Linguistic diversity, language vitality and the advancement of Linguistics as a science

Author:
Gabriela Pérez Báez
Curator of Linguistics
Smithsonian Institution
perezbaezg@si.edu

Department of Anthropology
National Museum of Natural History
MRC 112
Smithsonian Institution
10th and Constitution Avenue NW
Washington, DC 20560

Linguistic diversity, language vitality and the advancement of Linguistics as a science

Abstract
The study of linguistic and cognitive strategies for spatial referencing has seen an increase in studies reported in the literature since the 1970s. This research has moved from formulating theories based on a handful of international languages to documenting and analyzing a growing sample of the world's languages. This paper focuses precisely on the contributions that languages across a diversity of language families, socio-cultural contexts and geographic settings are making to the advancement of research on spatial referencing. A concrete example is illustrated by the Diidxazá language (Juchiteco, Isthmus Zapotec, Otomanguean) whose analysis has propelled significant explorations into spatial referencing in Mesoamerica and beyond. This paper explores the contributions made by Diidxazá and other lesser studied languages to the advancement of semantic typology, spatial referencing and language and cognition studies. This is placed in the context of current trends in the decline of the world's linguistic diversity. Thus, this paper advocates for a sustained, and, ideally, increased engagement of the sciences in documenting and revitalizing the world's languages.

1. Introduction

The study of linguistic and cognitive strategies for spatial referencing has been an area of increased scientific study since the 1970s with numerous publications on the topic published over the last 20 years. The symposium *Geographic grounding: Place, direction and landscape in the grammars of the world* and this special issue of *Acta Linguistica Hafniensia* edited around the symposium are proof of the research interest in these topics across a variety of disciplines including linguistics, psychology, geography and anthropology.[1] Further, both the symposium and this special issue illustrate the fact that linguistic diversity and its long-term sustainability are critical to the advancement of linguistics as a scientific endeavor. This is one of the two core topics that this paper addresses. It does so by presenting a summary of the contributions that the endangered language Diidxazá (Otomanguean, zai) has made to the advancement of studies in spatial language and cognition. This particular line of research is presented as an example in keeping with the scope of the symposium and this special issue. It should be made explicit, however, that linguistic diversity is critical to a broad range of topics and lines of research in linguistics.

Now, the relevance of studying a diversity of languages is not a new topic in the linguistics literature. Since the 1990s there has been an increase in the documentation and analysis of languages and in revitalization efforts. However, much remains to be done which warrants a continued discussion. This is the second point this paper addresses: that the engagement of the linguistics discipline with the endangered languages of the world is still very

---

[1] The symposium *Geographic grounding: Place, direction and landscape in the grammars of the world* was held at the University of Copenhagen, 30-31 May, 2016. See http://inss.ku.dk/english/calendar/geogram/, last accessed on January 25, 2018.

limited despite what the increased presence of the topic in the literature might suggest. The goal of bringing this to light is not to focus on the shortcomings of ongoing efforts. Rather, the purpose is to put the gains achieved to date into perspective with regards to the magnitude of the challenge at hand in order to motivate a greater and more informed engagement of the discipline with the ensemble of endangered languages of the world.

This paper provides in Section 2 an overview of some of the questions driving contemporary research into spatial reference, both at the linguistic and the non-linguistic level. A concrete example of the contribution that a lesser studied language can make to the scientific advancement of this research is illustrated in Section 3. This section shows that in-depth analysis of locative constructions and of the lexical semantics of the constituents of a locative utterance in Diidxazá (Juchiteco, Isthmus Zapotec, Otomanguean) has revealed basic strategies as well as dependencies across subsystems. Section 4 further explores the contributions of other lesser studied languages in addressing this and other related research questions and places them in the context of the decline of the world's linguistic diversity. Overall, this paper advocates for a sustained, and ideally, increased engagement of the linguistics discipline in documenting and also supporting the world's linguistic diversity.

2. Research into spatial referencing

Spatial referencing has been a long-standing area of study in psychology, linguistics, anthropology, geography, philosophy and the medical sciences. Questions about how humans process the spatial relations between objects in their surroundings have been addressed for centuries. In the 18th century, Kant (1991 [1768]) analyzed the relevance of the human anatomy, and specifically its medial and transverse axes, as a device for humans to partition space at various scales, including geographic scale. The symmetries and asymmetries of the human body continued to be considered well into the 20th century as basic devices driving the human cognitive ability to partition the space around us. Clark (1973), to take but one example, explains that the human perceptual space or *P* space is defined by the human body's symmetries and asymmetries at the anatomical and the perceptual level with spatial referencing being seated in innate human perceptual and anatomical features. At the time, however, studies relied on data from a relatively limited set of languages. Clark (1973) illustrates this point well as English and briefly German are the only languages mentioned in the article.

Nevertheless, researchers in the late 19th century and early 20th century were well aware of differences across the cultures of the world whose documentation became an increasing undertaking. A widely known proposal arguing for relativism in language, cognition and perception is the Sapir-Whorf hypothesis (Whorf 1940; Sapir 1958). Independently of the fact that the strong version of a linguistic relativity proposal has now largely fallen out of favor, this hypothesis served as a call to attend to the world's diversity in languages and cultures and to explore the extent to which this diversity has a bearing on human cognition. This exploration has been at the core of research efforts such as those led by the Language and Cognition Group of the Max Planck Institute for Psycholinguistics under the direction of Stephen C. Levinson. One early and seminal example of this work is Levinson and Brown (1994) which challenges Kant's analysis. The authors do so by referring to data in Brown and Levinson (1993) on practices of spatial referencing among Tzeltal Mayans from Tenejapa, Chiapas, Mexico.

Tenejapan Tzeltal (Mayan) is described in Brown and Levinson (1993) as having a system of reference based on an uphill vs. downhill referencing system. The terms *ta ajk'ol* 'uphill' and *ta alan* 'downhill' refer to inclines in the topography in Tenejapa. They may also be

used to describe the location of objects in small scale space and independently of whether the described direction is indeed in line with an upward or downward incline and independently of the speaker's orientation. The authors further clarify that while this system is in place, it is not in addition to a generalized use of a speaker-centered left-right asymmetry based on the human anatomy. In fact, the authors argue that complex body part terms that are used in reference to the left hand/arm in opposition to the right hand/arm do not comprise morphemes that can be readily extended to refer to left vs. right regions whether of the human body or elsewhere. Later studies examine the broader concept of Frames of Reference (FoRs) in which an object, a figure, can be located within a region projected from a reference object, a ground.[2] Numerous works have addressed the topic of FoRs. One notable study is presented in Pederson et al. (1998) which reports on a semantic typology analysis of FoRs based on data from 13 languages from ten different families including the Mayan languages Mopan, Tzeltal and Yucatec, Totonac (Totonacan), the Austronesian languages Kilivila and Longgu, Kgalagadi (Bantu), Hai‖om (Khoisan), Tamil (Dravidian), Belhare (Tibeto-Burman), Dutch (Indo-European) and Japanese (Japonic) as well Arrernte (Arandic, Pama-Nyungan). There is also a brief mention of FoR preferences in Warlpiri (Desert Nyungic, Pama-Nyungan).

Studies in topological relations comprise yet one more line of research delving into descriptions of spatial referencing across a number of languages in order to unveil the diversity of strategies of spatial reasoning and description across humans. Topological relations are considered to be some of the most basic spatial notions and include relations between objects on the basis of containment, support, contiguity and proximity (Piaget and Inhelder 1956). For example, in a containment relation, the figure may be described as contained inside a second reference object, a ground, as in *the apple is inside the bowl.* In a relation of support, a ground provides support to a figure against the force of gravity as in *the book is on the shelf.* Concepts such as IN, ON and UNDER, all referring to topological relations, have been posited as primitive or near-primitive concepts (Jackendoff 1983; Miller and Johnson-Laird 1976) and, in principle, universal (Piaget and Inhelder 1956; Landau and Jackendoff 1993).

In the case of topological relations, again, an expanded investigation of these concepts in a broad sample of languages from a diversity of language families provided important perspectives. For instance, Levinson and Meira (2003) take a semantic typology approach to the study of topological relations across nine languages. These include the isolates Basque, Lavukaleve, Trumai and Yélî Dnye, plus Dutch (Indo-European), Ewe (Niger-Congo), Lao (Tai-Kadai), Tiriyó (Cariban), and Yucatec (Mayan). The study is based on a uniform set of stimuli known as the Topological Relations Picture Series consisting of 71 images of objects in various topological relations. The data collected were mapped onto ethic grids that elucidated on a number of semantic categories. The authors describe considerable diversity in the composition of the categories and discuss potential factors driving such diversity. These factors include cultural practices. The authors cite, for instance, that the culturally salient use of trays or coolabahs as opposed to containers among Australian Aboriginals may be associated with the fact that concepts such as IN and UNDER are encoded as one category in a single spatial nominal (Levinson and Meira 2003, 514). Overall, the empirical data shows that not all languages recognize the same semantic categories, except for the apparent recognition across all nine languages of a category of ATTACHMENT. Further, those categories recognized by each language did not necessarily match the extension of the same categories in other languages.

---

[2] The terms ground and figure are used as per Talmy (2000).

Subsequently, the literature has been infused with works whose underlying research question is whether the differences that can be attested in the encoding properties of languages are indicative of differences in mental representations across humans and/or indicative of influences on language from factors such as culture, environment, education and/or language contact (Li and Gleitman 2002; Majid et al. 2004; Bohnemeyer et al. 2015 *inter alia*). These lines of investigation have produced numerous works, too many to cite in this section. For now, and to transition into Section 3, it will suffice to state that it is very much in keeping with the Language and Cognition Group's process of cross-linguistic and cross-cultural research with the aid of semantic typology methods that the research into Diidxazá has been conducted. The research initially focused on the topic of body part-derived terms to describe spatial relations. This required detailed analysis of the intension and extension of body part-derived meronyms – terms that refer to parts of a whole. In turn, these explorations have motivated work into the study of FoRs. In the section that follows, I summarize the phenomena that have been studied based on Diidxazá data, to show the extent to which the study of this single language can contribute to the advancement of research into spatial language and cognition. Prompted by the work on Diidxazá, studies into a number of other endangered and understudied languages have been made possible.

3. Frame of reference use in Diidxazá

Diidxazá is a Zapotec language belonging to the Zapotecan branch of the large Otomanguean stock of Mesoamerican languages.[3] It is VSO and head marking. In Diidxazá, body part terms functioning as nouns head ground-denoting phrases (see Pérez Báez 2012) as in the basic locative construction in (1).

(1) *n-uǔ   ti=bǒlla   zhaˀna   ti=taburěte[4]*
   STA-enter INDF=ball buttocks INDF=chair
   'There is a ball under a chair.'

Body part terms are used based on a mapping from the anatomy of the human body or an abstraction of it, as a source domain, and the geometry of objects as a target domain. It is the interaction between this structure mapping process and the use of FoRs that is the object of the analyses summarized in this section.

3.1. Frames of reference (FoRs) typology

FoRs are coordinate systems used to locate objects in space. In a FoR, a region is projected from a reference object – the ground. Within this region, a second object – the figure – is located. A number of typologies have been proposed outlining the types of FoRs that are known to be used

---

[3] Diidxazá is also known as Juchiteco, Isthmus Zapotec and *zapoteco de la planicie costera*. In this paper, the designation Diidxazá and its spelling is used in deference to the language community's practice to refer to the language as such. See Anonymous 1956. For a discussion on wordhood and orthographic representation as they relate to the spelling of the language designation, see Pérez Báez, Cata and Bueno Holle 2015.

[4] In order to be maximally informative in the orthographic representation of Diidxazá in this paper, the conventions used are those of the Project for the Documentation of the Languages of Mesoamerica. Some graphemes differ from the recommendations made by the *Alfabeto Popular para la escritura del zapoteco del Istmo* (Anonymous 1956) and tone is systematically represented. Glossing follows the Leipzig Glossing Rules, supplemented with the following abbreviations; HAB, habitual; I, inanimate; NEG, negative particle; NTRG, interrogative particle; STA, stative.

by humans and their classification (Levinson 1994; Wassmann and Dasen 1998; O'Meara and Pérez Báez 2011). For the purposes of this paper, it will suffice to make reference to three types of FoRs. The first is the relative FoR as defined in Levinson (1996), Pederson et al. (1998) and O'Meara and Pérez Báez (2001): a projection onto space is made from the perspective of the observer/speaker, based on the left-right asymmetry of the human body. The second is the FoR termed intrinsic in Wassman and Dassen (1998) or object-centered in O'Meara and Pérez Báez (2001/2011?). The third is the absolute FoR as defined in Levinson (1996), Pederson et al. (1998) and O'Meara and Pérez Báez (2001) where a region is defined based on an abstraction from an invariable system as would be that of cardinal points. Examples of utterances describing spatial arrays for each of these three FoRs are in (2).

(2) Relative: *the ball is to the right of the chair* (from the speaker's perspective)
Intrinsic: *the ball is at the back of the chair*
Absolute: *the ball is East of the chair*

In a relative FoR, the truth condition of a description depends on the orientation of the speaker and can therefore be invalidated if the speaker rotates. An intrinsic FoR depends on the anatomy of a reference ground and its geometry and orientation. In an absolute FoR, the orientation of the speaker and the particulars of the ground do not have a bearing as the FoR depends on an immutable system of coordinates.

FoR preferences have been a recurring theme in the literature since the 1990s. Studies such as Levinson 1994, Haviland 1996, Wassmann and Dasen 1998, Li and Gleitman 2002, O'Meara and Pérez Báez 2011 *inter alia* have shown that speakers of different languages encode different FoRs when describing the relationships between objects in their environment and also employ different FoRs at the pre-linguistic cognitive level. Much debate prevails, however, as to the relationship and possible influences between language, cognition and culture. But the variation in preferences in FoR use across the peoples of the world is a well-documented and accepted fact. For instance, speakers of languages such as English, Dutch, Japanese and Spanish (a contact language for Diidxazá speakers) have a preference for resolving small scale tasks by using a relative FoR (Levinson 2003). In contrast, speakers of Diidxazá readily refer to cardinal directions when locating objects even in small scale space and have a strong preference for the use of absolute FoRs, and make very little use of the relative FoR.

3.2. Frames of reference (FoRs) preferences in Diidxazá
FoR preferences in Diidxazá were documented in a study based on an experimental referential communication task carried out with 6 dyads of Diidxazá speakers (Pérez Báez 2011).[5] The task consisted of presenting identical copies of 12 photographs of a ball and a chair in various configurations to two speakers seated next to each other with a screen between them. A speaker acting as a director was tasked with describing one photograph at a time so that the partner would find the same photograph in their own set. Four rounds of this exercise using four

---

[5] The experimental tasks were developed by and carried out in the context of the *Spatial Language and Cognition in Mesoamerica* project (MesoSpace for short; NSF Award #BCS-0723694, PI J. Bohnemeyer, https://www.acsu.buffalo.edu/~jb77/MesoSpaceManual2008.pdf). The Ball and Chair and New Animals in a Row tasks were designed by the MesoSpace project after the Men and Tree and Animals in a Row tasks developed by the Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics (Danziger 1992; Levinson and Schmitt 1993).

different sets of photographs were carried out with each dyad. The data consisted in transcriptions of audio and video recording of the interactions as well as a record made of the matches and mismatches between the selections of the director and the matcher. Two types of descriptions were the focus of the analysis: descriptions of the orientation of the chair and descriptions of the location of the ball with regards to the chair.

Pérez Báez (2011) reports that out of 295 descriptions of the chair and its orientation, 231 descriptions equivalent to 78% of the total were descriptions in which an absolute FoR was used: the orientation of the chair was described in reference to cardinal points as in (3). In 2% of the cases, the description referred to a topological relation, in another 2%, the orientation was described in reference to a landmark, and in the remaining cases, no description of the orientation of the chair was offered. In descriptions of the location of the ball with respect to the chair, across a data set of 420 descriptions, both the intrinsic (object-centered) and the absolute FoRs were the most frequent. The location of the ball was described in relation to the geometry of the chair in intrinsic (object-centered) FoRs in 111 descriptions (26%) as in (4). Both (3) and (4) apply to the same image in Figure 1. The location of the ball was described in reference to cardinal points as in (5) in 138 descriptions (33%) of the total. Relative FoRs were not used in descriptions of the orientation of the chair and were used in only 11 descriptions (3%) of the location of the ball. An example is in (6) where, given the array shown in Figure 3, the truth condition of the description provided by the speaker can only be validated by the orientation of the speaker as facing the array. This description is therefore based on a relative FoR.

(3) *n-u=dxii*   *lu=nǐ*   *lǎdu getéʔ*
     STA-CAUS=turn   face=3I   side   south
     'Iit is turned to face the south.'

(4) *n-exéʔ*   *ti=bǒla*   *kweʔ=nǐ*
     STA-place   INDF=ball   flank=3I
     'The ball is placed lying by its (the chair's) flank.'

(5) *n-exéʔ*   *ti=bǒla*   *ladu getéʔ de laa=nǐ*
     STA-place   INDF=ball   side   south   of   3=3I
     'The ball is placed on the south side of it (the chair).'

(6) *n-uǔ*   *ti=pelǒta ladu bigáʔ x-tǐʔ*   *asyěntu*
     STA-enter   INDF=ball   side   left   POSS-INDF   chair
     'There is a ball left of the chair (from the speaker's perspective).'

An additional task was designed to document FoR preferences at the cognitive level. This was a non-linguistic task in that it did not require the production of language-based descriptions. In this task, participants were presented with 3 toy animals of symmetrical features, all placed on a table in line, facing the same direction. The participant stood between this presentation table and another one placed parallel to the first. The participant was presented with the initial array and asked to turn around 180 degrees and reproduce the array on the second table. This task could be solved in a manner consistent with a relative FoR i.e. with the toys facing the speaker's inherent left or right. It could also be solved by using other FoRs as might be the absolute, ex. all animals facing the same cardinal point as in the presentation array.[6] The task was administered to 19 native Diidxazá speakers. Participants overwhelmingly resolved the task in a manner consistent with the absolute FoR. Only one participant in one trial produced a response consistent with the relative FoR, further confirming a bias against this FoR type.

The notion that speakers of a number of Mesoamerican languages exhibit a bias against relative FoRs has been discussed in several works (Brown and Levinson 1993, 2009; Levinson 1996, 2003; Pérez Báez 2011; Polian and Bohnemeyer 2011; Hernández-Green et al. 2011). In the case of Diidxazá, the extent to which the use of the relative FoR is limited is striking. Further, and more interestingly yet, is the fact that the restricted use of the relative FoR can be shown to serve a specialized function.

3.3. Complementarity between meronymy and frame of reference use

Pérez Báez (under review) explains the process of semantic extension of body-part-derived terms (BPTs) to refer to parts of objects and to areas projected from them. Before presenting this analysis, it is worth discussing the etymology of BPTs briefly. Over 125 BPTs have been documented in Diidxazá (Pérez Báez under review). Most BPTs are native words that have cognates in other Zapotec languages and in a closely related variety of colonial Zapotec documented in the late 1500s. Further, most BPTs are characterized by the fact that they are inherently possessed and do not require the possessive prefix *x-* in order to enter into an obligatory possessive relationship with the noun that they precede in a canonical possessive phrase. In what follows, any lexical items referred to as BPTs are those whose etymology and/or morphosyntactic properties warrant the designation. These BPTs can be used to name parts of objects as was done in the Ball and Chair task: instrinsic descriptions generally utilized a BPT to refer to a part of the chair in relation to which the ball is to be located as in (4) above.

The analysis takes a Structure Mapping Theory (Gentner 1983 *inter alia*) approach to describe the types of comparisons that enable the semantic extension of BPTs to varying degrees. BPTs extended on the basis of a literal comparison between the human body part as a source domain and a part of an object as a target domain depend on a close match between attributes of the two domains. Thus, the BPT *dyaga* 'ear' can be extended to refer to the handle of a cup on the basis of a close match between anatomical features of a human ear and the geometric features of a cup handle. This dependency loosens when a BPT is extended on the basis of an analogy, rather than a literal comparison. In analogy, the most relevant matches between a source and a target domain are about relations rather than attributes. This is precisely what we see in

---

[6] To be more explicit, the task opposes solutions based on a relative FoR or a geocentric FoR. The latter term encompasses several FoRs that rely on features of the environment to project the region within which the figure (the ball) is to be located. For the sake of simplicity, the discussion in this section only makes reference to the relative, intrinsic and absolute FoR. Further details on other types of FoRs can be consulted in O'Meara and Pérez Báez (2011).

Diidxazá. Take for instance the extension of the term *năˀ* 'upper extremity' which can be extended to name, for instance, the branches of a tree. It is the relation between upper extremities and the rest of the human body, and not the attributes of the body part, that is mapped onto a tree. The upper extremities of a human are extensions of the body that may extend upwards and sideways. In humans, an important attribute of upper extremities is that there are canonically two of them. However, such an attribute does not condition the extension of the BPT to name tree branches (cf. Levinson, 1994 which raises questions regarding this very point).

In addition to participating in the processes of semantic extension described in the previous paragraph, a select set of six BPTs –a closed class– also participate in a process of abstraction in which the BPT may be used to refer to a part or region of an object with few to no discernable parts. Such would be the case of a sphere-shaped object as would be a ball. The six BPTs are *ike* 'head', *zhaˀna* 'buttocks', *lu* 'face', *deche* 'back', *kweˀ* 'flank', and *ndaani* 'stomach, belly'. The mapping of these BPTs onto objects of simple geometry depends heavily on the projection of an egocentric FoR onto the object. This is illustrated in Figure 4. Note that the BPTs *lu* 'face' and *deche* 'back' can only be mapped from the point of view of a speaker. In other words, in Diidxazá, an otherwise highly dispreferred FoR is indispensable for the mapping of BPTs in certain contexts.

PLACE Figure 4. BPTs mapped onto a ball HERE

The research that showed the interaction of the meronymic system and the FoR system also showed that the relative FoR has additional functions that seem rather specialized. The Diidxazá data suggests that the relative FoR serves in ambiguity resolution. For instance, in the Ball and Chair task, cases where a BPT might refer to more than one part of the chair, as when referring to one of the two sides of a chair as in (4) above, the relative FoR may be used to identify the correct part. The relative FoR is not the only disambiguating strategy documented in the experimental tasks. However, it is only in this context that the relative FoR was used.

3.4. Frame of reference use and disambiguation in Diidxazá
The data from the Ball and Chair task showed that the relative FoR was strictly used in cases where the use of an intrinsic FoR would have resulted in an ambiguous description. In other words, the relative FoR was used when the part of the chair in relation to which the location of the ball would have been described is not a unique part as in the case of the sides of the chair or its legs. This is illustrated in Table 1, which includes in the first row all of the photos that prompted the use of the relative FoR. The second row lists the photograph numbers as used in the stimuli. For instance, the array in Figure 5 was used as photo 3 in set 1 and also as photo 4 in set 4. The third row lists the code corresponding to the participants who provided the descriptions analyzed. The fourth and last row lists the relational terms used in the descriptions based on relative FoRs.

PLACE Table 1. Ball and Chair photos and relational terms in relative descriptions HERE

The images show that in all cases, the ball is next to a side of the chair and the use of the relators in a relative frame of reference allow the director to specify to the matcher at which of the two sides the ball is located. (7) is an example of a case in which a description based on an intrinsic FoR may result in ambiguity.

(7) *n-uǔ   ti=pelǒta kwe'=nǐ*
    STA-enter INDF=ball flank=3I
    'A ball is at its flank.'

Fig. 10. B&C 2.9

The description in (7) could apply to Figure 10 (same as 11) as well as to Figure 12 below, both of which were arrays that were presented within a single set and needed to be differentiated in order to resolve the task. The dialogue shown in (8) illustrates the process through which speakers may resort to the relative FoR to resolve the ambiguity. In (8a) the director describes Figure 11 by locating the ball in reference to the legs of the chair. This description is ambiguous in that it does not create contrast between Figures 11 and 12. In order to resolve this, the matcher produces (8b), a description based on a relative FoR in reference to Figure 12 which the matcher is considering selecting. We are able to confirm that the description produced by the matcher – *delǎnte nuǔ bǒla lá*– is relative because its truth condition depends on the observer's orientation and not on the chair's. If the matcher were sitting facing the back of the chair, for instance, the description would no longer be true. Based on (8b), the director then offered (8c) which also uses the relative FoR to confirm that the photo in question is in fact Figure 11.

 (8) (a) Director:
        r-uuya=dxǐ=nǐgeté' lá   pěru gí' di   bǒla ká
        HAB-see=calm=3I   south NTRG but adhered ball DEM
        'It is looking to the South, ok? but a ball is next to it.'

        gaxa pe'   de ka'nyee=nǐ ká
        near   precisely of PL leg=3I   DEM
        'Right by its legs.'

    (b) Matcher:
        delǎnte n-uǔ bǒla lá
        front STA-enter ball NTRG
        'Is the ball in front?'

    (c) Director:
        *ko', **atrá**   de laa=nǐ n-exé*   *bǒla ká   para   lǎdu ri-ndani   gubidxa*
        NEG behind of 3=3I   STA-place ball DEM towards side HAB-be.born sun
        'No, the ball is lying behind it (the chair) towards the East.'

It is not the case, though, that all participants systematically used the relative FoR in describing the five photographs in Table 1, hence the codes for the participants who did. In fact, only six of the 12 participants resorted to the relative FoR in these cases of potential ambiguity. For instance, participant 9 used the relative FoR in describing photo 2.9 but no other participant did. Indeed, participants made use of other strategies in order to resolve ambiguity such as the use of meronyms, topological relations and the use of cardinal points in an absolute FoR. However, the relative FoR seems to be restricted to resolving ambiguity. To my knowledge, this kind of specialized function for an otherwise dispreferred communicative and cognitive strategy has not been described before. In fact, the extent of what has been learned about spatial language and cognition based on Diidxazá –the relationship between FoRs and the semantic extension of body-part terms is also a unique contribution that the Diidxazá language has made to the advancement of semantic typology and spatial referencing.

3.5. Further cross-linguistic research motivated by Diidxazá data

The research on Diidxazá illustrates the value of studying a diversity of languages and especially languages that are not yet well documented or understood. It shows how a single one of these languages can propel scientific studies forward and at large scale. I first studied the use of BPTs in spatial descriptions and reported on their status as nouns even when heading ground phrases (Pérez Báez 2012). Based on this study, Jürgen Bohnemeyer suggested broadening the analysis of BPTs to a sample of Mesoamerican languages. An initial hypothesis was put forth that the prevalence of BPTs in spatial descriptions in Mesoamerican languages might correlate with a preference for intrinsic FoRs both at the linguistic and the non-verbal cognitive level. This gave rise to the MesoSpace project and therefore to the documentation of spatial referencing strategies in 13 Mesoamerican languages, in addition to two non-Mesoamerican indigenous languages of the Americas which served as control languages–Seri and Sumu-Mayangna located north and south of the language area respectively. In addition, four varieties of Spanish were included for control purposes. In the end, the original hypothesis was not validated. However, the cross-linguistic data did reveal that Mesoamerican languages exhibit a bias against relative FoRs.

Individual outputs from this research include language-specific descriptions of FoR preferences in eight Mesoamerican languages and the two control languages (O'Meara and Pérez Báez 2011). The MesoSpace project expanded to include 10 languages outside Mesoamerica including Turkic, Mon-Khmer, Atlantic Congo, Austroasiatic, Japonic, Sino-Tibetan and Indo-European languages in addition to three Mesoamerican languages, to investigate the validity of the hypothesis put forth in Bohnemeyer et al. (2015): that the innate bias for geocentric FoRs among great apes in Haun et al. (2006) might also be present in humans. Independently of whether this hypothesis is supported in future research, the point that matters here is that the advancements in spatial referencing studies within the two iterations of the MesoSpace project were initially motivated by research on Diidxazá and eventually enabled by the ability to carry out crosslinguistic research on languages that are sufficiently diverse in their typology so that the data can be representative of the diversity of the world's languages to the extent possible.

Yet the ability to derive such knowledge about the human language faculty is jeopardized by the rapid decline in the vitality of the world's languages and the influence of dominant languages on the native structures of other languages. In Diidxazá, for instance, of a total of 615 descriptions of the orientation of the chair and of the location of the ball with regards to the chair, 138 utilized a loan word as a spatial relator. Looking at the ensemble of data from 6 Mesoamerican languages and 5 non-Mesoamerican languages within the broader MesoSpace

project, an irreducible effect of the use of Spanish as an L2 in the increased use of relative FoRs was found, suggesting a strong contact-induced change in the native systems of spatial description in Mesoamerican languages which otherwise disprefer the relative FoR (Bohnemeyer et al. 2015). Spanish is the language used in Mexico in federal and state-level government, media and education. As mentioned earlier, Spanish speakers have been described as preferring relative FoRs for spatial referencing in small scale space (Levinson 2003; Bohnemeyer et al. 2015).

A point of clarification: by bringing up these findings, it is not my intention to vilify language contact or to advocate for purist views of languages and monolingual speakers living in isolation. The point to make here is that much has yet to be learned about the human language faculty. Yet, as languages continue to undergo pressures from a few highly dominant and often related languages, our ability to study the diversity of communicative linguistic strategies and associated cognitive functions is hindered. This brings us to the second point of discussion in this paper – that the engagement of the linguistics discipline with the endangered languages of the world is still very limited.

4. Linguistic diversity, language endangerment and their impact on typological studies

Estimates of the number of languages spoken in the world vary. The Glottolog (Hammarström, Forkel and Haspelmath 2017) reports 7,943 entries corresponding to individual languages from 430 different language families.[7] The Ethnologue lists 6,909 languages.[8] As introduced in Section 2, the linguistics literature has been infused with data from a diversity of languages. However, the very source of the data needed for these studies –the world's languages– is under threat, with the number of languages spoken in the world and their vitality being in steep decline. This is further complicated by the fact that our ability to ascertain the vitality of the world's languages is limited.

4.1. The vitality of languages around the world

Similar to the difficulty in determining the number of languages spoken around the world, establishing their vitality has presented significant challenges. By language vitality I refer to the rate at which the intergenerational transmission of a language continues to take place. The vitality of languages around the world has been in decline over the last few centuries. For instance, Garza Cuarón and Lastra (1991) estimate that at least 113 indigenous languages of what is now considered to be Mexico –one of the most highly linguistically diverse places in the world– have gone silent since colonization. Further, the early stages of the decline in linguistic diversity can be traced back to the expansion of Nahuatl and other languages of dominant societies even prior to colonization. One more example: in June 1791, ten years before becoming the President of the United States of America, Thomas Jefferson visited a place he referred to as Brookhaven, Long Island. During his visit, he collected vocabulary from the Unquachog language.[9] Jefferson's notes attest to the decline in vitality of this language stating that "There remained but three persons of this tribe now who can speak this language. They are old women.

---

[7] The number of languages was consulted at http://glottolog.org/glottolog/language while the number of language families was consulted at http://glottolog.org/glottolog/family in the Glottolog 2.7 edition (Hammarström, Forkel, Haspelmath and Bank (eds.)). Both web pages were last accessed on December 12, 2016January 27, 2018.

[8] The number of languages was consulted at https://www.sil.org/resources/publications/entry/6133 in the 16th Edition of the Ethnologue (Lewis, Simons, and Fennig 2016) last accessed on January 27, 2018December 12, 2016.

[9] Unquachog is an Algonquian language of the Southern New England branch. It is referred to as Unkechaug by the Unkechaug Nation and is also referred to as Quiripi-Unquachog in Goddard 1998.

From two of these this voc. was taken. A young woman of the same tribe was also present who knew something of the language." [10]

 Krauss (1992) is an initial attempt at estimating the extent to which the world's linguistic diversity is endangered. The author suggests that 90% of the world's languages could cease to be spoken within the 21st century. Simons and Lewis (2013) based on the Expanded Graded Intergenerational Disruption Scale (EGIDS; Lewis and Simons 2010) suggest that 37% of the languages that were spoken in 1950 are either no longer spoken by children or have gone silent. This percentage does not account for the languages that went silent prior to the mid-20th century. Further, a total of 2,444 languages out of 7,097 are reported to fall between the 6b level Threatened and the 9 level Dormant in the EGIDS scale.[11] In other words, 34% of the languages documented in the Ethnologue are either endangered or no longer spoken.

The estimate that about a third of the world's languages are endangered is likely overly optimistic and this is because assessing the vitality of a language is rather difficult. The case of Diidxazá shows how the assessment produced by EGIDS can fail to accurately capture the vitality, or in this case, degree of endangerment of a language. The Ethnologue (Lewis, Simons and Fennig 2016) considers Diidxazá as a developing language defined on the EGIDS as "The language is vigorous and is being used in written form in parts of the community though literacy is not yet sustainable" (Lewis and Simons 2010). A practical orthography was approved in 1956 by a committee comprised of speakers of Diidxazá in collaboration with linguists Velma Pickett and Morris Swadesh (Anonymous 1956; Pérez Báez, Cata and Bueno Holle 2015). There has been significant literary production in Diidxazá for decades, including literary magazines, poetry, short stories and the like. Several vocabularies have been produced such as Jiménez Girón (1979). Yet, literacy is not widespread.

Now, the critical point to make is that the focus on literacy obscures the fact that the language is undergoing a rapid process of language shift. Diidxazá is spoken by perhaps as many as 100,000 people (Instituto Nacional de Lenguas Indígenas 2000). However, as stated in Pérez Báez, Cata and Bueno Holle (2015) the literary production in Diidxazá is in stark contrast with the decline in the vitality of the language. In the 16th century Diidxazá used to be spoken well beyond the Isthmus of Tehuantepec into what are now the states of Chiapas and Veracruz. Nowadays, however, in the city of Tehuantepec which constituted the political center of the *binnizá* (Isthmus Zapotec people), Diidxazá is hardly ever heard. Only elderly speakers of the language remain, making up less than 10% of the city's population. Marcial Cerqueda (2014), a speaker of Diidxazá and researcher based in the Isthmus of Tehuantepec, carried out an analysis following the model termed *Índice de Remplazo Etnolingüístico* ('Ethnolinguistic Replacement Index', Ordorica et al. 2009). The author reports that only in one of the 22 municipalities where Diidxazá is spoken are children speaking it as their first language. For reference, in La Ventosa, where most of the data presented in Section 3 was collected, the youngest speakers of the language are in their late 20s. Therefore, Marcial Cerqueda (2014) considers Diidxazá as a language that is undergoing a process of *extinción acelerada* ('accelerated extinction'). The

---

[10] Based on Manuscript 28 of the National Anthropological Archives, Smithsonian Institution which is a handwritten copy made by A. S. Gatschet, from the original manuscript material in John Napoleon Brinton Hewitt's office based on Jefferson's original notes. Digital surrogates are available online at https://goo.gl/JwPPZX last accessed on January 25, 2018.

[11] Lewis, Simons and Fennig (2016), last accessed online at https://www.ethnologue.com/statistics/status on January 27, 2018December 16, 2016.

seemingly robust numbers of speakers reported in census results and the literary production do not reflect the reality of the rapid decline in the vitality of Diidxazá.

Simons and Lewis (2013) in their estimates report that 63% of the world's languages remain in everyday use and categorized above the threshold of Threatened category within the Vigorous to International levels in the EGIDS. Diidxazá, being labeled as a Developing language would fall well above the endangered levels (Lewis, Simons and Fennig 2016).[12] However, as explained above on the basis of a detailed analysis of the situation on the ground in the Isthmus of Tehuantepec, the EGIDS-based assessment of the vitality of Diidxazá is considerably flawed. Lee and Van Way (2016) evaluate methods of vitality assessment and propose the Language Endangerment Index (LEI) designed for the Endangered Languages Project. The LEI is based on intergenerational transmission of a language, absolute number of speakers, speaker number trends, and domains of use as key factors in assessing vitality. Based on LEI, Diidxazá is considered a vulnerable language.[13] This is perhaps a better assessment although optimistic when compared to the assessment proposed in Marcial Cerqueda (2014).

The case of Diidxazá suggests that the estimate in Simons and Lewis (2013) that two thirds of the world's languages are not threatened is overly optimistic. Indeed, it is highly likely that well over a third of the world's languages are in some degree of endangerment. The Endangered Language Project provides data on more than 3,000 endangered languages, i.e. close to 50% of the world's languages are considered to be endangered based on the LEI. Whatever the most accurate percentage might be, Krauss' (1992, 10) comment in closing his thought- and action-provoking article is relevant and very much a point that this paper wishes to emphasize: that linguistics stands to lose large percentages of its very object of study due to language endangerment. Below I illustrate the impact of the decline in the world's linguistic diversity on linguistics as a scientific endeavor.

4.2. The vitality of languages that enable research on spatial language and cognition

To analyze the vitality of other languages that have made the research into spatial language and cognition possible, I focus on three specific studies although certainly, these are not the only cross-linguistic studies in the area of spatial language and cognition. The studies selected for illustration are MesoSpace, Pederson et al (1998) and Levinson and Meira (2003). Altogether, these studies analyze data from a total of 37 languages. In what follows I present the assessment reports for each of these languages based on the Ethnologue and/or the Endangered Languages Project. The former provides vitality assessments on languages of the world whether they are endangered or not. The latter only provides vitality assessments for languages that it considers to be endangered. Data from the UNESCO's *Atlas of the World's Languages in Danger* is also mentioned to supplement the data.

According to the Ethnologue (Lewis, Simons, and Fennig 2016), among the languages analyzed in the aforementioned studies, Arabic, English and Spanish are international languages, Dutch, Japanese, Mandarin Chinese, Lao and Vietnamese are national languages, and two more languages are provincial, Basque and Tamil. These last two have very different language community sizes: Tamil has over 75 million speakers whereas Basque has around 600,000 speakers. Basque is actually considered Vulnerable by the UNESCO (Moseley 2010) although it

---

[12] Lewis, Simons and Fennig (2016), last accessed online at https://www.ethnologue.com/statistics/status on January 27, 2018~~19 December 2016~~.

[13] *Catalogue of Endangered Languages*. 2015. The University of Hawaii at Manoa and Eastern Michigan University. last accessed at http://www.endangeredlanguages.com/lang/1061 on January 27, 2018~~October 17, 2017~~.

is not listed in the Endangered Languages Project.[14] Overall, only 11 languages have over a million speakers: those that are international and national languages, in addition to Tamil as well as Ewe –considered as a language of Wider Communication with 4.6 million speakers– and Bashkir –considered as an Educational language with 1.2 million speakers (Lewis, Simons, and Fennig 2016).

All other languages including Basque have less than a million speakers. Of these, two languages have relatively small language community sizes yet are reported to enjoy a certain degree of vitality. The Ethnologue reports that Kilivila has some 20,000 speakers and is Developing, while Hai‖om is listed as a Vigorous language with some 46,000 speakers (Lewis, Simons, and Fennig 2016).[15] I would venture to question the validity of the assessment proposed for Kilivila and Hai‖om based on the Diidxazá case. All other languages, 21 in total out of 37, are endangered to one extent or another. These are listed in Table 2 based on the speaker counts and vitality assessment as per the Endangered Languages Project.[16]

<mark>PLACE Table 2. Endangered languages in spatial referencing research HERE</mark>

Of the 21 languages in Table 2, two languages are At Risk, six are Vulnerable, and 11 are considered to be Threatened. Belhare is Severely Endangered with only few speakers left, and Trumai is Critically Endangered. Kujirerai is not listed in Table 1 because there is no available data on this language in either the Ethnologue, ELCat or UNESCO's Atlas. Only an indication was found online about Kujirerai having a few hundred speakers with the speaker base declining.[17] Another observation to be made is that five of the languages that provided data for the aforementioned studies are isolates, with no known related languages in the world. As such, their data is especially valuable. Four of these isolates are threatened: Seri, P'urhépecha, Lavukaleve, and Yélî Dnye. Trumai, as was just mentioned, is Critically Endangered. Overall, this means that two thirds of the languages that have fueled important works in semantic typology, spatial referencing and cognition are endangered.

To drive the point home, these trends are also evident in the works presented at the symposium entitled *Geographic grounding: Place, direction and landscape in the grammars of the world,* some of which are featured in this special issue. The symposium presentations included 22 singularly identified languages from around the world. Of these, seven are languages spoken either at a national or at an international level: Danish, English, Faroese, Finnish, Icelandic, Marshallese and Mexican Spanish. Murrinhpatha is considered by the Ethnologue as Developing. Two languages are considered as Vigorous by the Ethnologue, Dhivehi with over 300,000 speakers and two of the three languages of Chiapas Zoque – Copainalá Zoque with 10,000 speakers and Francisco León Zoque with 20,000 speakers. The third language, Rayón Zoque is considered by the Ethnologue as Shifting with 2,100 speakers (idem). The latter is then listed in Table 3 along with the remaining languages which are considered to be at some stage of endangerment. No vitality report is available in neither Ethnologue nor ElCat for Acazulco

---

[14] *Catalogue of Endangered Languages*. 2015. The University of Hawaii at Manoa and Eastern Michigan University. http://www.endangeredlanguages.com/lang/search/#/?endangerment=U,S,AR,V,T,E,CE,SE,AW,D&sample_types=N,A,V,D,I,G,L&locations=known,unknown&q=Basque&type=code. <mark>Last accessed on October 18, 2017</mark>.

[15] Neither of these two languages is listed by the Endangered Languages Project <mark>as of October 18, 2017</mark>.

[16] *Catalogue of Endangered Languages*. 2015. The University of Hawaii at Manoa and Eastern Michigan University. http://www.endangeredlanguages.com. <mark>Last Accessed</mark> January 27, 2018~~December 21, 2016~~.

[17] https://elar.soas.ac.uk/Collection/MPI191249, <mark>last accessed</mark> January 27, 2018~~December 21, 2016~~.

Otomí. The UNESCO Atlas provides a vitality assessment for Ocoyoacac Otomí, Ocoyoacac being the municipality to which San Jerónimo Acazulco belongs. In the Atlas, Ocoyoacac Otomí is considered to be severely endangered with 609 speakers. Hernández Green (2015) reports, however, a much lower number of 200 speakers of Acazulco Otomí, all over the age of 65. The languages are listed in alphabetical order with their vitality status. In essence, about half the languages upon which research presented at the symposium was based, are endangered.

PLACE Table 3. Endangered languages amongst Geographic grounding presentations HERE

4.3. Discussion

The literature on the topic of language endangerment is robust (see for instance the overview in Pérez Báez, Vogel and Okura, in press). There are several resources that provide information on endangered languages including the Ethnologue (Lewis, Simons and Fennig 2016), the Endangered Languages Project, and the Atlas of the World's Languages in Danger (Moseley 2010), all of which were consulted for this article. Yet, there is seldom, if at all, a clear exposition of the impact of language endangerment for the advancement of linguistics as a science. For instance, Glottolog (Hammarström, Forkel and Haspelmath 2017) lists the languages of the world along with works in linguistic analysis of these, but does not provide an indication of their vitality status. Such is the case of the World Atlas of Language Structures (Dryer and Haspelmath 2013).

The documentation of the languages of the world is also long-standing. As Woodbury (2011) suggests, language documentation could be considered as old as writing itself. Icelandic Sagas represent documentation of the Icelandic language dating back 1200 years. Initial attempts at systematizing the documentation of European languages date back centuries, as in the case of Antonio de Nebrija's works on the Castillian language. Large numbers of vocabularies and grammars of languages encountered by European missionaries were produced during the 16th and 17th century. The example I provided of the Unquachog language, of data collection by the avid language documenter Thomas Jefferson in the 18th century feels recent in comparison. Language documentation, if anything, has increased, and enjoys greater systematicity and more stringent collection, processing and ethics standards than ever before (Himmelmann 1998; Newman and Ratliff 2001 *inter alia*). This has been made possible in particular by large grants devoted specifically to the mission of amassing some form of a record of the languages of the world. Dating back to 2000, the *Dokumentation bedrohter Sprachen* (DOBES) program by the Volkswagen Foundation enabled 67 language documentation projects around the world.[18] In a little over ten years, the United Kingdom-based Endangered Languages Documentation Program supported by the Arcadia Fund funded over 400 language documentation projects around the world.[19] The Endangered Languages Documentation Program, of the United States National Science Foundation in collaboration with the National Endowment for the Humanities (NEH) funded 115 projects between 2010 and 2017.[20] These three projects alone have enabled the documentation of about 10% of the languages of the world in less than two decades. The point I wish to make is that despite how robustly the problem of endangered languages is being

---

[18] http://dobes.mpi.nl/dobesprogramme/, last accessed on October 20, 2017.
[19] http://www.eldp.net/en/about+us/, last accessed on October 20, 2017.
[20] https://www.nsf.gov/awardsearch/advancedSearchResult?WT.si_n=ClickedAbstractsRecentAwards&WT.si_x=1& WT.si_cs=1&WT.z_pims_id=12816&ProgEleCode=7719&BooleanElement=Any&BooleanRef=Any&ActiveAwar ds=true&#results, last accessed on October 20, 2017.

addressed in the literature and the gains in language documentation made in this century, much remains to be done.

And there is one more, rather crucial point to make. While language documentation is critical at several levels, there is noticeable reticence or inability to engage with processes of language revitalization. In other words, while we are capable and willing –as a discipline– to document languages, we have yet to commit fully as a discipline to develop effective practices to support the sustainability of the languages of the world. Since 2013, I have directed the Global Survey of Language Revitalization Efforts.[21] 261 surveys were collected from language revitalization practitioners from all continents with the exception of Antarctica who responded to a 30-question survey. In order to stay within the scope of this paper, I will focus only on the topic of funding. The ensemble of survey participants reported a total of 699 distinct activities aimed at revitalizing a language including language nests, school-based language programs at various academic levels, non-school-based language classes and programs such as language camps and master-apprentice programs, teacher training and development of pedagogical materials, cultural events, language documentation, and the use of media and technology for revitalization. The survey asked about sources of funding for each of the activities reported including community, local and federal government, grants, and private donations. Out of the 699 activities reported, 168 (24%) revitalization efforts report receiving little to no funding. Respondents were also asked to articulate the greatest needs their efforts faced. A total of 353 tokens were collected in this section of the survey of which 168, or 48%, reported needing support, notably funding.

Certainly, linguists at the individual level and even some institutions are becoming increasingly engaged in language revitalization. Examples of this are reported in Pérez Báez, Rogers and Rosés Labrada (2016), Hinton, Huss and Roche (in press), *inter alia.* However, while funders have made strong commitments to language documentation, there is a clear avoidance to making similar commitments to language revitalization. The Documenting Endangered Languages Program in its 2017 solicitation states: "Documentation is a key complement to language revitalization efforts, but DEL does not support projects to revive or expand the actual use of endangered languages."[22] The Endangered Languages Documentation Programme in its Major Documentation Project Information Pack for Applicants 2018 states: "Although documentation and revitalisation are linked, projects aimed only at revitalisation without significant emphasis on documentation will not be funded. Nevertheless, applicants are strongly encouraged to create documentation in ways that assist communities to maintain and strengthen their languages. This may increase the possibilities for combining ELDP funds with revitalisation funds from other sources."[23]

Building revitalization efforts around language documentation projects can indeed be effective. For instance, a three-year language revitalization project for the community of La Ventosa where Diidxazá is spoken, was generated based on the Smithsonian Consortia-funded project entitled *Documentation and Revitalization of the Language and Traditional Ecological Knowledge of the Isthmus Zapotec Community*. However, while the documentation project could

---

[21] The Global Survey of Language Revitalization Efforts is a collaboration between the Smithsonian Institutions' Recovering Voices Initiative, the Auckland Institute of Technology and the Linguistics Department at The University of Hawaii at Manoa. Research and analysis has been carried out by a slew of research assistants, notably Rachel Vogel, John Uia Patolo and Sarah Johnson.

[22] https://www.nsf.gov/pubs/2016/nsf16576/nsf16576.pdf, last accessed on October 20, 2017.

[23] http://www.eldp.net/application/ last accessed on October 20, 2017.

be designed with an end-date, the language revitalization effort was only beginning to enjoy momentum when the funding for the documentation project ended. No further funding has been secured to date to continue the revitalization work despite various efforts. Overall, language revitalization has specific funding needs that have yet to be well understood and attended to by funding agencies. And yet, without direct action in support of the long-term sustainability of the world's languages, the linguistics discipline will not be able to secure the long-term sustainability of that which it studies. Beyond the needs that revitalization endeavors might have, we have yet to begin to understand how revitalization as a line of research might enlighten our understanding of how languages may realistically be sustained.

5. Conclusion

In this paper I have detailed the contributions that Diidxazá has made towards uncovering the fine details of the functioning of different spatial referencing strategies and how these interact. I have also explained the fact that the study of this one language has spurred large-scale, long-standing research into many other languages and a number of topics, thereby enabling the articulation of innovative proposals about the underpinnings of human spatial referencing strategies. I have then analyzed the vitality of the languages documented and analyzed in the three cross-linguistic projects in spatial referencing. This exercise shows the extent to which the languages upon which the advancement of linguistics as a discipline relies are endangered. I have done so in order to make it tangible to the reader what Krauss meant when he forecasted that the linguistics discipline is in the process of losing the very object of its study: the world's languages.

Much headway has been made in terms of documenting the world's languages and certainly, the role given to a diversity of languages in linguistics research is ever growing. The methods and ethics in research have all become scrutinized and led to significant improvements in the way that linguistics research is carried out as our work evolves into collaborative relations with many more language communities in many more contexts around the world. And yet, our engagement in support of endangered languages, as I have tried to show in this paper, remains insufficient. This paper has been written in the hope that a larger segment of the linguistics discipline will become engaged and with increasing reach, to contribute not only to the documentation and analysis of these languages but especially to the sustainability of the world's linguistic diversity.

References

Anonymous. 1956. *Alfabeto Popular para la escritura del zapoteco del Istmo.* México: La Sociedad Pro-Planeación Integral del Istmo, El Consejo de Lenguas Indígenas y El Instituto Lingüístico de Verano.

Bohnemeyer, Jürgen, Katherine T. Donelson, Randi E. Tucker, Elena Benedicto, Alyson Eggleston, Carolyn K. O'Meara, Alejandra Capistrán Garza, Néstro Hernández Green, María de Jesús Selene Hernández Gómez, Samuel Herrera Castro, Enrique Palancar, Gabriela Pérez Báez, Gilles Polian, Rodrigo Romero Méndez. 2015. "The contact diffusion of linguistic practices." *Language Dynamics and Change* 5: 169-201.

Brown, Penelope and Stephen. C. Levinson. 2009. "Language as mind tools: Learning how to think through speaking." In *Crosslinguistic approaches to the psychology of language: Research in the traditions of Dan Slobin*, edited by Jiansheng Guo, Elena V. Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura and Şeyda Özçalışkan, 451-464. New York:

Psychology Press.

Brown, Penelope and Stephen C. Levinson. 1993. "'Uphill' and 'downhill' in Tzeltal." *Journal of Linguistic Anthropology* 31: 46-74. DOI:10.1525/jlin.1993.3.1.46.

Clark, Herbert H. 1973. "Space, Time, Semantics and the Child." In *Cognitive Development and the Acquisition of Language*, edited by Timothy E. Moore, 28-64. New York: Academic Press.

Danziger, Eve. 1992. "Photo-photo matching games: Men and tree." In *'Manual' for the Space Stimuli Kit 1.0,* edited by Eve Danziger, 7-11. Nijmegen: Max Planck Institute for Psycholinguistics.

Dryer, Matthew S. and Maertin Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info, last accessed on October 20, 2017.

Garza Cuarón, Beatriz and Yolanda Lastra. 1991. "Endangered languages in Mexico". In *Endangered Languages,* edited by Robert H. Robins and Eugenius M. Uhlenbeck, 93-134. Oxford, New York: Berg.

Gentner, Dedre. 1983. "Structure-mapping: A theoretical framework for analogy." *Cognitive Science* 7: 155-170.

Hammarström, Harald, Sebastian Bank, Robert Forkel and Martin Haspelmath 2017. *Glottolog 3.1*. Jena: Max Planck Institute for the Science of Human History. Available online at http://glottolog.org, last accessed on January 27, 2018December 12.

Haun, Daniel B. M., Christian Rapold, Josep Call, Gabriele Janzen and Stephen C. Levinson. 2006. "Cognitive cladistics and cultural override in hominid spatial cognition." *Proceedings of the National Academy of Sciences in the U.S.A*. 103: 17568-17573.

Hernández Green, Néstor. 2015. *Morfoxintáxis verbal del otomí de Acazulco*. Unpublished doctoral dissertation. Centro de Investigaciones y Estudios Superiores en Antropología Social.

Hernández-Green, Néstor, Enrique L. Palancar and Selene Hernández. 2011. "The loanword *lado* in Otomi spatial descriptions." *Language Sciences* 33(6): 961-980.

Himmelmann, Nicholas. P. 1998. "Documentary and descriptive linguistics." *Linguistics* 36(1): 161-196. DOI: https://doi.org/10.1515/ling.1998.36.1.161.

Hinton, Leanne, Huss Leena and Gerald Roche (eds.). In press. *The Routledge Handbook of Language Revitalization*. New York: Routledge.

Instituto Nacional de Lenguas Indígenas. 2000. *Población de 5 años y más hablante de alguna lengua indígena por variante lingüística según bilingüismo lengua indígena-español. Localidades con asentamientos históricos, 2000*. Retrieved from http://site.inali.gob.mx/pdf/estadistica/VARIANTE/VARIANTE_C3_monolinguismo.pdf, last accessed on January 27, 2018.XXXXX

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, Mass.: MIT Press.

Jiménez Girón, Eustaquio. 1979. *Guía gráfico-fonémica para la escritura y lectura del zapoteco. Pa sicca rica Diidxazá xti Guidxiguie'. Cómo se escribe el zapoteco de Juchitán*. Juchitán: Editorial Vitoria Yan.

Kant, Immanuel. 1991 [1768]. Von dem ersten Grunde des Unterschiedes der Gegenden im Raume. Translation 'On the First ground of the Distinction of Regions in Space'. In *The Philosophy of Right and Left,* edited by James van Cleve and Robert E. Frederick, 27-33. Dordrecht: Kluwer.

Krauss, Michael. 1992. "The world's languages in crisis." In Endangered Languages, *Language*

68(1): 4-10.

Landau, Barbara and Ray Jackendoff. 1993. "'What' and 'where' in spatial language and spatial cognition." *Behavioral and Brain Sciences* 16: 217-238.

Lee, Nala H. and John Van Way. 2016. "Assessing levels of endangerment in the Catalogue of Endangered Languages ELCat using the Language Endangerment Index LEI." *Language in Society* 45(2): 271-292. DOI:10.1017/S0047404515000962

Levinson, Stephen C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Levinson, Stephen C. 1996. "Frames of reference and Molyneux's Question: Crosslinguistic evidence." In *Language and Space*, edited by Paul Bloom, Mary A. Peterson, Lynn Nadel and Merrill F. Garrett, 109-169. Cambridge, Mass.: MIT Press.

Levinson, Stephen C. 1994. "Vision, shape and linguistic description: Tzeltal body-part terminology and object description." *Linguistics* 32(4): 791-856.

Levinson, Stephen C. and Penelope Brown. 1994. "Immanuel Kant among the Tenejapans: Anthropology as empirical philosophy." *Ethos* 221: 3-41.

Levinson, Stephen C. and Sergio Meira. 2003. "'Natural concepts' in the spatial topological domain - adpositional meanings in crosslinguistic perspective: an exercise in semantic typology." *Language* 79(3): 485-516.

Levinson, Stephen C. and Bernadette Schmitt. 1993. "Animals in a row." In *Cognition and Space Kit Version 1.0.*, 65-69. Nijmegen: Max Planck Institute for Psycholinguistics.

Lewis, M. Paul and Gary F. Simons. 2010. "Assessing endangerment: Expanding Fishman's GIDS." *Revue Roumaine de Linguistique* 55(2): 103-120.

Lewis, M. Paul, Gary F. Simons and Charles D. Fennig eds. 2016. *Ethnologue: Languages of the World*, Nineteenth edition. Dallas: SIL International. Online version: http://www.ethnologue.com, last accessed on January 25, 2018.

Li, Peggy and Lila R. Gleitman. 2002. "Turning the tables: spatial language and spatial reasoning." *Cognition* 83(3): 265-294.

Majid, Asifa, Melissa Bowerman, Sotaro Kita, Daniel Haun and Stephen C. Levinson, 2004. "Can language restructure cognition? The case for space." *Trends in Cognitive Science* 83: 108-114.

Marcial Cerqueda, Vicente. 2014. *Plan de acción para el impulso de la lengua zapoteca en el Programa de Gobierno Municipal 2014-2016*. Ms.

Miller, George and Philip Johnson-Laird. 1976. *Language and perception*. Cambridge, Mass.: Harvard University Press.

Moseley, Christopher (ed.). 2010. Atlas of the World's Languages in Danger, 3rd ed. Paris, UNESCO Publishing. Online version: http://www.unesco.org/culture/en/endangeredlanguages/atlas, last accessed on January 27, 2018December 21, 2016.

Newman Paul and Martha Ratliff. 2001. *Linguistic fieldwork*. Cambridge, Mass.: Cambridge University Press.

O'Meara, Carolyn K. and Gabriela Pérez Báez. 2011. "Spatial frames of reference in Mesoamerican languages." In Spatial frames of reference in languages of Mesoamerica, edited by Carolyn K. O'Meara and Gabriela Pérez Báez. *Language Sciences* 33: 837-852.

Ordorica, Manuel, Constanza Rodríguez, Bernardo Velázquez and Ismael Maldonado. 2009. "El índice de reemplazo etnolingüístico entre la población indígena de México." *Desacatos* 29 (enero-abril 2009): 123-140.

Pederson, Eric, Eve Danziger, David Wilkins, Stephen C. Levinson, Sotaro Kita and Gunter Senft. 1998. Semantic typology and spatial conceptualization. *Language* 74(3): 557-589.

Pérez Báez, Gabriela. 2012. Semantics of body part terms in Juchiteco locative descriptions. In *Expressing Location in Zapotec,* edited by Brook Lillehaugen and Aaron H. Sonnenschein, 117-136. Munich: Lincom Europa.

Pérez Báez, Gabriela. 2011. "Spatial frames of reference preferences in Juchitán Zapotec." *Language Sciences* 33(6): 943-960. DOI: 10.1016/j.langsci.2011.06.013.

Pérez Báez, Gabriela. under review. Comparison types in the semantic extension of Diidxazá body part terms.

Pérez Báez, Gabriela, Chris Rogers and Jorge Emilio Rosés Labrada (eds.). 2016. *Latin American contexts for language documentation and revitalization*. Berlin: Mouton de Gruyter.

Pérez Báez, Gabriela, Víctor Cata and Juan José Bueno Holle. 2015. "Xneza Diidxazá: Retos en la escritura del zapoteco del Istmo vistos desde el texto Teria." *Tlalocan* XX: 135-172.

Piaget, Jean and Bärbel Inhelder. 1956 [1948]. *The child's conception of space*. London: Routledge and Kegan Paul.

Polian, Giles and Jürgen Bohnemeyer, 2011. Uniformity and variation in Tseltal reference frame use. *Language Sciences* 33(6): 868-891.

Sapir, Edward. 1958. *Culture, Language and Personality*. Berkeley, Cal.: University of California Press.

Simons, Gary F. and M. Paul Lewis. 2013. "The world's languages in crisis: A 20-year update." In *Responses to language endangerment: In honor of Mickey Noonan*, edited by Elena Mihas, Bernard Perley, Gabriel Rei-Doval and Kathleen Wheatley, 3-19. Amsterdam: John Benjamins (Studies in Language Companion Series 142)

Talmy, Leonard. 2000. *Toward a cognitive semantics* Volume 1. Cambridge, Mass.: The MIT Press.

Wassmann, Jürg and Pierre R. Dasen. 1998. "Balinese spatial orientation: some empirical evidence for moderate linguistic relativity." *Journal of the Royal Anthropological Institute* 4(4): 689-711.

Whorf, Benjamin L. 1940. "Science and Linguistics." *Technology Review* 35: 229-231, 247-248.