Scientific programme for the PICS:

"Mesoamerica and the syntax of the relative clause"

The goal of this PICS is to create an international research network involving linguists from France, Mexico and the US. One of the main objectives of the project is to revisit the notion of Mesoamerica as a linguistic area and it will be achieved by the in-depth study of the syntax of the relative clauses from a representative set of indigenous languages of Mexico and Central America to assess the role relative clauses have in the making of this diverse and important linguistic area of the world. The network will consist, on the one hand, of the collaboration of specialists in syntactic theory and linguistic typology, and on the other hand, of experts of endangered native languages of the Americas. Two intense syntax workshops to take place in Mexico will provide the research environment for the network to work and there will be an international conference in Paris for dissemination purposes. The new knowledge to be generated by this project will inform descriptive, areal, typological linguistics as well as theoretical linguistics.

1. The research context: Mesoamerica as a linguistic area. Linguistic typology is the field of linguistic science interested in human language through the study of its rich natural diversity (Evans and Levinson 2009). Paramount for the study of linguistic diversity is the understanding of linguistic change. But while two neighbouring languages often change in ways they become more different, they may also change in ways they become more alike. This makes of divergence and convergence two fundamental forces in the making and evolution of a given linguistic system through time. In this light, the interest in areal typology has come to the foreground in linguistic typology because linguistic areas are particularly informative about possible ways particular linguistic systems evolve and take shape in contact situations.

One of the most diverse linguistic areas in the world is Mesoamerica (henceforth MA), first introduced in Campbell et al. (1986). The area is sustained on a well-known continuum of material culture bearing the same name (Kirchhoff 1943), which extended from Central Mexico to Honduras from the dawn of agriculture in c. 7000 BC to the arrival of the Europeans in the 16th c. This cultural continuum provided a complex social network that nurtured a long-lasting, intense linguistic contact among the different linguistic communities in this geographical area. The contact among pre-Columbian societies brought about the convergence of many diverse linguistic systems around a number of structural traits which are still observable in the modern indigenous languages of Mexico and Central America. Such traits configure the Amerindian profile of MA. Later, the arrival of the Spaniards brought Spanish along to the region. In the new colonial order, through the contact with Spanish the different MA languages kept evolving into a convergence with Indo-European involving some of their features.

The MAn languages are extremely diverse. The core of the area comprises seven families of the Oto-Manguean phylum (Zapotecan, Popolocan, Amuzgan, Mixtecan, Tlapanecan, Chinantecan and Otomian), the three families of the Totozoquean phylum (Mixe, Zoquean and
Totonacan), the Mayan family, the small Tequistlatecan family, the isolate Huave and Nahuaatl, a late-comer Uto-Aztecan language. Beside the core systems, there are also a multitude of languages whose status with respect to the area still remains poorly understood as they do not show all the defining traits proposed for the area. Such non-core cases are considered as out-layers. In Central Mexico, out-layers include the Corachol family of Uto-Aztecan and the Chichimec-Pamean family of Oto-Manguean. Towards the South, one finds the languages of the Chibchan, Lencan, Xincan, Jicaquean and Misumalpan families of Central and South America. Recently, Munro (2013) has sustained that Garifuna, the Northern-most language of the Arawakan family, also shows MAn traits. In the discussion, the status as a MA language of Tarascan, a language isolate of Central Mexico, still remains problematic (cf. Smith-Stark 1994).

Since the proposals by Campbell et al. (1986), Smith-Stark (1988, 1994) and Yasugi (1995), research on MAn languages has mainly focused on increasing our knowledge of individual languages or language families. The challenges in this direction are real as the area has a large number of different languages, the majority of which seriously threatened. Besides the occasional contribution to the discussion (León and Levinson 1992, or Stolz and Stolz 2001), the tenets in Campbell et al. (1986) have remained unchallenged. However, our knowledge of the syntax of the languages of the area has increased considerably since the eighties, and a well-informed revision of the morpho-syntactic traits defining MA as a linguistic area is due in order. This project will serve as base to produce such a revision as it is inspired in the approach in Palancar and Zavala (2013a), where we take a first step in the right direction suggesting that syntax should be the next level of comparison for MAn languages. To attain this goal, we propose to carry out an in-depth study of the properties of the syntax of the relative clauses in a representative set of the languages of the area. We plan to achieve this by providing the research environment to nurture the collaboration between the experts of MAn languages working in France with a selected group of experts from Mexico and the US.

2. Object of study and justification: Why relative clauses in MAn languages? A relative clause (henceforth RC), such as for example the clauses [I want] or [where they live] in structures such as that’s not the car [I want] or the house [where they live] is nearby, is a subordinate clause that attributively modifies a nominal (called the head noun of the RC). RCs have played a prominent role both in formal approaches to syntax and in typological linguistics; the literature is, accordingly, vast. This role is justified because RCs are not only amongst the richest of constructions within the syntax of a given language, but they also display a rich array of types cross-linguistically. In this project, we focus on RCs because of two reasons:

- RCs are informative constructs about important aspects of the morpho-syntax of particular languages. By studying them in depth on a set of poorly-described and poorly-documented languages of Mexico and Central America we will generate new knowledge of the grammar of these endangered languages.
- As all human languages have them, RCs are convenient objects of study for the comparative method of areal typology. In a previous study, we have identified that aspects of the structure of RCs can be borrowed and were indeed borrowed in MA (Zavala 2013). The different domains of grammatical structure we want to study through them (see 2.2 below) will shed invaluable light onto a revision of the definitional criteria of MA as a linguistic area, which is the ultimate goal of this project.

2.1. The research background: Previous collaboration of the PIs.
Up to recently, the main focus in the description of the syntax of MAn languages had been on the fundamentals of main clause structure. Depending on family, such fundamentals often involve complex voice phenomena including obviation and inversion (Zavala 2007). From 2007 onwards, the PIs have been key actants in providing the research environment for the production of studies on the syntax of subordination in several MA languages, such as for example the syntax of secondary predication, complex predication and complementation. This was made possible through organizing several workshops to which collaborators attended which took place at CIESAS-Sureste in San Cristóbal de las Casas, Mexico, under the auspices of the research project SEP-CONACyT-47175. Research outcomes such as Aissen and Zavala (2010a) and Palancar and Zavala (2013b) can serve as proof of this success. The research achievement of these workshops lay in being able to produce a substantial amount of comparable knowledge of the grammar of different indigenous languages of Mexico and Central America by implementing a unified syntactic methodology. We plan to do the same for this PICS, but better.

The interest in RCs in MAn languages is recent and it comes in the right time for the discipline. In general, our understanding of such structures is limited as it is mainly based on a handful of unpublished materials (Hill and Hill, 1981, Peralta, 2004 and Vázquez, 2002, Gutiérrez 2013, being notable exceptions). Such studies for the most part are either first approaches or concentrate on only one aspect of RCs ignoring the rest. Alternatively, information about RCs comes from reference grammars such as Polian (2013) and Palancar (2009). More recently, Martínez (2007), de la Cruz (2010) and Jiménez (2014) are three up-to-date monographs on RC structure in three MAn languages, Chol, Tepostecan Nahuatl and San Miguel Chimalapan Zoque. The outcomes of this project can in turn be evaluated against the results of the different studies on the RCs of the non-MA Uto-Aztecan languages in Comrie and Estrada (2012). Such previous work can serve as control group to evaluate to what extent a given trait involving RC in the languages of our study is truly MA. In this PICS, we plan to focus on different aspects of RCs following a unified and well-informed methodology, both syntactically and typologically. For this we ask the research questions in §2.2.

2.2. Areas of study and research questions: We will research the following areas aiming to provide novel knowledge to answer a number of key research questions (henceforth RQ):
2.2.1. **RC and word order:** RCs have been given a prominent place in the discussion of basic word order typology since Greenberg (1963). Their position with respect to their head noun has traditionally been seen as a structural implicative (Bruce, 1978, but see Dryer 2005, 2013 for an update). The languages of MA have been traditionally identified as being verb initial, but many have shifted word orders, while others have preserved an old verb final order (e.g. Mixe-Zoquean), whereas the most have a word order split involving transitivity. This leads us to **RQ1. How the different types of RCs in the languages of study behave in these shifts?** This both with regard to their internal word order and to their placement with respect to their head noun. Likewise, as appositional structures are known to exist, but our knowledge of them remains limited, **RQ2. What is the degree of syntactic embedding of the RCs in different languages with respect to their noun phrases?**

2.2.2. **RC and the syntax of information structure:** Due to their particular make up, it is known that the syntax of RCs is intimately related to the syntax of focus by way of cleft constructions and wh-questions (Schachter 1973, Kiss 1998, etc.). This leads us to **RQ3. How is focus realized in the different languages of the sample in main clause and how these mechanisms relate to the internal structure of a RC?** and **RQ4. What is the structure of wh-questions in both direct and indirect discourse and what is their relation to the structure of a RC?**

2.2.3. **RC types, strategies and distribution:** The literature has long observed that in the languages of the world, RCs fall into a limited set of structural types (Keenan 1985, Lehmann 1986, Comrie 1998, Andrews 2007b). Lately, the different types are seen as revolving around four main strategies concerning the treatment of the head noun in the RC (i.e. gap, relative pronoun, non-reduction, and pronoun-retention). On the areal side of things, a specific strategy or even a specific type may be seen as characteristic of a linguistic area. In this direction for example, Haspelmath (1998) and Comrie (2006) propose that the relative pronoun construction is one of the defining features of Europe as a language area, rarely found elsewhere. This opens a crucial question for us: **RQ5. Can MA be characterized as for the existence of a specific relativizing strategy or a RC type?** A substantial portion of the research efforts of this project will be devoted to expand our knowledge to answer such a question. In this regard, Comrie and Kuteva (2005) have only three MA languages in their cross-linguistic map, and the three of them all have the gap strategy, but we know for a fact that there are other languages in Mexico and Central America of the other three strategies, for example Anderson (1986), Zylstra (2012), Jiménez (2014) and Knapp and Victoria (2013), leading us to **RQ6. How many types and strategies are found in MA languages and how spread they are both language internally and areally?**

On the other hand, since Keenan and Comrie (1977) we know that languages differ as to which syntactic functions are relativizable in the RC and by what syntactic means. This has provoked extensive discussion, and the literature is vast. It is also known that a language can have different RCs for different syntactic functions, making the areal study of RC in a specific area a
complex task as there are more variables in play. Because of this, a principle in our methodology shall be to study all possible relativizing functions (i.e. subject, primary object, direct object, indirect object) but paying special attention to possessors and adjuncts (instrumentals, comitatives, locatives) because Smith-Stark (1988) identified that pied-piping with inversion in wh-questions (whom with did you go? instead of with whom did you go?) was characteristic of the languages of his sample. This leads us to RQ7. What are the structural differences in relativizing a core function to adpositionally marked ones in MAn languages?

But what about if a language has alternative means for the same relativizing function? We know that this happens, but the literature does not go into further details leading us to RQ8. What is the distribution of similar RC strategies for the same function in natural discourse and what are the conditioning factors behind them? While most previous studies of RCs in MA have been based on elicitation, to answer this question both the PIs of this PICS and its collaborators have documentation materials at hand for the different MAn languages of study in the form of natural textual corpora to be in a position to establish the statistical significance of the distribution of the different RC types.

2.2.4. Free RCs: Crucial to the understanding of RC structure is the study of free relatives, a subtype of relative constructions functioning as full nominal phrases which occur in two types: head-less as in I like [what I see] or [who comes first] wins, or light-headed (Citko, 2004) as in Spanish [el que quiero] no es ese ‘The one I want is not that one’. Gutiérrez (2012) has claimed that free relatives in Yucatec Mayan are light-headed only, but our knowledge of such structures in MAn languages remains poorly although we know that they are common in natural discourse. In this PICS, we plan to tackle this gap to expand our knowledge of free relatives in natural usage.

2.2.5. RC and property concepts. In many MAn languages, the expression of most property concepts such as colour, size, etc. is by way of verbs and nouns and not by adjectives. This is so to such an extent that some languages have even been characterized as either having no adjectives at all, such as for example Cora (Vázquez 1994) and Northern Otomi (Palancar 2006), or as having just two of them as for example Tarascan (Capistran 2013). Consequently in such systems, in lieu of adjectives, nominal modification is carried out by participial clauses or RCs. We want to investigate what is the bearing the lack of adjectives has for the syntax of the nominal phrase across the sample and the extent to which there could be trends in this direction involving the specific function and labour RCs have across MA languages.

2.2.6. RC grammaticalization chains and borrowing? Since Campbell et al. (1986), it is known that MAn languages converge by way of calquing (Smith-Stark, 1994) and that this calquing often involves syntactic structures now known as ‘contact-induced grammaticalization’ Heine and Kuteva 2003, 2005, Chamoreau et Léglise 2012) as for example in Smith-Stark (1988), Zavala (2002), Gast and van der Auwera (2102), Peralta (2011), etc. In this respect, Zavala (2013) has shown that the Mayan language Chol has borrowed RC structure from Zoquean.
languages. This is a proof that RC structure can be and has been borrowed in MA, which leads us to RQ9. *What constitutes a RC calque in MA? Do these calques involve lexical material (i.e. relativizers), grammaticalization chains (i.e. indefinite noun > indefinite pronominal > relative pronoun), or are just syntactic (i.e. RC strategy or RC type, etc.)*? To answer this question, we first need to obtain a clear picture of what the origin is of the relative pronouns we observe (i.e. nominal classifiers, determiners, personal pronouns, etc.) and what the origin and types of the general complementizers.

3. Objectives: The previous research questions are driven by the following two objectives:

- **O1.** Produce new knowledge of a fundamental area of the grammars of a representative set of indigenous languages of Mexico and Central America by way of high-quality descriptions. This is important because most of the languages we will study are threatened, two of them on the brink of extinction, Ixcatec and Tilapa Otomi.

- **O2.** Produce new knowledge of the different aspects of RC structure that can be borrowed through language contact situations at a micro-level to shed light onto how a linguistic area is made up at a macro-level, and more particularly on what RC convergence reveals about the internal structure of MA as a linguistic area.

The outcomes of O1 and O2 will not only be of interest for the field of Amerindian linguistics but also they will feed into the general knowledge of linguistic science through its sub-disciplines in linguistic typology and contact and areal linguistics.

4. Research calendar. To achieve the objectives set up for this PICS in covering all research questions, we plan to create a research environment that is similar to the one that has been successful for us in previous occasions. Accordingly, in Y1 and Y2 we will organize two syntax workshops at the CIESAS-Sureste in San Cristóbal de las Casas, Chiapas, Mexico, each one of two-week's duration. The workshops involve intense working hours: a minimum of 6 hrs./day of group work plus an additional 4-6 hrs./day of individual work. They are thematically oriented and carefully planned. In Y1, the workshop will tackle syntactic methodology in the characterization of RCs in MA. In Y2, the scope will be typological. The workshops will be coordinated by the PIs in collaboration with two world-leading figures: Judith Aissen for syntactic theory and Bernard Comrie for linguistic typology. At the workshops, the different collaborators of the PICS will study in situ the different research questions on the languages of their expertise following a similar methodology and presenting their research results. In Y3, as part of a research dissemination strategy, we will organize an international conference at the SeDyL in Paris aimed at tackling the role of RCs in the characterization of areal typology as an important part of AXE 3 “Areal Typology” of the current research programme of the laboratory.
5. Research outcomes. The PIs will publish an edited volume on the RC constructions in MA (in Brill, De Gruyter, John Benjamins, etc.). This volume will consist of high-quality papers by the collaborators on representative different languages of Mexico and Central America. The book will also include a chapter on a revision of the typological trends found through the study of RCs which are relevant to expand our knowledge of MA as a linguistic area. Other different papers resulting from the third meeting will be published in the form of a special issue of an international journal on linguistic typology.

6. Funding under reciprocity conditions: See Annexe attached for details.

7. Network participants:

Principal investigators:
Enrique PALANCAR, SeDyL (UMR-8202), CNRS
Roberto ZAVAŁA, CIESAS-Sureste, CONACYT

French collaborators:
Claudine CHAMOREAU, SeDyL (UMR-8202), CNRS (Pech)
Denis COSTAQUEC, SeDyL (UMR-8202), CNRS (Ixcatec)
Evangelia ADAMOU, Lacito (UMR-7107), CNRS (Ixcatec)

Mexican collaborators:
Eladio MATEO, CIESAS-Sureste, CONACYT (Q’anjobal)
Gilles POLIAN, CIESAS-Sureste, CONACYT (Tseltal)
Néstor HERNÁNDEZ, CIESAS, CONACYT (Otomi)
Silviano JIMÉNEZ, CIESAS, CONACYT (Zoque)
Óscar LÓPEZ, CIESAS, CONACYT (Zapotoc)
Pafnuncio RAMOS, CIESAS, CONACYT (Zapotoc)
Godofredo MARTÍNEZ, CIESAS, CONACYT (Mixe)
Griselda GARCÍA, CIESAS, CONACYT (Mixtec)
José GÓMEZ, CIESAS, CONACYT (Tojolabal)
José del C. OSORIO, CIESAS, CONACYT (Chontal)
Juan Jesús VÁZQUEZ, Intercultural Chiapas (Chol)
Verónica VÁZQUEZ, UNAM (Cora)
Alejandra CAPISTRÁN, UAM (Tarascan)
Valentín PERALTA, ENAH (Nahuatl)
Rosemary BEAM, ENAH & UNAM (Zapotoc)
Michael KNAPP, INAH (Mazahua)

Collaborators from other countries:
Judith AISSEN, University of California at Santa Cruz
Bernard COMRIE, University of California at Santa Barbara
Eric CAMPBELL, U. of California at Santa Barbara (Chatino)
David BECK, University of Alberta (Tononac)
Telma CAAN, University of Texas at Austin (Kiche‘)
8. References cited:
    The Summer Institute of Linguistics and the University of Texas at Arlington.
Chamoreau, Claudine and Isabelle Léglise (eds.). 2012. *Dynamics of Contact-induced Language Change* [Language Contact and Bilingualism Series 2], Berlin: De Gruyter Mouton.
De la Cruz, Victoriano. 2010. Las oraciones relativas en el náhuatl de Teposteco, Chicontepec, Veracruz. MA Thesis in Indoamerican Linguistics. Ciudad de México: CIESAS.


Zylstra, Carol F. 2012. *Gramática del Tu’ún Savi (la lengua mixteca) de Alacatlatzala, Guerrero*. Ciudad de México: ILV México, AC.