P	patient-like argument of canonical transitive verb
AUX	auxiliary		
BEN	benefactive	PASS	passive
CAUS	causative	PFV	perfective
CL	clausal marker	PL	plural
COMIT	comitative	POSS	possessive
COND	conditional	POT	potential
CONJ	conjunction	PROG	progressive
CONT	continuative	PRS	present
DAT	dative	PST	past
DEM	demonstrative	REC	recipient
DES	desiderative	REP	reportative
DIR	directional	RES	resultative
DP	discursive particle	S	single argument of canonical intransitive verb
EMPH	emphatic		
EP	epistemic	SG	singular
HAB	habitual	SIM	simultaneity
IMPRS	impersonal	STAT	stative
INDEF	indefinite	SUB	subordinator
INSTR	instrumental	TM	temporal location
INT	intensifier		

INTRODUCTION

This work constitutes a statistical study concerned with the behavior of referential elements in Tarahumara. It retrieves its main concepts from typological studies and functional linguistic theory, while trying to touch upon their concerns. Accordingly, this study intends to address the need for more discourse studies in indigenous languages, mainly in the area of referentiality, referential mechanisms and information structure.

It is, for the most part, based on the Referentiality Project started by Christian Lehmann in 2012 at the University of Erfurt in Germany. This research project was envisioned towards a crosslinguistic text-based analysis of 12 unrelated, endangered languages that had nothing, but one thing in common: none of them had received much attention from researchers in general. Researchers that worked together in this project used data from The Language Archive (TLA). The 12 Languages studied as part of the Referentiality Project were: Aché (Tupi-Guarani, Paraguay), Baure (Arawakan, Bolivia), Beaver (Athabaskan, Canada), Chintang (Kiranti, Nepal), Gorani (Iranian, Iran), Hocak (Siouan, U.S.A.), Isubu (Bantu, Cameroon), Savosavo (East Papuan, Solomon Islands), Sri Lanka Malay (Austronesian, India), Teop (Austronesian, Bougainville), Vera'a (Austronesian, Vanuatu), Wichita (Caddoan, U.S.A.).

The aim of the Referentiality Project was not only to establish a protocol for cross-corpus analysis for future research, but also to reach a better understanding of referential operations in the communicative process. Additionally, the Referentiality Project had the purpose of demonstrating the importance of unknown languages in theoretical linguistics, as they have the potential to "change our views of how languages function" (Lehmann 2012). Hence, Lehmann´s research project was not aimed at simply organizing data or testing existing hypotheses on these

languages. Rather, the Referentiality Project was resolved to collect new or undocumented referential strategies and grammatical categories from unknown languages; it was open to language-specific categories that would later be mapped on cross-linguistic concepts, which can lead to novel research questions and hypothesis.

One of the most significant features of the Referentiality Project is that it only used natural spoken language data to draw statistical information for crosslinguistic comparison. To attain its purposes, the project focused on certain text genres, mainly narrative, procedural and interactive texts, as they typically present different discourse properties.

The main features and general purposes of the Referentiality Project are recaptured in this work: only natural, non-elicited material was used as data for linguistic analysis, and openness to language-specific structures and referential strategies is maintained. However, this is not a cross-linguistic analysis; instead, two texts of the same language were analyzed in this work, a narrative and a procedural text, as they feature a greater range of phenomena of the language in question. Thus, instead of comparing different text genres from different languages, this work makes a comparative analysis of different text genres of the same language.

The purpose of comparing two text genres of the same language is to reveal how diverse referential mechanisms can be within the same language, depending on the different communicative purposes which are embodied in different text genres. While it is understandable that it is impossible for a time-limited research project such as the Referentiality Project to look into every text genre of every language, individual researchers may well make an effort to collaborate with their own works. Ideally, linguistic theory and typological studies would take into consideration as many and as varied languages as possible, along with deeper studies of their structures and mechanisms. However, typologists cannot do all the work on their own, and many

times they have to make their hypothesis relying solely on the available data. It is descriptivists'
work to make available as much (processed) data as possible, so that linguistic research can
move forward.

With that being said, this work intends to contribute to the enrichment of typological
studies, providing a glimpse into the referential properties of two different types of discursive
texts in Tarahumara. It seeks to answer some questions about the Tarahumara language through
statistical values concerning referentiality (that is, Referential density):

i.   What does Referential Density say about informativeness in Tarahumara?

ii. How is referential information managed in Tarahumara?

iii. How does information structure relate to discourse genre?

Because discourse studies are such a complex subject matter, a wide range of concepts were
reviewed, and various tools were used in this work, leading to (a) statistical work to account for
referential density and information structure, (b) syntactic analysis for reference identification
and tracking, (c) semantic analysis to observe the level of informativeness of references, (d)
exploration of text genre to allow for predictions and have a better understanding of referential
behavior, and (e) examination of cultural and syntactic aspects that play a role in inference for
referential identification and maintenance.

While doing the analysis of two discourse texts for this work, some difficulties related to the
methodology of the identification and computation of references came across, resulting in a
couple of proposals that may help overcome such inconveniences for future works of the same
kind.

1. Work structure

The information needed to understand this work is organized logically in such a way that one concept leads to the next one. Chapter 1 begins with a brief account of some previous studies concerned with referentiality. Notice that it does not provide a historical overview of referential studies, nor does it enlist a large number of all authors or important articles that revolve about this subject. It rather looks at the variety of research questions and linguistic aspects that can be addressed by means of referential studies. These antecedents are followed by the theoretical framework which served as a base for the conception of this study. The theoretical framework incorporates all concepts that make part of this referential study, and explains the relationship among them. It first introduces the concept of referentiality and referential density in section 1.2, followed by informativeness and referential weight in 1.3. Section 1.4 explains the concepts of information structure, topic and focus, which gets us to discourse analysis, which is addressed in 1.5. Within discourse analysis, there is a brief portrayal of discourse/text typology; it also takes a look into the universe of discourse, which comprises referential anchorage, deixis, referential distance and inference.

Chapter 2 is a description of the subject language of this study: Tarahumara (spoken in the state of Chihuahua, Mexico). This chapter actually opens with some cultural and geographical aspects of Tarahumara which will be taken into consideration later on, as part of the analysis. Section 2.2 presents the linguistic family to which Tarahumara belongs. The language description begins in section 2.3. This description does not intend to be a thorough grammar; such descriptions can be found in other studies that have been made about this language. Here, the description includes only the linguistic features that are relevant for the understanding of referential behavior, that is, alignment type in section 2.3, which helps in the interpretation of

zero anaphora; morphological features (in 2.4), as some morphemes have referential value; valency changes that may affect the number of referents and referential mechanisms within a clause (section 2.5); nominal expressions, the main referential mechanisms, are presented in 2.6. Section 2.7 explains the usage of determiners and determiners, among other elements that may affect referential expressions in different ways.

Information about the data used in this work and the analysis are presented in chapter 3 and 4, respectively. Chapter 3 gives details about the material sources and the people who participated in the processing of the texts. Sections 3.1 and 3.2 describe the two discursive texts used, in terms of text genre. The last section of this chapter discloses some of the difficulties encountered in the interpretation process of the texts. Chapter 4 is the last chapter; it presents the actual referential analysis and statistical results. First, section 4.1 explains all the methodological details, including some problems that were encountered during the analysis of referential expressions. Then, section 4.3 shows the results for the procedural text, and section 4.4 shows those of the narrative text. Section 4.5 summarizes the outcome of the analysis. Chapter 4 is followed by the final observations and conclusions of this work. Finally, the appendix offers the two texts used for this study; they are transcriptions in Tarahumara accompanied by an English translation for easy reading.

CHAPTER I

THEORETICAL FRAMEWORK

1.1. Antecedents

In the last years referential studies have gained weight as another means for typological studies. There are works that explore strategies for referential codification in language and the discursive status of the referents encoded (Lehmann 2012). Other discursive analyses have described the relationship between reference and clausal transitivity (Nariyama 2001); it has also been associated with information structure across languages (Bickel 2003). These works are all based on analyses of reference in discursive texts.

By means of Referential Density studies, Bickel (2003) was able to observe when and how participants are introduced in discourse in the three languages of his sample (Belhare, Maithili and Nepali), which is associated to information management (or information structure), and to rhetorical styles that could be due to cultural values in the disclosure of information or individual choice (Stoll & Bickel, 2006). Additionally, the values obtained reflect the amount of information that is provided to the listener through the various explicit and implicit referential mechanisms, as opposed to what is left to the listener to infer. Bickel (2003) referred to this as information density (Lehmann 2012 uses the term *informativeness* of NPs). Moreover, referential density values (RDV) helped identify prominent syntactic arguments in discourse, which are the controllers or pivots that are subject to verb agreement, relativization, passivization and other syntactic phenomena.

Nariyama (2001) also used discursive materials in three languages (Sacapultec, Tshimshian, Acehnese) to show how argument structure can be used as a reference-tracking

system. In this study she identifies the referent of ellipsis in the clause, arguing that since the ellipted argument must be inferred by the hearer, it constitutes a reference-tracking mechanism. In addition, Nariyama (2001:122) observes that languages have preferred argument structure that induces a pattern of ellipsis, which serves as a reference-tracking of ellipted arguments. This referentiality study had interesting findings: languages of the sample taken by Nariyama tend to have only one lexical argument per clause (independently of the clause's transitivity). Nariyama (2001:118) points out a cognitive relevance of this fact: new information is cognitively more difficult to process, and since lexical arguments generally represent a new argument in discourse, the ones that are not new tend to be elliped. "It is also in line with the claim made by Hopper and Thompson (1980) about the transitivity and discourse relation, that low transitivity is correlated with backgrounding information, and high transitivity is correlated with foregrounding information" (Nariyama 2001: 122). This again is related to information structure in discourse.

Another aspect that can be observed through referentiality studies is the discourse status of referents. Lehmann (2012) maintains that different coding strategies respond to discourse status of referents: more explicit NPs tend to be the ones introducing new referents into the Universe of Discourse (creation of a referent), whereas less explicit or implicit coding strategies are used for referents that are already active and accessible for the interlocutor in discourse (endophora and anaphora).

By looking at the three examples mentioned above, we can see that referentiality studies can be used to render information about different aspects and properties of discourse in languages, such as:

a. information management (or information structure)
b. rhetorical styles
c. information density (or informativeness)

d. identification of privileged syntactic arguments

e. reference tracking

f. argument structure

g. clausal transitivity

h. coding strategies of referents

i. discourse status of referents

Of course, these are not the only reference related studies; many others have been made, as referentiality studies have proved to be a useful tool in typological studies.

1.2. Reference

Following Lehmann (2012), *referentiality* "is a property of a linguistic sign that functions in a discourse (as opposed to a linguistic sign as an element of a language system)". Linguistic elements that mention entities or referents are called referential/referring expressions or referential devices (Kibrik 2011: 5). The entities that people mention can be of different sorts, including (Kibrik 2001: 1123):

i. participants of the current speech act (*I, you*)

ii. other persons (*Mary, she*)

iii. living beings and objects (*the cat, that car*)

iv. abstract notions conceptualized as objects (*the Great French Revolution, my salary*)

v. locations in space (*right here, on the Red Square*)

vi. moments in time (*tonight, on New Year's eve*)

Referents do not have to be perceived entities of the external world; they can be brought to mind not only from perception but also from long-term memory or imagination. For a referent to be mentioned in discourse, it is not relevant whether it has an independent physical existence (Kibrik 2011).

Kibrik (2011: 7) considers the distinction between referents and events/states to be a fundamental one, and he also sets time/location as separate elements. As it is known, elements denoting time and location are commonly coded as peripheral or non-core arguments in the clause. This would restrict the study of reference to peripheral arguments of verbs only, leaving aside any mention of other entities in discourse, as if they were irrelevant in the processing of discourse.

Other authors consider this category to be broader. According to Givón (2005), "the grammar of reference cheerfully disregards denotation in the [referential world] RW, opting to align itself with denotation in the Universe of Discourse" (Givón 2005:127). Thus, referential expressions may represent any entity that is relatively stable in time, which would include any physical object, landmark, place, cultural institution or abstract concept. In even wider terms, Lehmann (2012) defines reference as the relation between the linguistic expression and an entity present in the universe of discourse. Lehmann (2012) does not ignore referential expressions that refer to peripheral arguments, or nouns that refer to spatial and temporal locations. Including these sorts of referents in the study of discourse means that all referents that are being mentioned by speakers and processed by listeners, are also being analyzed.

If we are to consider reference from a communicative/discursive perspective, it cannot be denied that there is reference and informative content even in those NPs that function as peripheral arguments or that appear in adverbial clauses. These NPs become relevant from the moment the speaker decides to include them as part of discourse. This may be explained by the principle of economy of language and by Grice's conversational maxims of relevance and quantity (Grice 1975). Speakers would not include these NPs if they were not relevant for their

communicative purposes. Based on this idea, this study considers peripheral arguments as well as core arguments of the clause, as will be noticed in the analysis.

1.2.1. Individuation and Referential Type

In discourse, speakers may refer to entities with a specific mental representation in mind, a referent with an identity (*John's car is red*), or they could refer to things as mere concepts (*Cars are not a luxury anymore*). According to Lehmann (2014), *individuation* is a linguistic operation leading from a unit of the linguistic system to a unit of discourse. Individuation involves separability and distinguishability of entities. Depending on the nature of entities, some may be taken as individuals more easily than others: a human entity, for instance, is more easily individuated compared to an inanimate object. The more specific the identity of the referent, the more individuated it is. Table (1) provides the individuation scale form low to high (based on Lehmann, 2014):

Table 1. Individuation scale		
low	*Generic reference.*	Making reference to a species comprised by the concept
	Non-specific reference.	Making reference to a subset of a certain species, without specifying the identity of the individuals.
	Specific reference.	Making reference to an individual or a set of individuals that can be identified by at least one of the interlocutors.
	Unique reference.	Making reference to an entity that is already stored in the memory of the interlocutors (encyclopedic or shared knowledge), so that the identity of the referent is known by both of them.
high		

Individuation is highly related to animacy: when people feel identified with another entity, they tend to individuate that entity in a higher degree, compared to those entities with which they feel no relation or similarity. Thus, Lehmann (2014) renames the animacy hierarchy as the *empathy* scale. Furthermore, people tend to pay more attention to and speak about those

11

entities that are higher in the animacy/empathy scale; therefore, they tend to occupy central positions in language, featuring higher in the topicality hierarchy. The animacy hierarchy and topicality hierarchy are shown below.

i.   **Animacy hierarchy** (Croft 1991): first/second person > third-person pronoun > proper name > human common noun > animate common noun > inanimate common noun
ii. **Topicality hierarchy** (Givón 1995): AGT > DAT/BEN > PAT > LOC > INSTR > OTHERS

Languages grammaticalize individuation in different ways and degrees. Higher individuation is usually attained through certain grammatical categories like determiners and definite articles, although it may be achieved by other mechanisms in languages. Lower degrees of individuation, by contrast, may be attained by indefinite articles or noun incorporation. In some languages it is possible to make constructions where no referents can be identified, like German passive constructions such as *Es wurde getrunken* 'there was drinking*'; however, an implicit referent can be inferred as a non-specific entity (Lehmann, 2014).

1.3. Referential strategies and informativeness

All languages of the world have a referential system. A system of reference is only feasible when it undertakes the semantic task, the discourse-pragmatic task and the processing task (Dooley & Levinsohn 2001). The **semantic task** is fulfilled when referents succeed to unambiguously identify the intended object from other possible ones. Referring expressions will be more specific when there is more than one possible referent. Sometimes only one argument of the clause is overtly expressed, and the others can be identified by means of contextual information and selectional restrictions. The **discourse-pragmatic task** is done by indicating a referent's activation status and its prominence. The amount of explicitness or coding material varies according to the activation status: highly active referents require less coding material. The

activation of a referent is commonly done using a full noun phrase; once activated, it will typically require only minimal coding to maintain its active status. Concerning the **processing task**, referential devices must overcome any disruptions in the flow of information. More coding material is needed whenever there are disruptions like thematic breaks in a narrative, or a change in the type of information (Dooley & Levinsohn 2001).

In general, people mention referents using nominal expressions such as nouns and pronouns; however, as Kibrik (2011) has stated, the boundaries of the category 'referent' are not clear-cut. He recognizes two basic types of referential devices or nominal elements (Kibrik 2011:37):

(i)   Full referential devices
      a. proper names
      b. common nouns (with or without modifiers), or descriptions

(ii) Reduced referential devices
      a. pronouns
      b. zero forms

Lehmann (2012) proposes a list somewhat different from Kibrik´s, resulting in five referential strategies:

Table 2: Lehmann's referential strategies.

---

a.   nominal expressions (lexical NP, free/clitic/affixal pronoun, inflectional person marker, zero anaphor)

b.   determiners (demonstratives, definite/indefinite and specific/non-specific articles, determinative/deictic clitics and affixes)

c.   case marking that is sensitive to referentiality (e.g. differential object marking)

d.   non-personal referential marking on verbs (switch-reference, obviation)

e.   constituent order

As we can observe, the list of referential devices provided by Kibrik (2011:37) is shorter, as he does not consider determiners, case marking, non-personal referential marking and constituent

order. Lehmann's inclusion of these referential strategies has a direct impact on the methodology and results of referentiality studies. His approach takes into account a larger range of coding possibilities for languages crosslinguistically, making it possible to analyze the different referential strategies of typologically diverse languages. For this reason, I will adopt Lehmann's approach in this work.

The referential choices of the speaker depend on the activation or accessibility of the referents in discourse. They also depend on the situational context and shared knowledge (see section 1.5.2). Overt NPs can have diverse structures as they can be more or less descriptive, rendering various degrees of information. Stoll & Bickel (2009) define *informativeness* as the degree of referential information given in a NP. Referential elements fall into a continuum of informativeness depending on their semantic structure, ranging from explicit to implicit mechanisms. Pronouns are closer to zero anaphora when they only encode number (for instance the pronoun *they* in English). They get closer to generic NP as they include other categories like gender, spatial distance or honorific degree (Stoll & Bickel 2009).

Lehmann (2012) proposes a scale that includes, as was mentioned before (Table 2: Lehmann's referential strategies), additional referential devices like pronominal affixes; he also makes a distinction among different types of pronouns:

Table 3. Explicitness of coding of referents (Lehmann 2012).

representation	explicitness
lexical NP	explicit
emphatic pronoun	
neutral pronoun	↑
clitic pronoun	
pronominal affix	
zero	implicit

Informativeness can be measured by assigning a numerical value to referential expressions, depending on their level of explicitness or complexity. This numerical value is called *referential weight*. Referential expressions were ordered by Ariel (1990) on a scale of decreasing complexity/explicitness as shown in table (4). According to Ariel (1990) explicitness is related to accessibility of the referent in the same way it is to Givón (1995), as we will see in section 1.4.2.

Table 4. Accessibility Marking Scale (Ariel 1990)
Low Accessibility
a.   Full name + modifier
b.   Full ("namy") name
c.   Long definite description
d.   Short definite description
e.   Last name
f.   First name
g.   Distal demonstrative + modifier
h.   Proximal demonstrative + modifier
i.   Distal demonstrative (+NP)
j.   Proximal demonstrative (+NP)
k.   Stressed pronoun + gesture
l.   Stressed pronoun
m.   Unstressed pronoun
n.   Cliticized pronoun
o.   Extremely High Accessibility Markers (gaps, including pro, PRO and *wh* traces, reflexives, and Agreement)
High Accessibility

Lehmann (2012) used another version of this scale to assign a referential weight to all references in his Referentiality Project. The relation between expliciteness and weight of references is shown in table 5.

Table 5. Scale of complexity of referential expressions.		
Referential expression		**Weight**
a.	clause (completive, e.g. the Obj of a say verb)	9
b.	(common) noun phrase with head and (clausal or adjectival) attribute	8
c.	(common) noun phrase with head and lexical possessive specifier	7
d.	(common) noun phrase with head and determiner	6
e.	proper name	5
f.	bare common noun	4
g.	emphatic pronoun	3.5
h.	neutral pronoun	3
i.	clitic person form	2
j.	bound person form (affix)	1
k.	ZERO person form	0

1.4. Information structure

Lehmann (2008) considers information structure to be another functional domain of language. It relates to the ways in which information is packaged in each sentence and in discourse, according to the speaker's evaluation of the Universe of Discourse. This functional domain has its own subdomains (Lehmann 2008:208):

1. presupposition vs. assertion of propositions (known vs. not yet known),

2. identifiability and activation of referents (status of mental representations in the addressee's mind, from active to brand-new anchored[1]), discussed

---

[1] The definition of referential anchorage is clarified in section 1.5.2.

3. topic vs. focus status of elements of propositions (predictability vs. unpredictability of relations between propositions and their elements).

According to Saeed (1997), information structure is the grammaticalization of the assumptions made by speakers about the knowledge of their listeners. Grammaticalization here refers to these assumptions being reflected in sentence form and referential choices. As participants of the conversation share their knowledge, it emerges the distinction between the information assumed to be known and the information that is being presented as new.

In general, the use of a definite nominal is the way that languages have grammaticalized the assumption that the information needed for referent identification is already given by the speaker, or known by the interlocutors. Typically, referents are introduced by an indefinite nominal; then a definite article can be used, later to be replaced by an anaphoric pronoun or zero anaphora, while the referent is still activated or salient in the conversation (Saeed 1997). These uses of referential expressions are called **coreference**. Coreferential elements are therefore the grammaticalization of the information status of nominals within discourse.

1.4.1. Topic and focus

Topic and focus markers are part of information structure. There are two notions of topicality: sentence topic and discourse topic. Discourse topics are referents maintained active throughout discourse by different conversational devices (see Saeed 1997: 200). According to Givón (1995) the discourse topic is the unmarked participant since it is more salient, that is, more recurrent and cognitively activated. Sentence topics or 'chain-initial topics' (Givón 1983), on the other hand, correspond to newly introduced referents in a sentence. Topics tend to be coded by the subject/agent and direct-object/patient, although peripheral arguments may also be topicalized

(by means of syntactic phenomena like word order). Both notions of topic are very close to each other, as the sentence topic is often the discourse topic. Focus markers tend to point out elements in a sentence as new. It can also mark a referent as salient when there are a number of alternative interpretations (Saeed 1997).

1.4.2. Referential distance

With this respect, Givón (1995) stated that a continuous or accessible topic (or referent) will need less linguistic material to be coded. Based on this knowledge of referential codification in languages, it is understood that referential choice in discourse is related to cognitive capacities and operations (like attentional focus and working memory) (Kibrik 2001). Givón (1995), measures the accessibility of referents by means of *referential distance*. Topical referents tend to have low referential distance. Referential distance is calculated by counting the number of clauses that separate a reference from its previous occurrence. When the previous occurrence of a referent is in the immediate preceding clause, it has a referential distance of 1. When it is in the second or third clause, a value of 2 or 3 is assigned. If there is no antecedent in the last 3 clauses, it is assigned a value of >3. Generally, highly topical referents (usually coded by zero anaphor) tend to have a referential distance of 1. Referents that are continuous in discourse also tend to be coded with zero anaphora or unstressed pronouns, whereas discontinuous referents are expressed by means of full lexical nouns or stressed pronouns. This relation is shown in the scheme below (Givón 1995):

maximal continuity ➔ zero anaphora, unstressed pronoun

discontinuity ➔ stressed pronoun, full lexical noun

Maximal continuity coding devices indicate the (cognitive) default choice of *continued activation* of the topical referent. Discontinuous devices, by contrast, indicate the termination of the current topical referent, along with the activation of a new topic. Stressed pronouns may serve as a short-distance device, which indicates activation of a new referent without terminating the current one (Givón 1995:364).

1.5. Discourse analysis

1.5.1. Typology of text/discourse genre

In a referential study it is important to identify the genre of the text that is being analyzed. Knowing the text genre will enable us to make some predictions about the flow of information and recognize the origin of certain phenomena. Smith (1985) analyses the importance of certain characteristics of texts needed for a taxonomy. This taxonomy considers two main features (agent orientation and chronological succession), that combine to render four text types, as shown in the table:

Table 6. Text types (Smith 1985: 231).

	+AGENT ORIENTATION	-AGENT ORIENTATION
+ CONTINGENT TEMPORAL SUCCESSION	NARRATIVE 1. 1/3 person 2. Agent oriented 3. Accomplished time 4. Chronological linkage	PROCEDURAL 1. Non-specific person 2. Patient oriented 3. Projected time 4. Chronological linkage
-CONTINGENT TEMPORAL SUCCESSION	BEHAVIORAL 1. 2 person 2. Addressee oriented 3. Mode, no time 4. Logical linkage	EXPOSITORY 1. No necessary person ref. 2. Subject matter oriented 3. Time not focal 4. Logical linkage

The characteristics of text types are reflected in particular linguistic forms and uses. The feature of [+agent oriented] is realized through the use of first or third person pronouns as agents. The feature of [+contingent temporal succession] is reflected in the use of the past tense, temporal conjunctions or temporal adverbials. Thus, the features of each text type would be distributed as follows (Smith 1985):

a) Narratives: 1/3 person pronouns, Agent/Experiencer subjects, Past/historical present tense, head-head/tail-head linkage.

b) Procedural: 1/2/3 person (general) pronouns, Patient as object, Present/future (customary) tense, head-head/tail-head linkage.

c) Behavioral: 1 incl., 2 or 3 person pronouns, imperative/jussive/hortative modes, linkage by condition/cause/purpose margins.

d) Expository: 3 person pronouns, equative/descriptive clauses, various tenses, linkage by sentence topic/parallelism.

1.5.2. Universe of Discourse

In real life, language users must combine semantic knowledge of language with knowledge of the world to make inferences and correctly interpret the intended meaning, which means that listeners also participate actively in the construction of meaning (Saeed 1997). Both kinds of knowledge (knowledge of the world and of language) are part of the context that helps hearers understand what the speaker says.

When a speaker makes an utterance he/she has to estimate the amount of knowledge the hearer possesses. According to Saeed (1997), this knowledge may come from three sources:

1. physical context (deixis)

2. discourse context

3. background or common knowledge

Lehmann (2014) considers that there are four ways for the speaker to create a representation of the referent in the hearer's mind. He calls these knowledge sources the *four spaces of the speech acts*, which are listed below.

1. speech situation

2. universe of discourse

3. encyclopedic knowledge

4. shared experience

The **speech situation** or physical context involves space and time. It comprises all physical objects present during the speech act (Lehmann 2014), and is coded in deictic expressions. In order to understand utterances, the hearer must be able to identify the referents of pronouns like *I* and *you*. Even though the interpretation of these pronouns is always different, the identification of referents is normally understood from context without problems. Linguistic elements that are contextually bound are part of deictic categories, which are of various kinds: temporal deixis, spatial deixis, person deixis and social deixis (Saeed 1997).

Spatial deixis refers to adverbs like *here/there* or demonstratives like *this/that*. The organization of space is different from one language to another, and the reference point of spatial deictic markers can be the speaker and the hearer. Spatial deixis is frequently used to point out information within discourse (which is then called discourse deixis), and it can also be employed to refer to time. Person deixis concerns the grammaticalization of participants (including the speaker, the hearer) and non-participants of the conversation; this information is grammaticalized

in the form of pronouns (*I*, *you*, *we*, *they*). Again languages differ in the kind of information that is packaged along with person, some may also include number and/or gender. Social deixis grammaticalizes information about social relationships and identity of the participants. Social deixis can be included in the system of person deixis, like the distinction between familiar and polite pronouns of the European languages (Saeed 1997).

The **universe of discourse** or discourse context is "an intensional world created, modified and maintained by interlocutors in a particular speech situation" (Lehmann 2014). According to Saeed (1997), the discourse context is given by the discourse topic, which influences the way listeners interpret the meaning of what is being said. One utterance may have very different interpretations depending on what the discourse is about. The current discourse topic provides information for the listener to make the right inferences about referents.

Encyclopedic knowledge or background refers to sociocultural and real-world knowledge that is shared by the participants in the conversation. This information is independent from any given information in the current or previous discourse, and is acquired by membership in a community (Saeed 1997). Encyclopedic knowledge is especially important as a referential anchor, given that in many cases it is the only means for listeners to understand what kind of entities are involved or what is the relation among them. This would be the case if a speaker makes an utterance like *yesterday I was visited by* _a colleague from Afghanistan_ with no previous introduction of the referent (Lehmann 2014).

Sometimes the identification of a specific referent can only be accomplished by **shared experience**. In such cases the referent is stored in the memories of both speaker and hearer by having lived in a situation jointly, or by having had the same experience independently. Lehmann (2014) points out that shared experience and encyclopedic knowledge can overlap:

22

"Given the context in which we live, a subset of those entities that make up our encyclopedic knowledge, importantly comprising certain unique objects, may be part of our daily experience. This subset will then have a special status in reference. That is the case in *I must go to the market*, where in the absence of indications to the contrary, the market in our horizon is being referred to. Note that in such cases, no special operation of recall from memory is conveyed."

In the universe of discourse of a given speech situation, the referents are anchored through different referential operations, which correspond to the four spaces of the speech act:

Table 7. Anchorage spaces		
Speech act space	**Referential operation**	**Comment**
Speech situation	Exophora (deixis)	Referent is outside the text, in the physical world, but within the universe of discourse
Universe of discourse	Endophora (coreference)	Referent is inside the text. It presupposes its existence in the UD
Encyclopedic knowledge	Concepts and representations	Referent is known by membership in a community
Shared experience	Anamnestic reference (recognition)	Referent is stored in the memories of both speaker and hearer.

CHAPTER II

THE TARAHUMARA LANGUAGE

2.1. Cultural and environmental context

Speakers of the Tarahumara language live in the state of Chihuahua, located in the Sierra Tarahumara region in North Mexico. Tarahumara people call themselves *Rarámuri*, which can be translated as "people", as opposed to mestizos or outsiders, who are called *chabochi* ('those who have beard') or *yori* (Pintado 2004). The term rarámuri is also associated with the meaning of 'foot runner' or 'light feet', given that these people are excellent runners (Bennet and Zingg 1930).

The Sierra Tarahumara region displays an extreme climate: while temperature in the *barrancas* can reach 40°C in summer, it descends several degrees below zero in the highlands during winter. In this region, water is a scarce resource, although they do have rainy periods in summer (from June to August) and in winter (from December to February) (Pintado 2004). In some areas of this region, particularly within caves, the dry atmosphere does not change, allowing for a natural mummification and preservation of corpses (Mansilla 2011). In spite of the dry conditions, the Tarahumara are able to produce corn, beans, potatoes, oranges, apples and chili, among other things.

Map 1. The Tarahumara region (Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013).

Because of the extreme weather of the region, they are a semi-nomadic society, living in a summer house in the highlands, and moving to their caves in the *barrancas* to spend the winter. The Tarahumara society is co-operative, helping each other in exchange for food and *tesgüino*, a corn based beer. Men and women have very well-defined, distinctive roles, which make them mutually dependent. Men are required to work outside the home, cultivating the field, watching over the cattle, and making all kinds of things needed: hand tools, musical instruments and houses. Sometimes they go out of the community to sell or trade their products. Government and leadership are also male tasks. Woman usually stay home to look after their children and take care of the herds that will provide them with the wool needed for making blankets. Many products for daily life are homemade by women, such as baskets, pottery and cooking instruments. In addition, women must prepare food and *tesgüino* for everyday meals, as well as for *tesgüinadas*, their traditional ceremonies (Bennet and Zingg 1930).

Tarahumara religion is a fusion of their prehispanic beliefs with those of Christianity. They believe in the *arewá* 'the soul' or 'energy', which keeps them alive. According to the Tarahumara, their *arewá* is strengthen by music (Pintado 2004); consequently, music and dancing play an important role in ceremonies and celebrations.

2.2. Linguistic family

The Tarahumara language is affiliated to the Uto-Aztecan (or Uto-Nahua) family, which is divided into northern and southern languages. According to Estrada (2004), this family takes its name from the designations of the indigenous groups that lived at the two poles of its geographic distribution, as they were found by Europeans: Ute at the northern end (Arizona, California, Colorado, Nevada and Utah) and Aztec or Nahua at the southern end (Central Mexico). The Tarahumara language belongs to the Tarahumara-Guarijío branch of the southern group. The organization of the entire family is shown in table 8 (adapted from Estradaet al. 2004):

Table 8. Classification of the Uto-Aztecan family	
Northern Uto-Aztecan	

1.	Numic
	    Western: Mono, Northern Paiute
	    Central: Timbisha Shoshoni (Panamint), Shoshoni, Gosyute, Comanche
	    Southern: Kawaiisu, Chemehuevi, Southern Paiute, Ute

2.	Takic
	    Cupan: Cahuilla, Cupeño, Luiseño
	    Serrano, Gabrielino-Fernandeño

3.	Tübatulabal (or Río Kern)
4.	Hopi
Southern Uto-Aztecan	

5.	Pimic
	    Pima-Tohono O'odham
	    Pima bajo (Névome) (extinct)
	    Pima bajo (from the mountain)
	    Northern Tepehuan

	Southern Tepehuan	
	Tepecano	
6.	Opatan	
	Opata	
	Eudeve	
7.	Tarahumara-Guarijío	
	Tarahumara	
	Guarijío	
8.	Tubar	
9.	Yaqui-Mayo (Cahitan)	
	Yaqui	
	Tehueco (extinct)	
	Mayo	
10.	Corachol-Aztecan	
	Corachol: Cora, Huichol	
	Nahuatl	

Tarahumara is one of the most widely spoken languages in Northern Mexico. It is spoken by approximately 85,316 people (INEGI 2010). The municipalities of Guachochi, Urique, Bocoyna and Balleza, in the state of Chihuahua, show the highest densities of Tarahumara speakers. About 13% of Tarahumara speakers are monolingual; they are mainly found in Guachochi, Batopilas and Balleza (INEGI 2009:121).

There are five varieties of this language, separated by natural geological formations of the Sierra Madre Occidental (Lewis 2015). Table 9 displays all five varieties along with their language code and municipalities where they are spoken.

Table 9. Tarahumara varieties (Lewis 2015)		
Variety	**ISO Code**	**Municipalities**
Southwest Tarahumara	639-3 twr	Tubare
Northern Tarahumara	639-3 thh	Santa Rosa Ariseachi, Agua Caliente Ariseachi, Bilaguchi, Tomochi, La Nopalera
Southeast Tarahumara	639-3 tcu	Chinatú, Turuachi, Balleza
Western Tarahumara	639-3 tac	Urique, Uruachi, Guazapares, Chínipas, Rocoroibo and Bocoyna.
Central Tarahumara	639-3 tar	Cuautemoc, Norogachi, Guachochi

According to Lewis (2015), Northern Tarahumara has 45% intelligibility with the Central variety (tar) and 25% intelligibility with the Western variety (tac). On this matter, nothing is said about the remaining dialects. Much of the dialectal variation in Tarahumara may have its origins in the free vocalic and consonantic variations this language exhibits. One can often find the following alternations, among many others:

(1) a. *bilé ~ biré ~ birée*   'one'

b. *beteame ~ pereame*   'inhabitant'

c. *ba'wí ~ wawí*   'water'

d. *onerá ~ onora*   'father'

e. *jéena ~ jináa*   'this'

f. *ko'a~ go'a*   'to eat'

28

An additional common variation source is reduction of words by dropping initial or final sounds or syllables. This variation may occur within the same variant and even produced by the same speaker.

(2)   a. *echi ~ chi*   'he/she/it, this'
            b. *rotori ~ toorí*   'hen'
            c. *ketasi ~ ke*   'no'
            d. *tabilé ~ ta*   'no'

Nevertheless, Villalpando (2010) presents some data to illustrate these differences among dialects through the comparison of some lexemes. Table 10 gathers some of the information he offers (Villalpando 2010).

Table 10. Phonological dialectal variation in Tarahumara					
	Western	**Northern**	**Central**	**Southeast**	**Southwest**
blood	*ela*	*la*	*la*	*la*	*la*
heat	*rahtá*	*ratalí*	*ratalí*	*ratá*	*ratá*
dough	*párisi*	*batusí*	*batusí*	*batusí*	*tusí*
far	*mehká*	*meká*	*meká*	*miká*	*mikiá*
house	*galí*	*galí*	*alí*	*karí*	*puruí*

Needless to say, this sort of superficial comparison is not enough to see how deep the differences among dialects are.

2.3. Word order and Alignment type

Tarahumara has a strong tendency to place the verb at the end of the clause. The most frequent constituent orders to be found in Tarahumara are APV, AVP and PVA (Cohen 1998). Another order that frequently appeared in the sample used in this work was VAP. Examples of these constituent arrangements are the following. In an intransitive clause the order is most often SV.

(3) APV

						A	
a	*bela*	*kó*	*bachá*	*rikó*	*atí*	*mukí*	*ko*
LOC	DP	DP	in.front.of	DIR	be	woman	DP

	P	V
walú	*kobísi*	*nawa-yá*
INT	pinole	make-HAB

'First the woman makes pinole'

'Primero la mujer hace mucho pinole'

(4) AVP

A						V	P		
échi	*mukí*	*ko*	*ayéna*	*bila*	*walú*	*manii*	*échi*	*batáli*	*bá*
DEM	woman	DP	CON	DP	INT	put	DEM	tesgüino	CL

'The woman also makes a lot of tesgüino'

'La mujer también le hace bastante tesgüino'

(5) VAP

V			A			P
alí	*katéwi-ka*	*achá*	*échi*	*mukí*	*ko*	*basíawi*
SIM	keep-STAT	be	DEM	woman	DP	corn

'The woman puts the corn away'

'La mujer deja guardado el triguillo'

Given that this language has no morphological case marking, it shows neutral alignment for full NPs; nevertheless, nominative-accusative alignment emerges in its pronoun paradigm. Pronominal forms in Tarahumara make a distinction between subject and non-subject for the first and second person, as shown in table (11). Third person pronoun, as in many other languages, is

expressed by means of a demonstrative pronoun and its form is invariable (see section 2.7 on referential expressions in Tarahumara).

Table 11. Pronominal paradigm for Tarahumara (Pahuírachic, Chih.)				
	SINGULAR		PLURAL	
	SUBJ	NON-SUBJ	SUBJ	NON-SUBJ
1SG.	*nejee*	*tami*	*tamuje*	*tami/tamo*
2SG.	*muje*	*mi*	*tamuje*	*tamuje*

Tarahumara combines neutral alignment (for full NPs) with nominative-accusative alignment (for personal pronouns), which is the most frequent alignment combination found among the languages that have a different case marking alignment of full NPs and pronouns, according to the results of the survey made by Comrie (2013), which included 190 languages. Neutral alignment is exemplified in (3-5). Example (6) shows the nominative-accusative alignment in the pronominal paradigm.

(6) The verb 'to give' in Tarahumara:

nejee *echi aari*	'I gave (it) to him'
échi **tamí** *aari*	'he gave (it) to me'
échi **mi** *aári*	'he gave (it) to you (sg)'
echi tamoó aari	'he gave (it) to us'
nejée mii *aaki*	'I gave (it) to you (pl.)'
nejée *echi aari*	'I have (it) to them'

2.4. Morphological features

Following Comrie (1989) on his description of polysynthetic languages, and considering the synthetic-analytic dimension, it can be argued that Tarahumara is more inclined towards the

analytical side of the continuum (Villalpando 2010:72). Words tend to be comprised by one to three syllables, Most Tarahumara verb stems take up to three suffixes, as shown in (7):

(7)   *ko'á-r-ti-nare*

        comer-CAUS-CAUS-DES

        'She wants to make him eat'

        'Quiere hacerlo comer'   (Caballero 2008:94)

As it was previously said, this language tends to truncate phonological material from words. However, the extent to which words are reduced varies from one dialect to another. Some dialects may drop one sound or one syllable, while others go as far as to *keep* one syllable, in some cases. Table 12 below illustrates this situation (Villalpando 2010).

Table 12. Word truncation across dialects					
	Western	**Northern**	**Central**	**Southeast**	**Southwest**
hill	*gawí*	*gawí*	*awí*	*kawí*	*awí*
pine tree	*okó*	*okó*	*okó*	*okó*	*ko*
ten	*makué*	*makói*	*makói*	*makói*	*mkó*
skin	*wichíla*	*wi'chí*	*wichíla*	*wichí*	*chíl*
red	*setákame*	*sitákame*	*sitákame*	*sitákame*	*sitákam*

Regarding the fusional criterion, Tarahumara is an agglutinative language, given that morphological segmentation is normally an easy task and morphemes are generally semantically simple. Apart from the difficulties that come with free vocalic and consonantic variation, morphemes are easy to identify once one is familiarized with the language, and with the specific dialect that is being analyzed. Some dialects, by instance, would prefer the allomorphs *–li* or *–ti* for the perfective marker *–ri,* although alternation of these allomorphs may as well occur within the same variant. This is illustrated in example (8) for the Choguita variant.

(8) a. atíʃ-**ri** ~ atísi-**li** 'sneeze-PST' (Caballero 2008:39)

  b. ritú-**r-l-**o ~ ritú-**ri-li-**o 'freeze-CAUS-PST-EP' (Caballero 2008:43)

This language has neither case, number nor gender markers on nouns or adjectives. It also goes without personal indexation on verbs. Tarahumara expresses syntactic function through word order: As tend to be positioned before Ps, regardless of the positioning of the verb (although pragmatic factors may allow exceptions); function may also be inferred by verb semantics. The idea of plurality is coded analytically by means of numerical elements and quantifiers such as *pé okwá* 'a couple', *walubé* 'many, more', *siné* 'some', *táa* 'little' etc. (example 9).

(9) a. *échi pé okúa rochí*

   DEM few two fish

   'a couple of fishes'

   'unos cuantos pescados'

  b. *táa batali*

   pequeño tesgüino

   'a little tesgüino (beer)'

   'un poco de tesgüino'

Tarahumara does not have gender as a grammatical paradigm, it only makes a natural gender distinction for humans and some animals, when relevant. Feminine and masculine distinctions on humans are made through a suppletive mechanism available for nouns and some adjectives, as shown in examples (10a-f) taken from Brambila (1953:13):

(10)　　　**Masculine**　　　　　　　　　　　　　**Feminine**

　　a.　*towí*　　　'boy'　　　　　　　*tewé*　　　'girl'

　　b.　*boní*　　　'younger brother'　　*waí*　　　'younger sister'

　　c.　*očíkari*　　'paternal grandfather'　*akáčuri*　'paternal grandmother'

　　d.　*si'ará*　　'father in law'　　　*wa'šira*　'mother in law'

　　e.　*wičuwati*　'crazy.M'　　　　　*wičulati*　'crazy.F'

　　f.　*o¢érame*　'old.M'　　　　　　*weráame*　'old.F'

When the situation requires a gender distinction to be made on animals, Tarahumara turns to an analytic mechanism: it uses a modifier of the noun that expresses the meanings of 'male' or 'female' (Brambila 1953:13):

(11)　a.　*totorí owira*　　　　'rooster'

　　b.　*totorí mukira*　　　　'hen'

Although it is generally accepted that Tarahumara has no morphological means to code person or number on verbs, Caballero (2008) has characterized the verbal suffix *–po* (*~ -bo*) as an indicator of number and time. These suffixes appear in various dialects but are not obligatory and do not occur frequently. She also identifies the prefixes *i-* and *na-* as plural markers and *ni-* as a singular marker for the variety she describes (from the municipality of Choguita). The latter morphemes are not seen consistently in other varieties. Even in the Choguita variety, Caballero (2008) mentions the verbal lexeme by itself as an alternative to code the singular (without using the singular morpheme. Examples are given in (12).

(12) a. *rapáa-m-**po***   (Caballero 2008:40)

split:APPL-TR-FUT:PL

b. *newá-**bo***   (Caballero 2008:46)

make-FUT:PL

c. ***i**-béwa*   (Caballero 2008:45)

PL-smoke:PL

d. ***na**-kúri*   (Caballero 2008:124)

PL-help

e. ***ni**-kúri*   (Caballero 2008:124)

SG-help

f. *ku'íri*   (Caballero 2008:124)

Moreover, there are some suppletive verbs that are used when a plural subject is involved in the event. This resource is not available for all verbs in the language. Although different authors have mentioned some of the verbs that count with a suppletive plural, to this day there does not exist a list of specific verbs, so as to know how many verbs or what kind of verbs have a plural counterpart. The examples here provided (13) are taken from Caballero (2008:124) and include transitive and intransitive verbs.

(13) | | **Singular** | **Plural** | **Gloss** |
|---|---|---|---|
| a. | *asi* | *močí* | 'sit' |
| b. | *wiri* | *hawá* | 'stand' |
| c. | *bo'i* | *bi'ti* | 'lie down' |
| d. | *bakí* | *mo'í* | 'go in' |
| e. | *mi'ri* | *ko'i* | 'kill' |

2.5. Transitivity and valency

As Nariyama (2001) has proved, knowing the argumental structure and semantic/syntactic properties of predicates is helpful in the process of inferring and identifying implicit arguments. This study, however, does not offer a description of the argumental structure of predicates in Tarahumara. It merely offers an overview of some syntactic mechanisms for valency changes that may, similarly, be useful in the identification and tracking of referents for this referential study.

Although there are different word orders in this language (see section 2.3), the basic order is APV. In an intransitive construction, the argument precedes the verb in most cases, as in (14a); the transitive construction is shown in (14b). Tarahumara seems to tolerate only two adjacent arguments before the verb: when an oblique argument is added to an intransitive construction, it appears in second position, after the subject and before the verb (14c); when it is added to a transitive construction, it goes *after* the verb (14d).

(14) a. *towí   maá*
          boy   run
          'the boy runs'
          'el niño corre'

      b. *Nejée kochí chóokame   raráa-ri*
          1SG   dog   black   buy- PFV
          'I bought the black dog'
          'yo compré el perro negro'

      c. *Echi   rejói kawi-chi   maa-ri*
          DEM   man   mountain-LOC run-PFV
          'the man ran on the mountain'
          'el hombre corrió en la montaña'

36

d.   *Nejee   músa   rarií-ki   Pedro*

      1SG   cat   buy- PFV   Peter

      'I bought Peter a cat'

      '(le) compré un gato a Pedro'

In Tarahumara we can find some valency reduction mechanisms like passive and impersonal constructions, as well as valency increasing mechanisms such as causative and applicative constructions. The **passive** construction is illustrated in (15). It is through the passive suffix *–ru* (or *–tu*) that the patient is promoted to the subject position and the original subject (the A) is omitted.

(15)   a.   *Nejée paan chikrée   ripayá nati*

            1SG   bread cut   knife INSTR

            'I cut the bread with the knife'

            'Yo corté el pan con cuchillo'

      b.   *Echi   paan ripiyáa   nati   chikire-**ru***

            DEM   bread knife   INSTR cut-PASS

            'The bread was cut with the knife'

            'El pan fue cortado con un cuchillo'

      c.   Active: Passive:

           SBJ —— OBJ —— OBL   (OBL) —— SBJ —— OBJ

            | | | | | |

           agent   patient   instr.   agent   patient   instr.
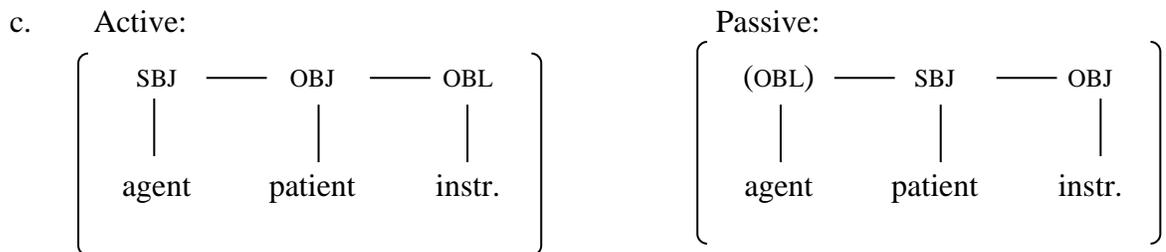
Valdez (2013) recognizes the morphemes *-ia*/*-riwa*/*-wa* as additional passive markers, indicating that they can occur with transitive and intransitive verbs, but receiving an impersonal interpretation in the case of intransitive verbs. In this study they are treated solely as **impersonal**

markers that make reference to a non-specific person, like the English *one* (as in "one must cultivate one's garden"). Thus, impersonal constructions are less transitive in the sense that they do not code a specific lexical referent. Nevertheless, valency is not really affected, since the agent/subject of the clause is still explicitly present (as a non-specific referent) through the impersonal suffix. Similarly, the **applicative** *–lta* (or *–la, -ra*) codes a non-specific object/patient in a verb where this argument could be easily omitted, that is, it could be coded as an intransitive clause and leave the presence of the argument to inference, but the speaker decides to code it without specifying it. In (16b), the non-specific argument refers to a beverage.

(16)   a. *a   bilá   ko   échi bela owirú-ame   bela maníi-**wa**   échi batáli*

             LOC   DP   DP   [DEM one   heal-NMLZ]   DP   put-IMPRS   [DEM tesgüino]

             'one makes the batali for the healer'

             'así pues, se le pone al curandero el tesgüino'   (Trabajos-Cl109)

      b. *ayéna chukú échi   mukí   pákia-**la**   ba*

             CONJ AUX [DEM woman]   drain-APPL   CL

             'also this woman drains (it)'

             'también se pone esta mujer a colar (lo)'   (Trabajos-Cl80)

In the **causative** construction the causer is introduced in subject position and the causee (the agent) takes the second position; this is exemplified in (17). The causative marker takes the form of *–ri* (or *–ti*). Caballero (2002) adds the allomorphs *–ri, -re, -tia, -ria,* registered in the Ojachichi variety; however, the morpheme *–ria* (along with its allomorphs *–lia* and *–ya*) seem to perform another function closer to an aspectual marker, in the varieties that have been analyzed in this work.

(17) a. *towí   maá*
          boy   run
          'the boy runs'
          'el niño corre'

      b. *mukí   ranára maa-r-ti*
          woman   child   run-CAUS-PFV
          'the woman makes the child run'
          'la mujer hace correr a su hijo'

      c. Intransitive: Causative:

          ⎛ SBJ   ⎞   ⎛ SBJ ───── OBJ ⎞
          ⎜   │   ⎟   ⎜   │   │   ⎟
          ⎝ agent   ⎠   ⎝ causer   agent ⎠
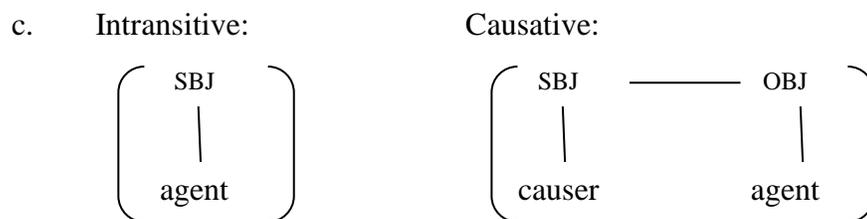
Causative markers may also adhere to transitive verbs, as in the case of the verbs *beneri* 'to teach' (lit. to cause to learn) and *riwiri* 'to show' (lit. to cause to see). In such cases, the causer and the agent precede the verb, and the patient or theme follows it. Causativity may also be expressed syntactically, using verbs like *orá*, 'make', as shown in (18).

(18) *echi   mukí   orá-ri   mapu towí   kochi-méa*
      DEM   woman   make-PFV   SUB   boy   sleep-IRR
      'The woman put the boy to sleep'
      'La mujer hizo que el niño se durmiera'

2.6. Nominal grammar

Given the analytical nature of this language, nouns are generally not very complex; they often consist of a lexeme by itself. This is not to say that there is no nominal inflexion at all. Nouns

39

may take the locative suffix *–achi* (or *–chi*) or the possessive suffix *-wa*. The locative morpheme usually appears with names of places, common nouns that refer places/physical spaces, and nouns that refer temporal spaces. It can also be used to refer events or to index body parts.

(19)   a.   *kawi-chí*   'at the mountain'
              mountain-LOC

       b.   *napí-chi*   'on the fire'
              fire-LOC

       c.   *wíwa-achi*   'harvesting season'
              to:harvest-LOC

       d.   *alíwa-achi*   'in the evening'
              evening-LOC

       e.   *batáli-chi*   'at the batali party' (tesgüinada)
              batali(beer)-LOC

       f.   *ropa-chí*   'in the stomach'
              stomach-LOC

The locative suffix also appears with verbal forms, producing nouns, as in (20):

(20)   *biti-chi*   'house'
              to:live-LOC

The possessive suffix does not seem to be obligatory, since there is also a possessive pronoun. Sometimes both the possessive pronoun and the possessive suffix appear around the noun, and sometimes only the possessive pronoun is used. The possessive suffix, however, does not appear on its own.

(21)   a.   **kini** *panero*   'my husband'

       b.   **kita** *retéma*   'our people'

       c.   **kípi**   *kúchu-**wa***   'her children'

Many nouns in Tarahumara have a *–li* or *–i* ending, however this is not a rule.

40

(22) a. *kobísi*   'pinole (beverage)'

b. *napíwili*   'corn' (sp. nixtamal)

c. *mukí*   'woman'

d. *rejói*   'man'

e. *muní*   'beans'

f. *namúti*   'thing'

g. *pachí*   'small garden' (sp. milpa)

h. *rochí*   'fish'

Even though it is unusual, this ending can adhere to verbal lexemes in order to make nouns. In example (23a), the presence of the modifier *walú* 'big', and the fact that the construction has a transitive verb (along with its subject) corroborate that the word *nóchali* 'work' is in fact functioning as a noun, meaning that the suffix –*li* is working as a nominalizer, perhaps by analogy. In (23b) we have the same modifier preceding the word *omáwali* 'celebration', which appears as the subject of an existential construction. However, this assertion of –*li* being a nominalizer should be further studied in order to see if this mechanism actually comes from this nouny ending. Another possibility is that it comes from the perfective suffix –*ri*, which has –*li* as one of its allomorphs. It could be that the perfective marker has been producing participles that are then used as nouns.

(23) *a.   ayéna bela   walú   nócha-**li**   níwi   échi   bilé   mukí*
         CONJ   DP   big   to:work-NMLZ   have   DEM   one   woman
         'And this woman has a lot of work (to do)'
         'También pues tiene mucho trabajo una mujer'

41

b. *walú   omá-wa-**li***   *niru-li-sa*   *siné   kaachi*

   big   to:celebrate-IMPR-NMLZ   to:be-PFV-COND   some   times

   'Sometimes when there are big celebrations'

   'En ocasiones cuando hay grandes festejos'

A very productive nominalizer in Tarahumara is the suffix *–ame*. This suffix can be attached to nominal and verbal lexemes to produce nouns and adjectives. Examples (24a) show the suffix *–(k)ame* with a nominal base and in (24b-d) it appears with a verbal base, all receiving a nominal meaning.

(24) a. *bantéli-ame*   'the one who carries the flag' (sp. el abanderado)

         flag-NMLZ

     b. *ko'a-ami*   'food'

         eat- NMLZ

     c. *wilí-ame*   'leader'

         put(in charge)-NMLZ

     d. *newá-ami*   'maker, manufacturer'

         make-NMLZ

Yet another nominalizer is the suffix *–ka*. A couple of examples are given in (25).

(25) a. *napabu-ka*   'gathering'

         gather-NMLZ

     b. *mani-'i-ka*   'food'

         put-PRF-NMLZ

This marker has traditionally been accepted as a stative/progressive morpheme in the literature, although it has been tagged in many ways. Nevertheless, there is currently some discussion over the real meaning of and adequate tagging for this morpheme and the former

42

morpheme *–(k)ame*. Table (13) makes a comparison of the different meanings/tags that have been assigned by different authors to each one of these suffixes.

Table 13. Interpretations and tagging for *–ka* and *–(k)ame*				
Guadalaxara (1683)	**Brambila (1953)**	**Lionnet (1972)**	**Burgess (1984)**	**Caballero (2008)**
	-rame ~ -kame Stative, Participal pres. *-reame ~ -keame* Participial. perf.	*-me ~ -mi* Participial, Adjective	*-ame ~ -me* Participial	*-ame* Participial
-go ~ -ga Pluscuamperfect		*-ga ~ -ka* Durative present	*-ga* Simultaneous, Continuative, Stative	*-ka* Gerund

It is also questionable whether they should be considered nominalizers or exclusively aspectual/temporal markers. Thus, the discussion is not a trivial one, and is not limited to the semantic domain, since the different interpretations of these morphemes may affect the morpho-syntactic description of this language.

2.7. Referential expressions

As seen in section 1.2.1, languages manage to code referents specifically, non-specifically, or as general concepts; and they do this using a wide range of constructions and resources that have different degrees of complexity. The degree of complexity is dictated by pragmatic factors, as well as situational and discursive context.

Tarahumara uses bare nouns mostly to code non-specific or general concepts. Indefinite entities may also be coded by the quantifier *bile* (or *bileana*) 'one'.

(26) **bilé** *mukí*   *ko*   **bilé**   *rawé*   *náli*   *wika*   *nátame*   *nócha*

[one woman] DP   [one   day]   COMIT mány   ways   work

'a woman can work in many ways in one day'

'una mujer en un día llega a trabajar en diferentes maneras'

It makes use of demonstratives as determiners when referring to a specific entity. Some

frequent determiners are *echi* and *jenaí*, among others. The demonstrative *echi* may also function

as a third person pronoun, accordingly it can occur by itself, referring to humans and to all kinds

of entities. Demonstratives/determiners, quantifiers and possessive pronouns precede the noun.

(27) a. **jenaí** *ayé*   *ralámuli-chi*   *ko*

DEM   CONJ   rarámuri-LOC   DP

'here where the raramuri (live)'

'aquí en donde (estamos) los rarámuris'

b. *ayéna* *raíchi-ka*   *cho*   **échi**

CONJ   hablar-STAT   LOC   DEM

'also (she is) talking to them'

'también platicando con ellos'

The following examples illustrate the diversity of referential constructions available in

Tarahumara. It attempts to cover all kinds of referential expressions that are applicable to this

language, taking Table 5 (list by Ariel 1990 and Lehmann 2012) as a guide, and using the

discursive texts that were analyzed for this work as the source for data.

	Table 14. Referential mechanisms in Tarahumara.
a.	clause (completive, e.g. the Obj of a say verb) *ø* *ané-lta* *ruá ru pi* **napawí-sa** [A] say-APPL CIT CONJ [gather-COND] '(they) told them they should gather' 'les dijeron que se juntaran' (Tubali-Cl9)

b.	(common) noun phrase with head and (clausal or adjectival) attribute	

pachákie　　　*muchúa*　　　*tiwéli* ***láname***　　***sunú*** ***welí*** ***ramé-kame***
inside　　　　to:be.PL　　　there [yellow　　corn　big　kernel-NMLZ]
'inside was the yellow corn (with) big kernel'
'ahí adentro estaba el maíz amarillo (con) grandes granos'　(Tubali-Cl37)

c.	(common) noun phrase with head and lexical possessive specifier	

kipu ***rana*** *gará*　*kú*　　*sawí-li*　　　*lá*　　*ba*
[POSS son]　good　DP　　heal-PFV　　　　CL
'her son healed well'
'su hijo se curó bien'　　　　　　　　　　　　　　　(Trabajos-Cl114)

d.	(common) noun phrase with head and determiner	

jériká　　　*bila* *kó*　*nócha la*　***échi*** ***mukí***　　*ko*　　*bá*
like:this　　pues　DP　work　　[DEM　woman]　　DP　　CL
'this is (how) the woman works'
'pues así trabaja la mujer'　　　　　　　　　　　　　(Trabajos-Cl116)

e.	proper name	

Tubáli　　　*ábaga*　　　*riwéwi échi*　*rijói*
[Giant]　　　like:that　　call　[DEM　man]
'Giant was the name of that man'
'Gigante, así se llamaba ese hombre'　　　　　　　(Tubali-Cl1)

f.	bare common noun	

ø　　***napíwili***　　*batú-a*
[A]　[corn　]　　　grind-PRS
'(she) grinds corn'
'(ella) muele maíz nixtamalizado'　　　　　　　　(Trabajos-Cl7)

g.	emphatic pronoun	

*ø   machí **binói**   ko   mápu-ali   nayuu-sa   échi   taa*
[S]   know   same/self   DP   SUB-SIM   get:sick-COND DEM   small
'(she) knows (by) herself when the little (kid) gets sick'
'(ella) misma sabe cuando se enferman los pequeños'   (Trabajos-Cl97)

h.	neutral pronoun	

*mukí rarámuli   ko   a   bilá   **taá**   nócha-ka   rejói   cho   pa*
[woman rarámuri]   DP   LOC   DP   [1PL]   work-STAT   man   LOC   CL
'we, the raramuri women, work like men'
'las mujeres rarámuris, nosotras trabajamos como el hombre' (Trabajos-Cl2)

i.	clitic person form	

N/A

j.	bound person form (affix)	

*na   jaré iná-**lta**   ruá ru tibulua ka   chio*
PART   [some] burn-APPL   CIT   [last] EMPH   too
'some (of the giants) were burnt, it is said, the last one too'
'algunos los quemaron a los últimos también   (Tubali-Cl48)

k.	ZERO person form	

*ayéna cho   **ø**   newá kowa-ami   beá*
CONJ LOC   [A]   make [eat-NMLZ]   early
'(she) also makes food early'
'y también hace comida temprano'   (Trabajos-Cl9)

CHAPTER III

MATERIAL DESCRIPTION AND REFERENTIAL ANALYSIS

The linguistic materials used for this work are discursive texts that were taken from different sources. The text "*Trabajos de la mujer*" was documented in the city of Chihuaha by Dr. Zarina Estrada Fernandez during field work in September 2012, as part of her project entitled "Syntactic complexity and typological diversity in northwestern languages of Mexico" ("Complejidad sintáctica y diversidad tipológica en lenguas del noroeste de México"). It was transcribed and translated by Luis Sandoval, a bilingual speaker of Tarahumara and Spanish; the gloss and morpho-syntactic analysis, as well as the English translation has been proposed by Ana Tona.

The text "Tubali" was previously analyzed by Edgar Adrián Moreno Pineda in his master's thesis entitled "*Chabé nirúame natári: Análisis de los referentes en el discurso mítico ralámuli*" in 2013. The English translation for this text has also been proposed by Ana Tona. Additionally, some changes to the original gloss have been made in order to be consistent with the other text used in this work and for the sake of comparative referential analysis.

In sections 4.1 and 4.2 we will take a look to both texts in terms of discourse typology: their expected behavior and the challenges they posed in their analysis.

3.1. The text "Trabajos de la mujer"

The text "trabajos de la mujer" contains a total of 118 sentences[2]. The narrator is a Raramuri woman who was a teacher at Gabriela Mistral School located in Norogachi; she is now living in Pahuirachic, Chihuahua. In this text she talks about the Raramuri women, focusing on their duties within the community; some of the activities are well described, whereas others are only mentioned. The following text is the translation I have proposed for this text; it is presented here in order to better understand which entities appear and what is the relationship among them.

According to the taxonomy of texts presented above, the one used here belongs to the procedural text type. Although it has a logical linkage, rather than a chronological one (since it is explaining how the work and house chores are done), it shows all the other features of a procedural text:

1. Most of the references in subject positions are non-specific persons (see table 22: Referential type of references), as it talks about the Raramuri woman in general; they appear in the text either as bare/indefinite nouns, or as impersonal markers (-*wa*).

(28) a. **mukí**   **rarámuli**   ko   a   bilá
           woman   rarámuri   DP   LOC   DP

           taá nócha-ka   rejói cho   pa
           1PL trabajar-STAT man   LOC   CL
           'we, the raramuri women, live working like men'
           'las mujeres rarámuris pues nosotras vivimos trabajando como el hombre'

        b. échona   biti-chí   échi   ona   kawi-chí   ayé
           LOC   vivir-LOC   DEM   LOC   monte-LOC   CONJ
           'in the house, at the mountain'
           'en nuestras casas, en nuestro monte también'

---

[2] This text was analyzed by sentences, to be consistent and comparable with the text "Tubali" which was analyzed the same way: each sentence may contain from one to three clauses. The text "Trabajos de la mujer" contains up to 169 clauses.

c. mápu   ona   ø   ichi-**wá**   má
   SUB   donde [P]   sembrar-IMPRS   LOC
   'where we sow'
   'donde se siembra'

2. Patient orientation is reflected in the high number of overtly mentioned patients in the text (see table 4: References by weight and function).

Table 15. References by weight and function (PT)

	S	A	P
w=0	25	56	13
w=1	1	9	1
w=2	0	0	0
w=3	1	0	5
w=4	11	4	28
w=5	0	0	0
w=6	22	18	35
w=7	5	1	2
w=8	5	0	4
Overt Ref.	45	32	75

3. The present and customary tenses are used throughout the text, which in Tarahumara can be seen in the forms of the present marker –*a*, the habitual marker –*ya*, and the stative marker –*ka*.

(29) a. mápu-reká ø   beá   asísi-**ka**
        SUB-así   [S]   temprano   levantar-STAT
        'for instance, getting up early…
        'por ejemplo, levantándose temprano'

      b. biré mukí a   chukú   sakí   newa-**yá**,
        uno mujer LOC   AUX   esquite   hacer-HAB
        'a woman makes corn'
        'una mujer hace esquite (para hacer pinole)'

      c. ø   a   chukú   remé-**a**
       [A] LOC   AUX   tortilla-PRS
       '(she) makes tortillas'
       'está haciendo tortillas'

The description of text genres raises certain expectations in relation with the behavior of the text; however, as can be observed, not all expected characteristics are completely fulfilled. Some texts may present features from other text genres, and the decision of putting it into one genre or the other will depend on how strong are the features for one genre above others. In this case, the text presents a logical linkage rather than a chronological one, but this alone is not enough to classify it as a behavioral or expository text. Since the use of a non-specific person and the patient orientation of the text are strong features, it is best classified as a procedural text.

3.2. The text "Tubali"

This text has a total of 79 sentences[3], the narrator is a Raramuri woman from the Chínibo locality. She tells a myth about a one-legged giant who ate children from that same community. This text has its place within the group of narrative discourse:

1. Although there are some first person references, most of the references in subject position refer to a third person.

2. In this text, the subject appears as an experiencer (S or P) most of the times (see table 19). The most agentive references include the giant (only a few times) and a group of men that watched and took care of the people in the village.

Table 16. Distribution of functions (PT)		
Function	Human	Inanimate
S	59	17
A	31	0
P	17	16

---

[3] I have respected Moreno´s organization of the text, which is analyzed by sentences rather than individual clauses: one sentence may be simple or complex, containing from one to three clauses).

3. Throughout the text, the past and the (historical) present markers are used. As we have seen in the typological analysis of the previous text, the temporal markers take the form of the suffix –*ri* (-*li*, -*ti*) for the past tense. The present tense marker is –*a*, the progressive or habitual is –*ya*, and the stative marker is –*ka*.

(30)

    a. *Tubáli ábaga riwéwi éch   rijói,   riwá-la   niká   ka   iyén-ti*
       Gigante así   llamar DEM   hombre nombre-POSS así   EMPH andar-PFV
       'Giant was the name of that man, that was his name'
       'Gigante, así se llamaba ese hombre, ése era su nombre'

    b. *chabé ka   ma   ku   wiká-li   jípe   ka*
       antes EMPH ya   otra:vez   perder-PFV   hoy   EMPH
       'long time ago, today (he) is lost/gone'
       'hace mucho, hoy ya no está'

(31) *ke   ka   nokó-ya   ichúa-li   ko   ba   chió   ko'á-ya*
      no   EMPH mover-PROG sentar-PFV   EMPH CL   también comer-PROG
      '(they) didn't move / they were siting / (and) eating too'
      'no se movieron, estaban sentados y estaban comiendo'

3.3. Observations: difficulties in the interpretative process

Different discourse types embody a variety of challenges and levels of difficulty in discourse analysis. In the case of a procedural text, for instance, clause structures are very consistent, making referent identification and referential tracking an easy task. In the procedural text "Trabajos de la mujer", referential continuity is maintained throughout the text and newly introduced entities are then easily processed. A narrative text, on the other hand, presents a wider variety of clause structures where referents are treated in many different ways, giving place to more ambiguity.

As we have seen in section 2.5.2, the interpretation process in discourse implies a series of inferences that are possible due to the hearer's knowledge about the language and the world. In this matter, theory assumes that speaker and hearer (or in a given case, narrator and reader) share a series of knowledge about the world. It is therefore assumed that both of them have access to the same referents in such way that the hearer (or reader) can make the correct inferences, according to what the speaker (or narrator) wanted to communicate. But what happens when speaker and hearer are not members of the same socio-cultural group and do not have access to the same referents?

Even though it is possible to do a linguistic analysis solely based on linguistic data and pertinent theories, sometimes they are insufficient for *discourse* analysis. Occasionally, semantic interpretation (for instance, particpants' semantic role or referent identification) can be obscured by the unawareness of certain possibilities within a linguistic system and of specific sociocultural features of the communities where the languages in question are spoken. This unawareness can mislead the interlocutor so as to make false inferences. When a linguist is confronted with a text, his/her interpretation is limited by his knowledge of the language and the community where it is spoken. These inconveniences are not limited to the semantic aspect of linguistic analysis; they also affect syntactic analysis, as we will see.

During the process of interpretation of the myth "Tubali" some sentences emerged with different possible interpretations, where choosing one or another possibility would change the story entirely. Some of the ambiguities of this text are related to the identification of the referents, mainly, the agent of the event. In other cases the ambiguity was associated to the nature of the events or state of affairs.

3.3.1. Syntactic and semantic ambiguity

The ambiguities related to referent identification are generally connected to the use of zero anaphor. However, this is not the only source for ambiguity. It was previously mentioned that the discourse topic creates a scenario that limits the possibilities for interpretation (Saeed 1997). In the myth "Tubali", the discourse topic consists of a giant man that was supposedly cruel to the children. Taking this as a base for interpretation, one may predict that the giant will act, in the majority of cases, as an agent, an instigator of events. Hence, when we are exposed to a sentence where something is being destroyed or someone is being affected, one is tempted to infer beforehand that it is the giant who would be taking the role of the agent. Throughout the text, this inference was not always the most adequate. Contrary to these predictions, the giant appears in several constructions as an experiencer and sometimes even as a patient.

The following examples are two possible interpretations for the same piece of text. In (32a) it is inferred that the zero anaphor refers to the giant (referential index r1), who would take the functions of A and S in sentence (i), A and S in (ii) and again S and A in (iii). Thus interpreted, the myth is telling a story where the giant was able to walk even though he had only one foot, and that he destroyed everything as he was standing.

(32a)

i.
r1 A		r4 P		r1 S		
ø	*biré-pi*	*raláti,*	*pi*	ø	*simí-li*	*ka*	*riká*
	uno-RES	pie	CONJ		ir-PFV	EMPH	así

'(he had) only one foot, that's how (he) went'
'sólo un pie (tenía él) pues así iba'

ii.
r1 A			r5 P	r1 S		
ø	*nawó-la*	*ruá*	*kárga,*	ø	*rolomáchi*	*niwá-la*	*ruá*
	tumbar-APPL	REP	todo		mochar	hacer-APPL	REP

'he destroyed everything, he shattered (everything), they say'
'tumbó todo, dicen, hizo mochar (todo) dicen'

53

iii. r1 S   r1 A r6 P

ø *we* *rigá* *wilíga ruá ru,* ø *kárga rolomá-la* *ruá ru*
   mucho así estar REP   todo mochar-APPL REP
'they say (he) was standing (and) (he) shattered everything'
'dicen que estaba parado, (y) mochó todo'

Example (32b) shows a different interpretation of the events: in (i) zero anaphor refers once again to the giant (referential index r1), with the same functions as in the previous example. However, in (ii) both of the zero anaphora refer to the giant's lost foot (rather than the giant as an entire entity). The inference of the latter referent is enabled by the previous sentence, and takes a P function. In (iii) the first anaphor refers to the giant with an S function and the second anaphor refers again to his foot (the one that was cut off) with a P function.

The second interpretation tells a story somewhat different from the one shown in (32a): now the giant is walking even though he only had one foot, the other foot was cut off from his body as he was standing. In this version the giant loses all agentivity to become a syntactic P, and the victim of this part of the story.

(32b)
i. r1 A   r4 P   r1 S

ø *biré-pi* *raláti, pi* ø *simí-li* *ka* *riká*
   uno- RES pie CONJ   ir- PFV EMPH así
'(he had) only one foot, that's how (he) went'
'sólo un pie (tenía él) pues así iba'

ii. r6 P   r6 P

ø *nawó-la* *ruá* *kárga,* ø *rolomáchi* *niwá-la* *ruá*
   tumbar-APPL REP todo   mochar hacer-APPL REP
'(his foot) was completely destroyed, (it was) broken/cut off, they say'
'le tumbaron todo (el pie), se lo mandaron cortar, dicen'

iii. r1 S   r6 P

ø *we* *rigá* *wilíga ruá ru,* ø *kárga rolomá-la* *ruá ru*
   mucho así estar REP   todo mochar-APPL REP
'they say (he) was standing (when) (his foot was) completely cut off'
'dicen que estaba parado (cuando) (se lo) cortaron todo'

According to Givón (1995), topical referents tend to have a low referential distance, that is, they are constantly mentioned in the clause chain. In the myth "Tubali", however, this is not the case. Once the topic (the giant man) is introduced into discourse, it is not clearly mentioned again throughout the text, not even to reactivate this referent after a thematic change where other agentive entities are mentioned. This feature makes the interpretation process more complex, as the entity "the giant" is not the only implicit referent. Other implicit referents occur that could be equally agentive or equally patientive; in sum, they are all competing as possible referents. Table 17 shows the referential weight and distance of the discourse topic.

Table 17. Referential weight and distance of the discourse topic "the giant" (r1).

Mention	Referential weight	Function	Sentence
1.	6	S	1
2.	0	S	2
3.	7	A	3
4.	0	A	4
5.	0	S	4
6.	0	P	5
7.	0	S	6
8.	0	P	7
9.	**0**	A	**23**
10.	0	A	23
11.	**3**	S	**30**
12.	0	S	32
13.	0	S	32
14.	**8**	S	**58**
15.	0	A	59
16.	3	S	60
17.	1	P	60
18.	1	P	60
19.	8	S	61
20.	0	A	62
21.	0	S	63
22.	1	P	64
23.	0	S	65
24.	0	S	66
25.	0	S	66
26.	0	S	67

27.	0	S	67
28.	3	S	68
29.	0	S	68
30.	0	S	69
31.	0	S	69
32.	0	S	70
33.	0	S	70
34.	0	S	71

We can observe that the giant was overtly mentioned in sentence 1 and 3, then it only occurred implicitly until sentence 7. After sentence 7 other referents were mentioned and it wasn't until sentence 23 that the giant appears in discourse again, but not overtly. Even though there was a thematic change in discourse, the discourse topic was not properly reactivated (that is, with an explicit NP). The discourse topic is later overtly reactivated in sentence 30, but using only a pronoun *échi yobi* 'the same one'. No clear, full NPs are used until sentence 58. From this sentence on, the discourse topic is maintained active, mostly through zero anaphora, and it occurs along with other referents.

The correct inferences about the referents in this text can only be made after having an understanding of the deeds of the myth. Nevertheless, this is not an easy task, as we have already seen. Sentence 23 is shown in (33). Here we can see that it is only after understanding the myth in general that we can infer the referents for these anaphora.

(33)

	r1 A	r25 P				r1 A		r25 P
sinémi	*ø*	*kúchi*	*kalá*	*timútuka*		*ø*	*chumisá*	*kúchi*
principio		niño	bien	llevar-STAT			chupar	niño

'first (he) took a child , (he) suck the child'
'primero se llevaba un niño, chupaba al niño'

56

In this case, it is mostly through sociocultural knowledge that inferences are made. Anyone who has no access to this knowledge will not be able to make the right inferences, given that syntactic cues are not sufficient.

There are other instances of ambiguous sentences where knowledge of the world plays a role in making the right inferences. When the hearer lacks the necessary knowledge of the world, the text can only be interpreted literally, as in (34a). Thus interpreted, sentence (34a) presents a situation in which a group of men are warning other men about something being burnt, and summon them to gather in order to do something about it. However, this interpretation raises many doubts about the referents; neither the agent nor the patient are explicitly mentioned, except for an applicative mark. Then, what is burning and why? Doubts persist after revising the discourse context: before this sentence there are no hints of this kind of scenery, and in the following sentences there are no mentions of damages caused by fire, nor is it mentioned again in the story. Clearly, there has to be an interpretation error, or the text is not very coherent. I am inclined to think there is an interpretation error, since the speaker is less likely to produce such semantic faults.

(34)
    a. *ø*   *ané-lta*   *ruá ru pi*   *napawí-sa*   *mina*   *ikoú-lta*   *ruá ru*
           decir-APPL   REP   CONJ   juntar-COND   allá   quemar-APPL   REP
           'they say they told them to gather, (something) was burning over there'
           'dicen que les dijeron que se juntaran, allá se estaba quemando (algo)'

Once one is familiarized with some cultural and world facts about the Tarahumara region, it is possible to make some sense out of this sentence, which is actually somewhat idiomatic. In this region, the weather and minerals of the earth contribute to mummify corpses by a natural process. In the view of the Tarahumaran people, the aspect of a mummy is the same of someone

57

who has been burnt, hence the expression of "something being burnt", which actually refers to

dead people. The correct interpretation of the sentence would be as shown in (34b).

(34)
    b. *ø*   *ané-lta*   *ruá ru pi*   *napawí-sa* *mina* *ikoú-lta*   *ruá ru*
            decir-APPL   REP   CONJ juntar-COND allá   quemar-APPL REP
          'they told them to gather, as there were dead people'
          'les dijeron que se juntaran, que allá había quemados (había muertos)'

With enough cultural information and world knowledge available, we can now make the

correct inferences about the meaning of the sentence and the true course of the story that is being

told.

CHAPTER IV

REFERENTIAL DENSITY IN TARAHUMARA

In this section I present a referential analysis of two texts in Tarahumara, as well as a comparison between them in terms of referential behavior and information management. The analysis seeks to answer some questions about referentiality in discourse:

i.   What does RD say about informativeness in Tarahumara?

ii.   How is referential information managed in Tarahumara?

iii. How does information structure relate to discourse genre?

Furthermore, I propose to take into consideration some additional phenomena in the analysis of discourse, specifically the functions performed by elements of narration such as direct speech. In some texts it is necessary to account for direct speech, whereas others can leave them out of the referential analysis of the discourse, according to each discourse/text genre. In addition, I propose to take into consideration headless noun phrases in the scale of complexity of referential expressions, since they are not accounted for in the one used for this work.

4.1.   Methodology

The methodology used in this study is based on the Referentiality Project started by Lehmann in 2012. The criteria for the identification of referential expressions and data management are very similar to the ones used for the Referentiality Project, whereas name tagging or labeling may differ from it. Differences originate from the nature of both studies: the Referentiality Project is a crosslinguistic study that is designed to include many typologically different languages, as well

as texts of different sorts. This work, by contrast, is a comparative study among two different text genres of the same language. Thus, tagging differences seek a simpler, more transparent terminology and comprehensive paradigms that adapt to the language in question, in this particular analysis.

Referential density values were obtained by identifying, classifying and counting all referential elements, following the list of possible referential expressions from Lehmann (2012), which include: noun phrases, pronouns, verbal indexation and deictic elements that denote physical entities (people, things, places). It also considers abstract entities, which are entities with no physical existence like *knowledge* or *soul*, temporal location and manner, when expressed as entities like 'one day/October/the celebration', 'this way/many ways'. All sorts of predicates such as verbs, adjectives and nominal predicates were excluded, provided that they constitute non-referential expressions, according to Lehmann's approach (see section 2.3).

Each reference or referential expression received several tags to account for different features it may have in relation to other references within the clause and throughout the entire text. Firstly, references were assigned an identification number called *referential index*. Only first mentions of entities receive a new referential index, elements that are coreferential to a previously introduced referent receive the same (corresponding) referential index. However, two generic NPs may have the same referential index, as they refer to the same group. Two non-specific references should have different referential index, as they refer to random members of the same group, which may or may not be the same in a given situation. In the case of non-specific references, the linguistic system is open for both possibilities of semantic representation, that of same or different referent.

References are then tagged by their *referential type*. This tag classifies all references according to the degree of individuation, which is manifested by means of linguistic coding (e.g. presence/absence of determiners, etc.). This classification is based on the individuation scale shown in table 1 (section 1.2.1); therefore, available tags for referential type are:

Table 18. List of tags for referential type

a. Generic reference (gen)

b. Non-specific reference (nspec)

c. Specific reference (spec)

d. Unique reference (uni)

Referential expressions also received a tag that corresponds to their *semantic type*. The semantic type of references was defined by its place within the animacy hierarchy (by Croft 1991), but adding other kinds of entities that are being accounted for in this study (e.g. physical and temporal location, manner, textual deixis). This expansion of possible semantic types renders the following possible tags:

Table 19. List of tags for semantic type

a. human (hum)
b. animate (anim)
c. inanimate (ina)
d. physical location (loc)
e. temporal location (temp)
f. manner (man)
g. textual deixis (text)
f. action

Additionally, references were characterized by their *syntactic function* within the clause. Available functions are:

Table 20. List of tags for syntactic function

a.	Subject (S)
b.	Agent (A)
c.	Patient (P)
d.	Oblique (OBL), including benefactive, theme, instrumental, comitative, origin, recipient
e.	Locative (LOC)
f.	Time (TM)
g.	Manner (MN)

Finally, referential expressions received a *referential weight* value, depending on their complexity or explicitness degree as explained in section 2.3. Table (21) shows one example for each level of explicitness in Tarahumara, (see table 14 in section 3.7 for complete sentences) along with their corresponding weight values.

	Table 21. Referential weight in Tarahumara		
	Explicitness degree	**Example in Tarahumara**	**Referential weight**
l.	clause (completive, e.g. the Obj of a say verb)	*napawí-sa* gather-COND	9
m.	(common) noun phrase with head and attribute	*láname sunú welí ramé-kame* yellow corn big kernel-NMLZ	8
n.	(common) noun phrase with head and possessive specifier	*kipu rana* POSS hijo	7
o.	(common) noun phrase with head and determiner	*échi mukí* DEM mujer	6
p.	proper name	*Tubáli* Giant	5
q.	bare common noun	*napíwili* corn	4

r.	emphatic pronoun	***binói*** same/self	3.5
s.	neutral pronoun	***taá*** 1PL	3
t.	clitic person form	N/A	2
u.	bound person form (affix)	*iná-**lta*** burn-APL	1
v.	ZERO person form	***ø*** [A]	0

In sum, each reference was tagged for referential index, referential type, semantic type, syntactic function and referential weight. Example (35a-c) illustrates how referential expressions were tagged within each clause throughout both texts.

(35) a. | ***mukí*** | ***rarámuli*** | *ko* | *a* | *bilá* | ***taá*** | *nócha-ka* | ***rejói*** | *cho* | *pa* |
|---|---|---|---|---|---|---|---|---|---|
| woman | rarámuri | DP | LOC | DP | 1PL | work-STAT | man | LOC | CL |
| r3 | | | | | r4 | | r5 | | |
| gen | | | | | nspec | | gen | | |
| hum | | | | | hum | | hum | | |
| S | | | | | S | | OBL | | |
| 8 | | | | | 3 | | 4 | | |

'we, the raramuri women, live working like men'
'las mujeres rarámuris, nosotras vivimos trabajando como el hombre'

b. | *échona* | ***biti-chí*** | *échi* | *ona* | ***kawi-chí*** | *ayé* |
|---|---|---|---|---|---|
| LOC | live-LOC | DEM | LOC | hill-LOC | CONJ |
| | r6 | | | r7 | |
| | gen | | | gen | |
| | loc | | | loc | |
| | LOC | | | S | |
| | 4 | | | 4 | |

'in the house, at the mountain too'
'en las casas, en el monte también'

c.	*mápu*	*ona*	*ø*	*ichi-**wá***	*má*
SUB	where	P	SOW-IMPRS	LOC
		r8	r9	
		gen	nspec	
		ina	hum	
		P	A	
		0	1	

'where one sows'
'donde se siembra'

4.1.1. Methodological problems: headless phrases and direct speech.

During the classification and counting process of referential expressions, there were two main referential devices that made the analysis somewhat complicated: the occurrence of headless phrases and first person pronouns (direct speech). Neither one had a high percentage of occurrences; nevertheless, their treatment and inclusion/exclusion in the analysis invariably has an impact on the results of referential density values.

4.1.1.1. Headless phrases

While examining the referential mechanisms of the texts, there emerged some cases where an argument of the verb was implicit (zero anaphor) but its modifiers remained. In other words, there were cases of headless phrases. According to Croft (1996), the head of a phrase is a word "whose form-class is the same as that of the phrase". In a sentence like *all this fresh milk*, the head of the phrase is *milk*. A phrase that somehow refers to a noun but lacks a word of that category is then a headless noun phrase. In the texts we can find constructions like the ones illustrated in examples (37) and (39) below where arguments are omitted, leaving behind modifiers like adjectives and quantifiers, which do not denote properties. Instead, the property

denoted is specified for the entity omitted, and the presence of that entity is contextually understood.

With the modifier in place, and knowing that it belongs to a noun which can be tracked in discourse and related to that modifier, the referential weight of such referential mechanisms cannot be easily assigned. Even though there is no overt noun, such referential expressions raise the possibility for zero anaphora to have a referential weight different than zero. However, they could not take the same referential weight as those referential expressions with an overt noun (a head) and its modifier because they lack the head.

The following is an example from the text "Trabajos de la mujer". In (36) we can see an overt referential expression with a demonstrative, an adjective and its head referring to "the little boy". We can compare this example to the one in (37), where the noun is absent.

(36) *mápu-alí*   *ropa-chí*   *rikáachi*   *okó-sa*
       SUB-SIM   estomago-LOC   ocasión   doler-COND

       échi   taa   towí, *taa*   *tewé*   *kere*
       DEM   pequeño   niño   pequeño   niña   POSP
       when (his/her) stomach hurts sometime the little boy, (or) a little girl'
       'cuando le duela la panza a este niño (o) a una niña'

(37) *mápu-alí*   *échi*   *náli*   **taa**   **ø**   *majá-sa*
       SUB-SIM   DEM   PREP   [pequeño]   temer-COND

       siné   *kaachi*   *biléana*   *bají-chi*
       algunos   ocasión   INDEF   aguaje-LOC
       'when a little (kid) gets scared sometime at a baji'
       'cuando un pequeño (niño) se asusta alguna vez en algún aguaje'

There were 10 instances of headless referential expressions in the text Tubali. Example (38) is one of them; here the noun comes with the adjective *kúchi* 'small', and the quantifier *jaré* 'some'. Then we have the headless phrase on (39), which also comes with a quantifier.

(38) *pi* **kúchi** **jaré** **ba** **ochúla** ka ø *ochúla rítea*
      CONJ pequeño algunos CL huesos ENF [A] huesos ver
      'and some small bones,(I) have seen bones'
      'y algunos huesitos, huesos es lo que vi'

(39) **wiká ø** *iníla* *pi* *ichúa* *ta* *kalí* *itíwela* *bo*
      mucho junto CONJ estar.PL chica casa como:nosotros CL
      'many (bones) were in a small house like ours'
      'muchos (huesos) juntos estaban en una casita como la nuestra'

These modifiers do not function as predicates in these constructions, and they certainly do not stand alone. They appear in an argumental position (or more exactly, by a zero anaphor), but are not nominalized either.

Headless phrases like these ones are not included in the list of referential mechanisms in Table 5 (scale of complexity of referential expressions); only noun phrases with head and different kinds of overt nouns are taken into consideration, with the exception of zero anaphora. For this reason, headless phrases were simply counted as zero anaphora in this study, only to be methodologically consistent.

4.1.1.2. Direct speech

Although they were equally analyzed and received a referential index, deictic personal pronouns that refer to the speaker (what I call here direct speech) were treated differently, depending on their function within discourse: they were left out of the computation for the procedural text "Trabajos de la mujer", since they scarcely occurred and do not form part of the procedural text *per se* (36a). In this procedural text, direct speech was only used at the beginning of discourse as a means for introducing the subject or topic. In the myth "Tubali", however, direct speech is used as a means to make evaluations related to the story (36b).

66

(36) a.

jípi	ko	naa	mi	che	ø	raíchi-ma	la	bá
today	DP	LOC	LOC	LOC	1SG.S	talk-IRR	DP	CL

échi	mukí	ralámuli	kítira
DEM	woman	rarámuri	POSP

'Today I will talk about the raramuri woman'
'Hoy voy a hablar sobre la mujer rarámuri'   (Trabajos-Stnc1)

b.

ijínti	ko	ba	laya	ijínili	**nijé**	ka	we	uchúa	opória
maybe	EMPH	CL	think	maybe	1.SG	EMPH	very	be	grave

'I think maybe there is a grave'
'Tal vez, yo pienso, ahí está una tumba'   (Tubali-Stnc43)

Direct speech is also used to defend the relevance of the events described in the myth into the present, as shown in (37).

(37) a.

ø	napeti	olá	ko'á	itú,
A	together	make	food	quick

ø	jemi	biré	nomi	basóli	kosá
A	DEM	one	place	pozole	eat

'together (they) quickly made food, (they) ate pozole somewhere'
'juntos hicieron comida rápido, en algún lugar comieron pozole'
          (Tubali-Cl52)

b.

chiyó	nokwá-la	tamujué,
also	make-apl	1.PL

ø	ko'á	warú	biré	nomi	basóli	kosá
A	eat	much	one	place	pozole	eat

'we make (food) too/ we eat much pozole, somewhere'
'nosotros también hacemos comida/comemos mucho pozole'
          (Tubali-Cl53)

Consequently, in the text "Tubali" first person references were not left out of the computation when relevant to the story or the community traditions (as in 37), but they were set aside when used to make evaluations (as in 36b). Additionally, complement clauses of saying verbs were rather analyzed as separate/independent clauses whenever such verbs constituted reportatives (e.g. *people say he came late that day*), as in (38).

(38) a. *[chió ø   napawíp-sa   éna]* ***ruá***   ***ko***   ***ba***
        also S   gather-COND go   REP   EMPH   DP
        '(they) were also together, that's what they say'
        'también andaban juntos, es lo que dicen'   (Tubali-Cl21)

   b. *[boná-ka   mawé-ru   ba]* ***ania*** ***ani*** ***ochérame*** ***ka***
        rip:out-STAT cut-PASS   CL   say   say   old:people   EMPH
        '(it was) ripped out and cut, old people say'
        '(fue) arrancado y cortado, dicen los ancianos'   (Tubali-Cl45)

   c. *[chabé ka   ø   napawí-ka   rijóla]* ***ruá ru***
        befor EMPH S   gather-STAT   man   REP
        'it is said that long ago men (were) together'
        'antes todos estaban juntos, dicen'   (Tubali-Cl73)

This kind of reportative forms were very frequent throughout the text "Tubali" and counting them as one reference, without analyzing its arguments, would leave out much of its real content (content concerned with the events and referents of the story that is being told).

4.2. Analysis of the text "Trabajos de la mujer"

In the procedural text "Trabajos de la mujer" there are a total of 293 references of different kinds: human, inanimate, spatial location, temporal location, nouns that refer to manner (e.g. this way), and nouns that refer to actions (e.g. the celebration, work). There was also textual deixis, as in "they do *these things* (mentioned before)".

4.2.1. Referentiality in a procedural text

Most references are non-specific in this text, which is due to the type of text, as we have seen in section 3.1. The only specific references are deictic pronouns that take their meaning from

68

situational context or from discourse. The four specific references that are reflected in table (22) (Referential type of references) correspond to textual and locative deixis.

Table 22. Referential type of references (PT)

Ref Type	Occurrences	Percentage
specific	4	1.4
non-specific	198	67.6
generic	90	30.7
other	1	0.3
Total	293	100

References for human and inanimate referents are of main importance here, since they are the kind of referents that typically occupy central positions. They can also appear in non-central positions (recipient, benefactive, theme, instrument, comitative) that could, at some point, take a central role in discourse through diverse syntactic mechanisms. Textual references are also important, as they tend to take central positions, completing the meaning of the clause. References that mention location, time and manner are of adverbial nature and are hence not relevant in the same way; however, their importance in the communicative process must not be underestimated, as we will see (in 4.2.2).

Out of the 258 (246+12) argumental referential expressions, 152 were human, 104 were inanimate, and 2 were textual references (see table 23: Semantic type of references).

Table 23. Semantic type of references (PT)

Sem Type	Ocurrences	Percentage	Total
human	152	51.9	
animate	0	0	
inanimate	104	35.5	
text	2	0.7	258
location	5	1.7	
temp loc	21	7.1	
manner	2	0.7	
action	7	2.4	35
		100	293

Human referents come first in the animacy hierarchy; thus, they are usually the ones in S and A positions. Likewise, inanimate entities tend to be patients. Both assertions are attested in table (24):

Table 24. Distribution of functions (PT)[4]

Function	Human	Inanimate	Total
S	47	18	65
A	86	2	88
P	14	70	84
REC[5]	1	0	1
BEN	1	0	1
COMIT	0	2	2
THEME	1	3	4
INSTR	0	2	2
ORIGIN	0	1	1
	151	98	249

4.2.2. Function: the importance of adverbial expressions

As can be seen in table (25), out of the 293 references, 246 were central arguments of verbs constituting 84% of references, in sum, which is not unexpected. Only 12 references were non-central arguments, and 35 were adverbial expressions (location, time or manner).

Table 25. Function of references (PT)

Function	Ocurrences	Percentage	Total
S	70	24	
A	88	30	
P	88	30	246

---

[4] There are 6 inanimate references and 1 human reference that function as LOC. Also, there are 4 Ps and 5 Ss that are neither human nor Inanimate, hence the discrepancy in the totals of humans and inanimates, and in the counting of S and P

[5] Oblique arguments are separated by type and Adverbial references are left out of this table.

OBL	12	4.1	12
LOC	13	4.4	
TM	20	6.8	
MN	2	0.7	35
		100	293

The distribution of adverbial expressions should not be overlooked: 57% of adverbial references expressed time location, which is very important in a procedural text where the ordering of ideas in a logical, consecutive or chronological way is crucial to the understanding of the procedure described.

4.2.3. Referential density: topicality

The density and distribution of the references points to a clear difference in status among them. "Raramuri woman" is by far the most persistent referent from the beginning to the end of the text, with 102 mentions, as can be seen in table (23). It is followed by a group of referents that are maintained active in small sections of the text. The number of mentions for these referents ranges from 5 to 19. All the remaining references in the text have only four mentions or less.

Table 26. Main Referents of the text (PT)

	implicit	explicit	total
woman	68	34	102
child	4	15	19
corn	4	11	15
batali	1	14	15
sanali	6	9	15
healer	4	7	11
IMPRS	0	7	7
basiawi	2	5	7
food	0	5	5

As we have said in section 1.3.2, discourse topics are referents maintained active throughout discourse by diverse devices (Saeed 1997). The more active a referent is in discourse, the less linguistic material it requires. This fact is reflected in the number of explicit versus implicit references in table (23): more than half of the mentions for "woman" are implicit (that is, zero anaphor). According to Givón (1995: 51), zero anaphors and unstressed pronouns are the unmarked coding choice for discourse topics. They code currently activated referents.

The entities with which the "raramuri woman" interacts as the Agent correspond in most cases to the sentence topic. The group of referents with 5-19 mentions is more salient than the rest of the referents in the text, but less salient than the discourse topic (the raramuri woman). It is worth noticing that these referents (child, healer, corn, *batali*, etc.) are cultural objects, entities that are present in the reality of every raramuri woman. She interacts with some of these entities on a daily basis, while other entities have a special/ritual value (like the *sanali*) and belong to specific occasions/celebrations.

Thus, the "the raramuri woman" is unquestionably the *discourse topic* of this procedural text. The entities affected or manipulated by "the raramuri woman", on the other hand, are the *sentence topic* or focus of individual clauses and sentences. The focus is usually the new argument introduced in a clause (Givón 1983). Although the raramuri woman is the topic and main actor of almost every activity described, the entire text is actually focused on her work: the activities she does, the objects she produces and the things she knows. Let us remember that, as part of a co-operative society, her importance as woman is defined by her work and participation within the community (see chapter 2, section 2.1, cultural and environmental context).

4.3. Analysis of the text "Tubali"

In the narrative text "Tubali" there are a total of 166 references of different semantic types: human, animate inanimate, spatial location, temporal location. This text lacks nouns that refer to manner, nouns that refer to actions and textual deixis. However, the presence of reportatives is abundant and first person pronouns are also relatively frequent.

4.3.1. Referentiality in a narrative text

In this text again, more than half of the references were non-specific. Out of the 51 specific references, 34 belong to the main character of the story, the giant man. 6 of them correspond to textual deixis, and the remainder corresponds to other entities mentioned in the story. In a narrative text, one would expect more characters to be specific; however, this is not a story about one giant and one hero or one victim. The myth tells a story about a group of brave men (not one specific person) that get involved in a battle against the giant, and at the same time, they look after the women and children of the community. Here, we can once again perceive the co-operative nature of this community. Generic references were used to mention entities like men, bones and food.

Table 27. Referential type of references (NT)

Ref Type	Occurrences	Percentage
specific	51	30.7
non-specific	97	58.5
generic	18	10.8
other	0	0
Total	166	100

73

As was previously mentioned (in section 1.5.2, typology of texts), the expectation of having an agentive topic is not fulfilled in this text. The narrator mostly speaks about the topic (the giant) and other characters in a very neutral way, focusing more on the state of affairs then on the actions perpetrated by them. This tendency is reflected in the high number of occurrences of S (intransitive constructions), as opposed to A.

Table 28. Distribution of functions (NT)

Function	Human	Inanimate
S	59	17
A	31	0
P	17	16
REC	1	0
BENEF	0	0
COMIT	0	0
THEME	0	0
INSTR	0	0
ORIGIN	0	0
Total	108	33

Naturally, humans play a main role in this narrative, which is manifested in the percentage of human references (65.1%). Even some of the inanimate entities are also related to human references, as they refer parts of the body such as feet, blood and bones. This is not to say that all inanimate referents are like that, other sorts of entities (like corn, grave) are also mentioned. However, all interactions described in the text happen between humans, and other entities are part of the context.

Table 29. Semantic type of references (NT)

Sem Type	Occurrences	Percentage	
human	108	65.1	
animate	2	1.2	
inanimate	33	19.9	143
text	6	3.6	
location	16	9.6	
temp loc	1	0.6	
manner	0	0	
action	0	0	23
Total	166	100	

As was conveyed in the analysis of the procedural text, contextual or background information should not be overlooked. It is worth noticing in this particular story, that when we compare table 28 and 29, adverbial elements emerge as significantly more relevant than oblique arguments (receptor, benefactive, instrumental, etc.).

4.3.2. Function: the importance of adverbial expression

Locative references are very important in a narrative text, since every story usually happens in a specific setting or context. Moreover, spatial location in this myth is very particular since the story that is being told supposedly happened in that same place, or very near. For this reason, in addition to having nouns referring to places (like house, barn and cave), we can find demonstratives like *jémi patú* 'up there'.

Spatial location is in fact relevant in the understanding of the events as the characters move around. References of spatial location represent 9.6% of all references in the text, and 70% (16/23) of references other than human, animate and inanimate (the ones that can take central positions), as can be observed in the data presented in table (29) of the previous section.

Temporal location would usually be part of the setting of a narrative text. Nevertheless, this myth is narrated simply as something that occurred a long time ago, and the narrator places no emphasis on time. She rather makes a contrast between that time and the present, especially in the beginning and at the end of the story. No nouns are used to code the past, only the adverb *chabé* 'before (in relation to another event)' is employed. The only temporal reference reflected in the table corresponds to the temporal locative *jípe* 'today' (meaning 'nowadays', in this context).

(39) *Tubáli ábaga riwéwi échi   rijói,   riwá-la   niká   ka   iyénti*
Giant so   call   DEM   man   name- POSS like:this   EMPH   andar
'Giant was the name of that man, that was his name'
'Gigante, así se llamaba ese hombre, ése era su nombre'

(Trabajos-Cl1)

chabé ka   ma ku   ø   wiká-li   jípe   ka
before EMPH   now again   S   lost- PFV   today   EMPH
'long time ago, today he is gone'
'hace mucho, hoy ya no está'

(Trabajos-Cl2)

4.3.3. Referential density: topicality

The myth "Tubali" revolves around the giant, and it is seemingly the most mentioned referent. However, it is very closely followed by a group of referents who are the actual agents of the story; this group of referents comprises "men", "watchmen" and "pascolas", which add up to 39 mentions. The indefinite codification of referents makes it difficult to make a clear distinction between these three groups, consequently their members seem to overlap as the story progresses.

Table 30. Main Referents of the text (NT)

referent	implicit	explicit	total
giant	25	9	34
men	14	10	24
watchmen	9	1	10
towns people	7	2	9
child	4	3	7
house	0	6	6
pascolas	3	2	5
bones	3	2	5

Town's people and children are the next most mentioned entities in the story; they are the ones being protected by watchmen.

4.4.   Contrastive analysis

Differences and similarities between both discourse types can be found in various dimensions of texts that exist, in one way or another, in every discourse type. These dimensions of discourse composition can be intuitively identified and include: self-identity establishment, entity construction, setting construction and discourse focus argumentation.

Self-identity establishment: is similar to what is commonly known as voice. The establishment of self-identity occurs when the narrator identifies him-/herself as an entity that is either involved in or external from the narration, giving hints about his/her position or view in relation with the discourse topic or focus. Self-identity can be overtly established in the beginning of discourse as the narrator introduces him-/herself; however it can also be established throughout discourse, without any introductions, by using mechanisms (e.g. 1/2/3 pronouns, impersonal markers or passivization) that will help the interlocutor infer the view or position that the narrator has taken.

Topic/entity construction: the narrator introduces the entities that will be the discourse topic, as well as other relevant referents. The descriptive complexity of each reference will vary according to the discourse goal or purpose. Some entities may be carefully described through highly complex reference mechanisms, whereas others may be barely mentioned. Another aspect to take into consideration is the (in)consistency with which entities are referred to: in some discourse types, one entity may always be named the same way: in a given text, "the queen" may consistently be referred to as "the queen" from the beginning to the end. However, a variety of referential mechanisms and naming choices are available in every language in order to fulfill pragmatic necessities; thus, in another text, "the queen" may also be referred to as "the beautiful queen", "her majesty", "the king´s wife", etc.

Setting/context construction: the construction of setting only refers to the location of entities in space and time. The importance of tempo-spatial location might differ from one discourse type to another: in one discourse type spatial location can be of great importance, and time may be almost irrelevant (this may apply to narrative texts); the opposite may be true for a different discourse type (for instance in procedural texts). In a given context, it could also be the case that neither spatial nor time location are crucial (some expository and behavioral texts).

Discourse focus/goal argumentation: as the discourse progresses, the narrator defines the real theme, the discourse focus (on top of the discourse topic as an entity), by leaning the attention towards one aspect of the discourse content. The discourse can focus on the entities, their importance in a particular context or situation, their identity, nature, characteristics or level of affection. On the other hand, it can focus on the facts, events or states described (frequency, intensity, impact or manner). It can also turn its attention on the processes involved by making emphasis on the knowledge required, the methodologies involved and level of complexity.

78

Finally, it can focus on the product such as changes, consequences, teachings or traditions which are the result of the events or actions perpetuated by the entities involved.

The discourse focus is directed by the discourse goal. Whether the goal is to transfer knowledge, communicate some fact or situation, present a concept or entity, give instructions or advice, convince the interlocutor to do something, or simply tell a story, the narrator will take every opportunity to make it clear. Consequently, all information given within discourse is relevant, and the mechanisms used to render this information are not random.

In each one of these dimensions that comprise discourse, the way that the speaker choses to handle information is important. This means that the inclusion/exclusion of certain reference types and the high/low frequency of specific coding mechanisms have a purpose. These dimensions of discourse are thus reflected in syntax (referential mechanisms) and information management (referential density, informativeness, etc.).

The manifestation of these dimensions of discourse is observable and comparable in the two texts used in this work. Tables 31-34 show the mechanisms used in each text, along with its functions, as a manifestation of these discourse dimensions.

Table 31. Discourse dimension: establishment of self-identity.

Dimension		mechanism	function
Self-identity establishment	Procedural text Trabajos de la mujer	use of 1p as introduction and self-identification in the beginning of discourse use of impersonal marker –*wa* to generalize throughout discourse	The narrator identifies herself as part of the group which will be the topic. This procedural text is an explicative one; it is based on personal experience but is intended to emphasize the group.

		use of 1p as self-identification and in modal constructions in some sections of the discourse	The narrator identifies herself as part of the community to which the myth is relevant. She makes evaluations about the entities involved.
Narrative text Tubali			
		use of citatives (38% of clauses) throughout discourse	As a mythical tale, it is based on tradition or collective knowledge; the narrator seeks validation through citatives.

Table 32. Discourse dimension: entity construction.

Dimension		mechanism	function
Entity construction	Procedural text Trabajos de la mujer	Use of non-specific references (67.6%)	As a procedural text, it does not focus on a particular entity or event, it speaks entities the entities that are generally involved in daily routine.
		Consistency in the coding/naming of referents	Being an explicative text, it seeks clarity, facilitating reference tracking.
	Narrative text Tubali	Use of non-specific references (58.5%) Specificity for the topic	As a narrative (mythological) text, it does talk about one specific entity (the topic), and other characters whose identity is unknown.
		Diversity in the naming of coreferential entities.	It pursues expressive richness and adds connotative values. It also avoids repetition, sometimes obscuring reference tracking

Table 33. Discourse dimension: setting construction.

Dimension		mechanism	function
Setting construction	Procedural text Trabajos de la mujer	Temporal references (7.1%) Locative references (1.7%)	This type of text requires a chronological order of events. It describes the adequacy of specific activities according to the moment (time of the day, season, and occasion).
	Narrative text Tubali	Temporal references (0.6%) Locative references (9.6 %)	In a myth, time is not of particular importance, it is only coded to establish a parallelism between past and present. It does require spatial location, given that the entities of the story move to different places. Spatial location is also especially relevant in this case because the story develops in places known to the community and the narrator.

Table 34. Discourse dimension: discourse focus.

Dimension		mechanism	function
Discourse focus	Procedural text Trabajos de la mujer	Reactivation of referents through explicit mechanisms High frequency of A (57%)	This text is focused on the agentive nature of the topic. However, the overall focus of the discourse is on the processes and knowledge required to perform the activities described.
	Narrative text Tubali	Use of zero anaphor even for referents that are expected to be explicitly reactivated (high referential distance) High frequency of S, that is, intransitive clauses(54.6%).	The interlocutor is expected to make many inferences about entities. The discourse is focused on states and the nature of entities.

After revising all the information given by the data and the analysis, we can now go back to our initial questions, which served as a guide for this work:

   i.   What does RD say about informativeness in Tarahumara?

  ii.   How is referential information managed in Tarahumara?

 iii.   How does information structure relate to the typology of text genre?

Along this work, RD values were used to place the two texts analysed within a typology of texts, and observations about referential information and informativenes were made through referential and semantic types of references, and their referential weight. Information structure was observed through the identification of the topic and other referents in both texts. This suggests that referential density, referential information, and consequently information structure, are all related, in a way, to the typology of text genre. Referential information is managed differently in each text, as each one of them has different pragmatic intentions.

CONCLUSION

Throughout this work we have observed the details and complexities of referential analysis (mainly informativeness and information structure) in discourse by means of referential density and semantic analysis. Several factors may take part in making a thorough referential analysis, so as to exploit the totality of the data while being loyal to the communicative purposes of speakers and the structural properties of the language. We have also seen that even though many good referential studies have been produced, studies in this area and discourse studies in general still have some problems that need to be addressed.

One proposal that is made in this work is considering different text genres of one language, at different levels of analysis, including syntactic, semantic and pragmatic aspects of discourse. It is through the comparison of different text types of the same language that generalizations can be made. For typological studies, having a clear view of the behavior of discourse elements within languages is central to the advancement towards a contrastive analysis crosslinguistically. Nevertheless, it is easier said than done, and several factors contribute to prevent the production of full systematic discourse studies. Disregarding or underestimating some information that emerges in discourse is one of these factors.

Cultural information and world knowledge has traditionally been left out of discourse analyses, or have been relegated to footnotes or small annotations to clarify the meaning of a word or idiomatic sentences. This study demonstrates that such information is valuable in discourse analysis (see section 3.3). Let us remember that languages are not isolated systems that exist for their own sake: languages are tools used by speakers with specific communicative purposes in particular contexts.

Additionally, not much has been said about the treatment of direct speech and headless phrases in discourse analysis. As we have seen in chapter IV, both phenomena can complicate the analysis of a text and affect the results of referentiality studies. It is important to carefully analyze these phenomena along with others that have already been included in discourse studies. In the case of direct speech, I propose to take a closer look at the functions performed by direct speech in specific discourse genres of particular languages. All languages behave differently as they serve communicative purposes within different socio-cultural contexts. For this reason, it is impossible to affirm that X discourse genre behaves in a particular way in all languages. Hence, the decision of including or excluding direct speech references may not completely rely on the discourse genre. Ideally, each discourse must be analyzed not only in terms of syntactic structure and referentiality, but also in terms of cultural context, keeping in mind that cultural information implies specific world knowledge and perhaps different views or interpretations of that knowledge. Moreover, communicative necessities vary depending on socio-cultural properties, which may lead to a diversity of features in discourse genre across languages.

Regarding headless phrases, for the sake of consistency with the methodology and theoretical framework, they were counted as implicit references or zero anaphora in this work. However, I consider that they should be accounted for along with the other overt referential expressions. In my view, adjectives and quantifiers that are left in a headless referential expression can function as specifiers. Filling an argumental place with a modifier that is more appropriate to one specific referent, as opposed to other active competing referents, helps in the identification or tracking of referents. Therefore, headless referential expressions should be analyzed as *explicit* references, with a referential weight greater than zero.

Other components of languages that are typically overlooked in discourse studies are adverbial expressions. Locative and temporal expressions have been carefully studied in many languages; nevertheless, they always manage to stay out of discourse studies. Locative and temporal expressions are generally assumed to be coded by adverbs and other mechanisms different than nouns. Consequently they are not considered as referential expressions. This presupposition is not always correct, locative and temporal notions may well be expressed by various mechanisms, including nouns. In languages where this is true, locative and temporal expressions should be included in referential studies, as they do refer to things: places in space and places in time. Additionally, they are always regarded as non-central information in clauses as well as in discourse. This non-centrality is correct at a clause level, as they normally fill oblique positions; very few predicates require locative expressions as central arguments. However, locative and temporal information can be very important in discourse, as we have seen in chapter IV.

In sum, this work intends to instigate the expansion of discourse analysis through the revision of different aspects of referentiality. Referential analyses have much to offer to typological studies; however, some concepts required to make an exhaustive analysis are still obscure. Similarly, methodological tools for analysis depend on theoretical bases, which depend on the results of typological studies, which likewise depend on available data analyses. This seemingly endless dependency circle could perhaps be interrupted by the contributions of more language-specific studies that are open to different language phenomena. However, being open to structural differences is not enough: it is also essential to confront the data as it is and try to describe everything that we *observe*, instead of limiting the results to everything we can explain, because *that* is what will enrich typological studies.

APPENDIX

Procedural text: "Trabajos de la mujer" (PT)	
Tarahumara	**English**
Jípi ko na a mi che raíchi-ma la bá échi mukí ralámuli kítira: mukí rarámuli ko a bilá taá nóchaka rejói cho pa, écho na bitichí échi ona kawichí ayé mápu ona ichiwá má. Bilé mukí ko bilé rawé náli wika nátame nócha, mápureká beá asísika, biré mukí a chukú sakí newayá; a chukú reméa beá asísika; napíwili batúa, kobísi newayá, ayéna cho newá kowaami beá mápu kíte kóama kípi kúchuwa .	Today I will talk about the Raramuri woman: we, the Raramuri women, live working like men in the house and at the mountains where sowing takes place. A woman can work in many ways in one day, for instance, as she gets up early, a woman makes corn; she makes tortillas when she gets up early; she grinds corn, she makes *pinole*, and she also makes food early so that her children eat.
Má pé rawé ko bilé mukí ayéna si á nikúuru siné káachi, mápualí wasachí ayé nirúlisa nóchali, mapuriká namúti ichirúame, mápu-riká nócha cho échi muní iyéri, reyáwi boneyá nócha cho; échi muní iyéri napóa. Nócha cho siné kaachi échi mápualí kuuchi pachí jáwi, reyáwi cho boneyá. Napóa má. Ayéna cho nócha échi mukí nikúuru cho mápualí má sono rékiachi jú, mápualí má wíwaachi jú, mápualí má boniwá, échi muní má kalá rakésa échi ná wasachí ayé mápuoná ichiwá.	During the day a woman also goes to help out sometimes, when there is work to do in the field, for instance, when something was sown, like working on the beans, pulling up the weed; she works there cleaning the weed from the beans. Sometimes they work, when a small garden is set, by removing the weed. She weeds there. They also work and help there when it is time to cut the cane, or when it is harvesting season, when we weed, when the beans sprout there in the field and where we sow.
Jériká nikúrisa échi mukí chécho kú alíwaachi kó ayéna a chukú-ro napíwia; napíwili newayá reméa má. Kowáami newayá, muní waseyá. Jériká isimí mápualí nochása la: nikúruka iyéna, kuírosa échi kípu kunála sine cho rikáachi. Ayéna mukí a nócha mápualí namúti newaami nísa échi mukí: púra iyéri machisá newayá, walí iyéri newáami nísa, sekolí iyéri newáami nísa, kemáka newáami nísa.	Once she has helped like this, the woman again in the evening also makes corn; she makes corn and tortillas. She makes food and boils the beans. That's how she goes when she works: she goes around helping, she helps her husband. Sometimes a woman also works when she is a product manufacturer: if she can make belts, is she is a basket producer, a pot maker, or a blanket maker.

Ayéna bé échi mukí beá sébali namúti nochása, ma karewésa namúti: kowámi iyéri, remé iyéri muchuwása, kobísi achása kíti kalá ku níwima kowáami échi kuuchi. Ayéna a chukú échi mukí ko newayá échi púra, newayá échi kemáka, newayá échi sekolíki. Sébali námuti machi, machíame nisa bilé mukí ayéna bela á. Ká rewesa échi namúti bitichi nocháliame, ayéna bela chukú pa rawé ko newayá jépu na namúti.	A woman also works early on other things, she arranges everything: the food and tortillas are made, she cooks *pinole* so that her children have good food. A woman makes the belt too, she makes the blanket and the pot. She knows other things, a woman is also wise. Everything looks right concerning the house chores, as she does these things during the day.
Shine kaachi bile mukí mápuali omawaaliachi nirúlisa. bilé wilíame niisa échi rejói wala, ayéna bela walú nócha-li níwi échi bile mukí. Mapualí maná batáli... mápualí manámare échi mukí batáli, a bela kó bachá rikó atí mukí ko. Walú kobísi newayá... walú newása échi kobísi bachá rikó. Ayéna bela á paké échi shunú pa mápu batáli newáma ba.	Sometimes, when there is a celebration and the husband is a chosen leader (for the celebration), then the woman also has a lot of work. When she makes *batali*…when the woman is going to make *batali*, the woman is there first. She makes a lot of *pinole*…she makes a lot of *pinole* first of all. She also soaks the corn which will make *batali*.
Ayéna bela á amíba échi sunú mápu gará awi-sá repa, ayéna pakésa sunú má moá kíti gará awiméa échi tói. Má gará awésa échi tói eneríame, má bilá chukú échi mukí ko matárli batúa ba. Ayéna bela á sebárisa ba. Mápu uchulú kulusá re échi bile wasí, má wasé échi nail batusí pa—tói batúsa bachá ba. Ayéna iyéna mukí ko bawí túa, manaáya napíchi. Páki échi nail sanalí napíchi ayé, ayéna bile rawé aliché, échi mukí ko tibúa échi sanalí napíchi kíti gará wasiméa, kíti gará kétási ikóta ma.	Also, one looks for the corn that has grown tall, then she soaks the corn and puts it in the pot so that the corn pops up well. When the so-called *toi* has popped up, the woman sets to grind it with a *metate*. Then she is done there. When the beverage starts to thicken, she boils it with the ground corn—the corn she first ground. The woman goes to fetch water and puts it on the fire. She drains the *sanali* from it by the fire, and for one day the woman watches over the *sanali* on the fire so that it is well done, so that it doesn't burn.
Má wasésa échi mukí ko échi sanalí, má su´ubúla aliiwachi bá. Má bokowíachi chukúru échi mukí ko sanalí su´ubúa kíti rulabaama. Ayéna wikaná bila amána bá kíti gará rulabaama échi sanalí pa. Arí maa chukúru la ramea bá.	When the woman has boiled this *sanali*, she takes it out in the evening. In the evening the woman is taking out the *sanali* so that it cools down. And she puts it in many places so that it cools down very well. Meanwhile she makes tortillas.
Chécho a kóa chukúru cho kípu kúchuwa kípu panéro má, alíi bila isába échi mukí bá, risiká	Her child and her husband eat again, while the woman rests. She is tired as she does

bilá maná ba, mápualí walú maná ko bile batáli. Walú omáwali nirulisa sine kaachi, ayéna ba´alí miko chécho a chukú échi mukí ko. Chécho batúa échi tói, má karewésa, má achása échi batusí ki. Ayéna chukú échi mukí pákia la ba, mápu rapáko nalaka wasíli échi sanalí ki. Échi sanalí má ruláami ma roá be echo ná sekolíchi kíti gará ronoméa kíti gará rikútaame níma la echi batali. Bilé mukí gará machíame manayá, wabé bilá alá waseká roá sekolíchi ba échi sanalí. Ayena cho mápualí má kuwaní la échi muki ba, waséa échi sanalí.	everything, when she makes a lot of *batali*. Sometimes if there are big celebrations, then again the woman makes *batali*. Again she grinds the *toi*, then she arranges everything, then she cooks this ground corn. Also this woman drains (the beverage), the *sanali* she cooked yesterday. She puts the cold *sanali* into the pot so that it cools down very well so that it makes you drunk. A woman knows well how to make *batali*. She puts the *sanali* into the pot very well cooked. Then, when this woman finishes, she boils the *sanali*.
Má chukúpa échi basíawi anelíame batúa ba, tá bila kasalá júpa mápu wabé bilá juerte newá échi batáli ba. Siné kaachi échi achésiwa. Wabé bila chukú échi mukí matárli batúa échi nali basíawi anelíame. bachá rikaachi napabúa achása échi basíawi mápualí nirúla baláa pá. Alí katéwika achá échi mukíko basíawi kíti gará níwimala bá mápuali newámare bile batáli ba sine kaachi	One sets to grind the so called *basíawi* (wheat), which is a small herb that makes the *batali* stronger. Sometimes this herb is added. The woman grinds a lot of this so called *basiawi* on the *metate*. First we put together the *basiawi* when there is (any) in summer. The woman puts away the *basiawi* so that it is good at the time when she makes some *batali*.
Maa lá kuwanísaka échi batáli waséa, má simí échi mukí la ba o rejói má bajurésia la ba, échi ná mápu oná peré la ba, bajuréka bela bakía ba, échi siiame bá échi koliame la batáli ba ne. Jérika ju echo na batálichi nirá ko.	Once she is done cooking the *batali*, then the woman or a man goes to get the guests, there where they live, the invitees, the guests, the visitors drink *batali*. That's how it is at a *batali* party.
Ne wika nátame bela jú échi mukí nocháa la ko, a bilá ko machi cho échi kuuchi tibúa ba; machiká bilá ináro cho échi kuuchi kito weelia bá. Machí binói ko mápuali nayuusa échi taa, a machiká ináro échi reyáwi; píri reyáwi owáami bajirisa reke échi taa mápualí witabóasa, mápualí we ratálisa, mápualí rosowása, mápualí ropachí rikáachi okósa échi taa towí taa tewé kere.	It is in many ways that a woman works, she knows how to take care of her children; she knows how to educate the children. She knows when the little kid gets sick; she knows the herbs, which healing herb is the one she must give to the little kid when (s)he has diarrhea, when (s)he has fever, when (s)he is coughing, or when his/her stomach hurts sometimes.
Ayéna bilá machiká ináro chabée jonsa cho, kípu wénowala be ko benérili cho a échi mápu	Woman goes around wisely since long ago, her parents taught her how to use healing herbs. A

ikí reyáwi owáami náti ko. Ayéna bila a tibúka lainàro échi mukí ko kípu kúchuwa, ayéna raíchika cho échi, échi ónami la ba échi ónami la ba owirúame ba mápu gará tibúka inároma kípu kúchuwa. A bilá ko échi bela owirúame bela maníiwa échi batáli, sine kaachi ówiisa la échi taa bá, mápualí échi nail taa majása sine kaachi biléana bajíchi. Ayéna bilá ruyelíwa échi owirúame bá mápu ku machí, paama échi alewála. Ayéna kalá ku sawisá échi taa, échi mukí ko ayéna bila walú manii échi batáli bá, ayéna kowáami cho a kóa échi owirúame. Kipu rana gará kú sawíli lá ba. Jériká bila júpa ne. jériká bila kó nócha la échi mukí ko bá, échi mukí ralámuli ko ba, jénaí ayé ralámulichi ko né. Matétera bá.	woman also takes care of her children by talking to them, to the doctor, the doctor, the healer, so that her children go around safely. One makes the *batali* for the healer, sometimes when he heals the little kid, when this little kid gets scared sometime at a baji. Then the healer is told so that he knows (that the kid is scared), and releases the soul (of the child). And if the child is well cured, the woman also makes a lot of batali and also food that the healer eats, so that he is well pleased. Her son healed well, This is how. This is how the woman works, the Raramuri woman, here where the raramuri live. Thank you.

Narrative Text "Tubali"

Narrative text Tubali (NT)	
Tarahumara	**English**
Tubáli ábaga riwéwi échi rijói, riwála niká ka iyénti, chabé ka ma ku wiká-li jípe ka. We aʹlá riwáti ke éche rijówi, chisíme we a'lá riwáti. Birépi raláti pi simíli ka riká, nawólaruá kárga rolomáchi niwála ruá. We rigá wilíga ruá ru kárga rolomála ruá ru.	Giant was the name of that man, that was his name long time ago, today he is gone. His name was well known, it was a good name. He had only one foot, that's how he walked. His foot was completely destroyed, it was broken, cut off, they say. They say he was standing when it was completely cut off.
*Muséchi ko'ubá ruá, tumó ke napáwi*sa ka. *Jémi noba ruá basíkolo; We a'lá suwá sirimé nawá ruá raléti. Anélta ruá ru pi napawísa miná ikoúlta ruá ru. Ke ka nokóya ichúali, ko ba chió ko'áya. Chió ichúa chíria rijóle ke suwábsa ka. Ichuála chabé piréme chió Biré a iyéntala ra'ícha, pi ichúa tibúka umuwí pocháme. Rikóchi muchúa niséla ruá ru, kúchi ma sísaga chabé ka pi*makói *ralámuli muchúa ruá ru tibúka.*	They say people from Munerachi went to kill him, they gathered. There they saw pascolas; all governors got there from the bottom (of the cliff). They told them they should gather because there were dead people. They didn't move, they sitted eating. Those men were siting like that, and woud not finish. Before them were the people too. One of them spoke, they were looking after a pregnant woman. They were sitting in the barn, they say, and the child was born too soon; ten men were there watching.
Biré nomí ka koyá énachi, tibúria bitichí ma ke itibú ko'á ruá ru. Pi usánipi muchása ka biréna bitichí, aria ómpa ko'á ruá ru. Ábaga pe wampi inla mapijí churú tamujué. Inárkake ochúla ka chió achúkarke wiká ka la ruá ko ba. Chió napawípsa éna ruá ko ba. Napawípsa ka énala ruá, wiká ka ka pi usántiri kichátiri	They were eating somewhere, watching over the house, and they would not stay to eat. Only six were in the house, so they could eat, they say. That way they are strong, like we are. Their bones were different and also their blood, they say.They were together, they say. They were together. They were many many, six or seven.
Sinémi kúchi kalá timútuka chumisá kúchi. je ne laláya inlábo kúchi ruá. Pikúchi jaré ba ochúla ka ochúla rítea, pi kúchi ka itíwela bo a rosánarka na iwílabo onákie. Wiká iníla pi ichúa ta kalí itíwela bo. Píri ga itíweli pachá-kie karíla.	First they took a child, they suck the child. I believe he was small, they say. There are some small bones; I have seen bones, small bones like our own, yes, white and strong, over there. Many bones were in a small house like ours. Like us, they were inside the house.
Pi échi yobi ichua karga, mo'ola we rosána ji iwíla ruá ru. Je mi onáye Ayórarle ka, chabéka tamí ta ka, chiká a'la suwérika itiwéli. Wíká	This same man was like this, his head very white and strong, they say. Yes, over there in Ayorare, long ago he was small like me; he

nirú chabé ka chomí, bitóli ma ke biri telépa jipé ka. Sunú jiwé ichúa matókela ruá. échi ka we pe we rojoránti sunú wilípa. Pachákie muchúa tiwéli láname sunú welí ramékame. We rojóranti níli ko ba, layá ijínti nijé ka. Pi jikúlitiri níli ko ba we konáteri a bitia itiweli onákie, rosochí pi ke chi jíki ko ba. we rosochí ka pachákie rosochí; konáka a ma wéka wiká koyánti. Ijínti ko ba laya ijínili nijé ka we uchúa opória, aní ruá ochérame ka we jiwétire ko ba.	was beautiful like us. There were many things over there, now there is not one bowl left. The big corn was taken on the shoulder. Very, very hard is the corn standing. Inside was the big, yellow corn with kernel. It was very hard, I think. Some time ago there was a lot of corncob like ours over there, nothing could go into the cave. Deep inside the cave was his house; he was eating lots of corncob. I think maybe there is a grave, that's what old people say.
Bonáka mawéru ba, ania ani ochérame ka; owísua ka maweya ania aníla ochérame, ke tibi ka. Pi mawétisa suwábalta ruá ru. Na jaré inálta ruá ru tibulua ka chio. Bitichí pi ma mulúbe mi ku, sisia suwíla ruá ru. We korúkame inla lepa chírika kumáya pichí yiri kumóya ji muchíla. Napeti olá ko'á itú; jemi biré nomi basóli kosá.	They ripped out and cut (the corn), old people say; the men asked for help to cut it, old people say, nothing was left. It is said that everything was cut. It is said that some (of the giants) were burnt, the last one too. It is said that very close, back in the house, he arrived and they died. He was eating, probably very hungry, and was eating this kind of food. Together they quickly made food; they ate pozole over there, somewhere.
Chiyó nokwála tamujué, ko'á warú birénomi basóli kosá; échi abe ka raísa anía anéla ba. Ke obé ichúsa ka chabé piréame komé jémi patú milílta ría. Niwái jeme patú milílta ria niwái	We make it too, we eat lots of pozole somewhere; they say it was delicious. Men are not eaten anymore. People from before are not watching, they were killed up there. They say they were killed up there, that's what people say.
Kanó anilíame ke chi olówame kária anía, narké awéme téri ka ría níli. Éche na ju ba rikábalta ria nia. nakárke ju ba rikábalta ruá. Échi kanó anilíame awéme ka ruá ru, awéme karí riá napé yíri wakási pi ikósa ka rikábalta ria.	The so-called giant was not tough. He also had a small horn, it was like that. He was there and was defeated like that. He was last one and was defeated. It is said the so-called giant has a horn, a horn like that of a cow and he was burnt, defeated.
Na karké jo'lú rosorása ikósa ga rikábalta riá. Na karké ke machíniga Asála ría. Niná karké we majála. Jolúla níla chukúya ría. Eche nolá-ya li na pi ke machíni ka. Asála majáya nabí biré ripisa ka. Miná ke machíti ka asala majaya pi ma ka nila kuba na kúchi suwábsa ka.	The big nest was also burnt, it was destroyed. He did not go out, he stayed like that. He was scared, very scared. He was the chief and was like that. He commanded but didn't go out. He was scared when he was left alone. He didn't go out, he was afraid, and in the end he was small.
Wiká chío níla ka napawíka rijóla ruá ru. Chabé ka napawíka rijóla ruá ru. Chabé ka	It is said that many men were together. It is said that long ago men were together. Long

achúa tami ke riwémela Pi ría ka muchúa chiké awénili. Ta niráma wiká ka rijóla ruá ru Chabé ka pi makói ka we ralámuli rijóla ruá. Chabé ka tibúya noléla ruá ru.	ago I was not allowed to be all alone. Few people wandered around, but there were many men, it is said. Long ago there were ten very brave men. It is said that long ago they went around watching.
Rikóchi jiena jicho ka jiena ta nalácha lachi. Rikóchi nalánti aria irápa wisá ria ruá ru jaré nomí.	There at the barn, they hit small kids when they criy. Crying babies are taken away, to the barn.

REFERENCES

Ariel, M. 1990. Accessing Noun-Phrase Antecedents. London: Routledge,

Bennett,W. and Zingg, R.1935. *The Tarahumara: An Indian Tribe of Northern Mexico.* The University of Chicago. Chicago, IL.

Bhat, D. N. S. 2004. *Pronouns*. Oxford: Oxford University Press.

Bickel, B. 2003. "Referential density in Discourse and Syntactic Typology." *Language*. 79, 4: 708-736.

Burgess, Don. 1984. Western Tarahumara. (Studies in Uto-Aztecan Grammar, 4.) Dallas: The Summer Institute of Linguistics and The University of Texas at Arlinton.

Caballero, G. 2008. Choguita Rarámuri (Tarahumara) phonology and morphology. Ph.D. Dissertation, University of California, Berkeley.

Caballero, Gabriela. 2002. Mecanismos de transitividad en rarámuri. Tesis de

Caballero, Gabriela. 2002. Mecanismos de transitividad en rarámuri. Tesis de licenciatura. Universidad de Sonora.

Cohen, D. 1998. A Grammatical Description of Tarahumara. MA Thesis. Arlington: The University of Texas.

Comrie, B. 2013. Alignment of case marking of Full Noun Phrases. Dryer, M. & Haspelmath, M. (eds.). *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology: Leipzig. (Available online at http://wals.info/chapter/98, Accessed on 2014-01-05.)

Comrie, B.1989. *Language Universals and Linguistic Typology. Syntax and Morphology*. 2nd ed. University of Chicago Press: Chicago.

Croft, W. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.

Croft, W. 1996. What's a head? In *Phrase Structure and the Lexicon*, eds. Johan Rooryck and Laurie Zaring, 35-37. Dordrecht: Kluwer Academic Publishers.

Dooley, R. & Levinsohn, S. 2001. Basic notions of reference. *Analyzing discourse: a manual of basic concepts*. Dallas: SIL International.

Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info, Accessed on 2015-03-01.)

Estrada et al. 2004. Diccionario yaqui-español y textos. Obra de preservación lingüística. México: Plaza y Valdés.

Givón, T. 1983. *Topic continuity in discourse: a quantitative cross-linguistic study*. Amsterdam: John Benjamins Publishing Company.

Givón, T. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins Publishing Company.

Givón, T. 2005. *Context as other minds. The pragmatics of sociality, cognition and communication*. Amsterdam: John Benjamins Publishing Company.

Grice,P. 1975. "Logic and conversation." In Cole, P., and J.L. Morgan, eds. *Speech Acts*. New York: Academic Press, 41–58.

Instituto Nacional de Geografía, Informática y Estadística. 2009. *Perfil sociodemográfico de la población que habla lengua indígena*. México: INEGI.

Instituto Nacional de Geografía, Informática y Estadística. 2010. *Conteo poblacional*. México: INEGI.

Kibrik, A. 2001. "Reference maintenance in discourse". *Syntactic Typology*. In: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*. Berlin: de Gruyter, 1123-1141.

Kibrik, A. 2011. *Reference in discourse*. Oxford: Oxford University Press.

Lehmann, C. 2008. Information structure and grammaticalization. Lopez-Couso, M., Seoane, E. (eds.), *Theoretical and empirical issues in grammaticalization 3*. Amsterdam: John Benjamins Publishing Company. 207-229.

Lehmann, C. 2012. "Approaches to the typology of referentiality" (Presentation). XIX International Congress of Linguists. Geneva, 21-27 July 2013. In http://www2.uni-erfurt.de/sprachwissenschaft/referentiality/links.html, accessed on 2014-04-05.

Lehmann, C. 2014. *Comparative corpus research on referentiality: theory, methodology, description*. Manuscript in preparation.

Lewis, M., Simons, G. and Fennig, C. (eds.). 2015. *Ethnologue: Languages of the World, Eighteenth edition.* Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Lionnet, A. 1972. *Los elementos de la lengua tarahumara*. México: Universidad Nacional Autónoma de México.

Mansilla, J. 2011. Archaeological and contemporary human hair composition and morphology. *Chungará (Arica)* [online]. 2011, vol.43, n.2, pp. 293-302. ISSN 0717-7356. http://www.scielo.cl/scielo.php?pid=S0717-73562011000200010&script=sci_arttext

Moreno, E. 2013. Chabé nirúame natári: Análisis de los referentes en el discurso mítico ralámuli. Tesis de maestría. Universidad de Sonora.

Nariyama, S. 2001. Argument Structure as Another Reference-tracking System with Reference to Ellipsis. *Australian Journal of Linguistics*, 21:1, 99-129.

Pintado, A. 2004. *Tarahumaras*. México: Comisión Nacional para el Desarrollo de los Pueblos Indígenas.

Reboul, A. 2001. Foundations of reference and predication. In Haspelmath, M. et al. (eds.), *Language typology and language universals. An international handbook*. Vol 1. Berlin: de Gruyter. 509-522.

Rodríguez, A. 2010. *Compendio del arte de la lengua de los tarahumares y guazapares de Thomas de Guadalaxara, 1683*. Chihuahua: Universidad Autónoma de Ciudad Juarez.

Saeed, J. 1997. *Semantics*. Oxford: Blackwell Publishing.

Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.

Smith, E. 1985. Text type and discourse framework. Amsterdam: Mouton Publishers.

Stoll, S., Bickel, B. 2009. How deep are differences in referential density? In *Crosslinguistic Approaches to the Psychology of Language: Research in the Traditions of Dan Slobin*. 543-555. University of Texas at Arlington.

Valdez, Yolanda. 2013. Predication in rarámuri (Urique Tarahumara). Ph.D. Dissertation, University of Oregon.

Villalpando, J. 2010. Hacia una caracterización del tarahumara de Turuachi: de lo básico a lo complejo. Masters degree thesis. Universidad de Sonora.