



Annotation of information structure in an Ixcatec text

Evangelia Adamou

► To cite this version:

| Evangelia Adamou. Annotation of information structure in an Ixcatec text. 2019. halshs-02417771

HAL Id: halshs-02417771

<https://shs.hal.science/halshs-02417771>

Preprint submitted on 18 Dec 2019

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Annotation of information structure in an Ixcatec text

Evangelia Adamou

CNRS

“The Typology and Corpus Annotation of Information Structure and Grammatical Relations” (Axe 3/GD1) of the excellence cluster *Empirical Foundations of Linguistics* (Labex EFL) funded by the French National Research Agency (Investments for the Future, ANR-10-LABX-0083)

1. Information on Ixcatec

Name and ISO code: *ʃ^hwa²ni³*, better known in the literature under the name Ixcatec (IXC), ixcateco (in Spanish), based on Nahuatl *ichcatl* ‘cotton’ + *-teca/-tecatl* ‘inhabitant of a place (whose name ends in *-tlan* or *-lan*)’.

Speakers: Ten identified speakers, of whom only four are fluent. Most of them - with one exception - are in their late 80s. All are bilingual in Spanish. They have had little formal education in Spanish and no formal education in Ixcatec.

Region: Ixcatec is spoken in the municipality of Santa María Ixcatlán in the state of Oaxaca, in Mexico. Today, Santa María Ixcatlán has some 400 inhabitants but at the time of the arrival of the Spaniards in 1522 it was an important centre for the Mixteca zone with an estimated population of 10,000 to 30,000 people.

Classification: Ixcatec belongs to the Popolocan branch of the Otomanguean stock together with Ngiba/Ngigua (also known as Chocho), Popoloc, and Mazatec.

Dialectology: There are no known dialects.

Status: Ixcatec is a critically endangered language, with less than ten speakers. An orthography was developed in the 1950s by a native Ixcatec speaker, Doroteo Jiménez, in collaboration with linguists of the Instituto Lingüístico de Verano, the Mexican branch of the Summer Institute of Linguistics. Doroteo Jiménez's orthography uses the Latin script and relies on the graphic correspondences with Spanish with some additions when necessary.

Main typological features: Ixcatec is a tone language, with three lexically contrastive tones: a high tone, transcribed with a superscripted¹, a mid tone, transcribed with², and a low tone, transcribed with³. Its phonology is complex and not yet well understood. The existence of stress is under discussion. Consonant inventory ranges from 24 to 52 depending on whether glottalized and aspirated consonants are analyzed as clusters of two segments, complex single segments or simple onsets followed by simple nuclei. It has five vowels which may be oral, /a e i o u/, or nasal /ã ë ï õ ù/. Ixcatec makes a clear distinction between verbs and nouns; some adjectives may also function as predicates. It is a head-marking

language, i.e., grammatical relations are marked on the verb. It has accusative alignment in indexing ($A = S \neq P$), i.e., only the single argument of intransitive verbs (S) and the agent-like argument of transitive verbs (A) are indexed on the verb through suffixes. A dozen experience predicates take a different coding, namely through possessive suffixes. Ixcatec is a pro-drop language, i.e., free pronouns are optionally used for all functions, and NPs are generally omitted. It has a VS/SVO unmarked order. When an S argument is moved to the preverbal position, a cross-reference morpheme is suffixed on the verb. The Ixcatec cross-reference morphemes ($-da^2$ ‘male’, $-k^wa^2$ ‘female’, and $-ba^3$ ‘animal’) corefer to nouns formed with the *noun classifiers*, d^2 - ‘man’, k^wa^2 - ‘woman’, βu^2 - ‘animal’, to some animate nouns even though they have no classifier, and to the masculine and feminine third singular pronouns which bear the same suffixes as those used for the cross-reference morphemes, i.e., $su^1wa^1-da^2$ ‘he’ and $su^1wa^1-kwa^2$ ‘she’. Noun classifiers are distinct from so-called class terms which partake in word formation for inanimates but are not associated with any cross-reference morphemes.

2. Annotation of information structure in an Ixcatec spoken text

For the annotation of Information Structure in Ixcatec (Otomanguan) I apply the “Annotation guidelines for questions under discussion and information structure” by Arndt Riester, Lisa Brunetti, and Kordula De Kuthy (2018). The Question under Discussion (QUD) approach aims to specify how the listener recovers the information structure from a text. QUD relies on three principles:

Q-A-Congruence:

QUDs must be answerable by the assertion(s) that they immediately dominate.

Maximize-Q-Anaphoricity:

Implicit QUDs should contain as much given material as possible.

Q-Givenness:

Implicit QUDs can only consist of given (or, at least, highly salient) material.

Questions under discussion are annotated in the Ixcatec ELAN file as follows:
{What are you doing?}

In the Ixcatec ELAN file, the following IS categories were annotated:

- **Focus:**

Definition: “The part of a clause that answers the current QUD” (Riester et al. 2018: 441).

In Ixcatec: Contrastive and corrective focus can be marked in Ixcatec with the focus marker $-na^2$ and prosodic marking (Adamou et al. 2018).

- **Restrictive focus:**

Definition: It restricts the scope of X (e.g. referent, location, temporal span) in relation to a set of alternatives.

In Ixcatec: Marked with an adverb ^h*ngwa*² ‘just’.

- **Topic:**

Definition: A discourse referent identifying what the sentence is about (Riester et al. 2018: 417).

- **Frame setting topic:**

Definition: A topic that sets the location and time.

- **Abouteness topic:**

Definition: “A referential entity (“term”) in the background which constitutes what the utterance is about” (Riester et al. 2018: 441).

- **Contrastive topic:**

Definition: “The instantiation of a variable within the background, which signals the existence of a superquestion-subquestion discourse structure. CTs are backgrounded w.r.t the subquestion and focal with respect to the superquestion” (Riester et al. 2018: 441).

In Ixcatec: CTs are constituents related to alternative questions and are optionally marked in Ixcatec by the focus marker *-na*² (Adamou et al. 2018).

- **Antitopic:**

Definition: Serves to reactivate a previously active discourse-topic (Lambrecht 2001).

In Ixcatec: Always in last position, separated by prosodic phrasing.

- **Afterthought:**

Definition: Right-detached constituents which express focus.

This is a tentative annotation that needs to be improved with further work on IS categories in Ixcatec.

Bibliography

Adamou, Evangelia. (2014). L’antipassif en ixcatèque. *Bulletin de la Société de Linguistique de Paris* 109(1): 373–396.

Adamou, Evangelia. (2017a). Subject preference in Ixcatec relative clauses. *Studies in Language* 41(4): 872–913.

- Adamou, Evangelia. (2017b). Spatial language and cognition among the Ixcatec-Spanish bilinguals (Mexico). In Bellamy, Kate, Michael Child, Antje Muntendam, and Maria Carmen Parafita Couto (eds.), *Multidisciplinary Approaches to Bilingualism in the Hispanic and Lusophone World*, 175–209. Amsterdam and Philadelphia: John Benjamins.
- Adamou, Evangelia and Denis Costaouec. (2013). El complementante *la* en ixcateco: marcador de clausula relativa, completiva y adverbial. *Amerindia* 37(1): 193–210.
- Adamou, Evangelia, Matthew Gordon, and Stefan Th. Gries. (2018). Prosodic and morphological focus marking in Ixcatec (Otomanguan). Adamou, Evangelia, Katharina Haude, and Martine Vanhove (eds), *Information Structure in Lesser-described Languages: Studies in Prosody and Syntax*, 51–83. Amsterdam and Philadelphia: Benjamins.
- Lambrecht, Knud. 2001. Dislocations. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (eds.), *Language Typology and Language Universals: An International Handbook, Vol. 2*, 1050–1078. Berlin: Walter de Gruyter.
- Riester, Arndt, Lisa Brunetti and Kordula De Kuthy. (2018). Annotation guidelines for Questions under Discussion and information structure. In Evangelia Adamou, Katharina Haude and Martine Vanhove (eds), *Information Structure in Lesser-Described Languages: Studies in Syntax and Prosody*, 403–443. Amsterdam and Philadelphia: Benjamins.