The contact diffusion of linguistic practices: Reference frames in Mesoamerica

1  Introduction: The Diffusion of Linguistic Practices

In this article, we attempt to make a contribution toward answering two questions: First, are practices of language use contact-diffused? And secondly, can such practices constitute areal features? To answer these questions, we explore the domain of spatial reference frames – cognitive axis systems involved in (linguistic and nonlinguistic) representations of space (e.g., Levinson 1996, 2003; Bohnemeyer & O’Meara 2012; Bohnemeyer 2012). Specifically, we investigate whether speakers of eight indigenous languages of Mexico and Nicaragua are influenced in their use of spatial frames of reference by the frequency with which they use Spanish as a second language. To this effect, we compare the use of reference frames among speakers of these indigenous languages and among speakers of three varieties of Spanish spoken in rural areas of Mexico and Nicaragua and in Barcelona. We also test whether there is an areal effect on reference frame use. Six of the indigenous languages belong to the Mesoamerican sprachbund, while the remaining two are spoken just north and southeast of the Mesoamerican area, respectively. The present section motivates the research questions. Our study is, to our knowledge, the first to empirically address the role of bilingualism in the domain of spatial frames of reference, which has played a key role in the debate about language-on-thought effects (cf. Section 6). Furthermore, we believe we are the first to examine practices of second language use – as opposed to categories of grammar or the lexicon – as a possible influence on first language use.

Languages borrow from one another anything from sounds and sound patterns via free and bound morphemes and morphological patterns to idioms, constructions, and semantic patterns (see, for example, Thomason (2001: 59-98) for an introduction and overview). The existence of any of these types of contact diffusion is not difficult to establish: prima facie evidence comes in the form of any single instance in which the relevant (type of) sound, morpheme, construction, etc., occurs at some stage in time $t_1$ in languages $A$ but not in $B$, then at a later stage $t_2$ occurs in $B$ as well, but is restricted to varieties of $B$ spoken by speakers who are to some extent bilingual in $A$, and finally at some $t_3$ has spread to speakers of $B$ with no knowledge of $A$.

But what about practices of language use, or linguistic practices for short? Practices may be loosely defined as any kind of behavioral patterns that are learned (rather than innate) and culturally transmitted or diffused (Bourdieu 1977, 1990). Here and throughout, we use transmission for the “vertical”, intergenerational transfer of cultural knowledge and behavior and diffusion for the “horizontal”, intra-generational transfer, adapting and generalizing this dichotomy from Labov (2007), where they are restricted to the transfer of linguistic knowledge. Linguistic practices, then, are patterns of language use that are culturally transmitted or diffused.

By this definition, any conventional aspect of language use is a linguistic practice – including, say, the practice of English speakers to call birds birds or to combine infinitives with the linker to. Such practices are incorporated into the lexicon, grammar, or phonology of a language. We may call these codified practices to distinguish them from non-codified practices, which do not form part of the linguistic code of a language, even though as linguistic practices they of course involve the use of linguistic code.
Codified practices can be defined as mappings of a sound (pattern), morpheme, or construction into a set of contexts, however the latter are defined. In contrast, non-codified practices transcend the use of any particular elements of linguistic code and in this sense are more loosely related to the code of a language. It is the non-codified practices that this article focuses on. In the remainder, the term *linguistic practices* is therefore used as shorthand for non-codified linguistic practices unless otherwise noted.

Let us briefly consider four kinds of non-codified linguistic practices: speech acts; speech events; metaphoric domain mappings; and cognitive frames, including spatial references frames. We do not intend this list to be exhaustive. The (non-codified) linguistic practices *par excellence*, or at least the best-studied kinds of linguistic practices, are speech acts (Austin 1962; Searle 1969) and speech events (Hymes 1972). Speech acts, for example, depend on what one might call a cultural ‘script’ for their execution. Thus, the reason why it is possible in some cultures, but not in others, to dissolve a marriage by saying ‘I hereby divorce you’ is quite simply that those cultures recognize this speech act as a conventional practice. The domain mappings involved in metaphors (Lakoff & Johnson 1980) likewise tend to be conventional (even as an individual metaphor licensed by a given mapping may be entirely creative) (e.g., Foley 1997: 179-191). This is the reason that metaphors in other languages and cultures may not immediately make sense to us – they involve domain mappings with which we are unfamiliar. An example is vertical metaphors of time in Mandarin (Boroditsky 2001 and references therein).

Frames in the sense of the FrameNet project (Fontenelle 2003) are a subtype of cognitive frames – knowledge structures we rely on to interpret our experiences – invoked by the semantics of linguistic utterances (Fillmore & Baker 2009). Spatial reference frames – the practice phenomenon our study focuses on – can be considered a special kind of frames in the FrameNet sense. Whereas the frames studied by the FrameNet researchers relate primarily to the semantics of event descriptions – especially the semantic roles of event ‘participants’ (semantic arguments) – spatial reference frames serve to interpret spatial representations, in particular, representations of the location, motion, and orientation of entities in space. We discuss the practice character of the use of reference frames further in Section 2.

Do linguistic practices diffuse through language contact? Since diffusion is a definitional property of cultural practices and linguistic practices are merely a special case of cultural practices, it follows that the diffusion of linguistic practices must be possible. However, this does not entail that non-codified practices are diffused *in tandem* with elements of linguistic code and codified practices. It is conceivable that the relation between non-codified practices and linguistic code is so loose that the two simply do not “travel together” in any systematic fashion. In this scenario, we might be reluctant to consider the diffusion of non-codified practices as a case of language contact even when it happens to occur among speakers of different languages.

Empirical evidence of contact diffusion of practices is surprisingly hard to come by. We are not aware of studies of the contact diffusion of speech acts or domain mappings, for example (which is not to say that such studies do not exist). The most likely explanation for this absence (or paucity, as the case may be) is that comparative

---

1 Enfield & Levinson (2010) provide a framework for the study of the *metalanguage* for speech acts across languages – i.e., the semantic classification of speech acts.
research into usage phenomena such as speech acts and domain mappings is still very much in its infancy. Ethnographers of communication have been studying speech events from a comparative perspective since the early 1970s (e.g., Bauman & Sherzer 1974); but most of the research on phenomena that are traditionally considered to fall under the rubric of ‘pragmatics’ has focused on individual languages only, and most commonly on the usual handful of WEIRD (Western, Educated, Industrialized, Rich, and Democratic) languages (Henrich et al. 2010).

The search for empirical evidence of the transfer of linguistic practices in language contact has important potential implications beyond the theory of language contact. On one possible interpretation of the Linguistic Relativity Hypothesis, language may serve as a conduit for the transmission and diffusion of culture-specific thought patterns – or cognitive practices. The diffusion of linguistic practices may be seen as a necessary prerequisite for the diffusion of such cognitive practices. We briefly discuss this connection in Section 6.

The rest of the paper is organized as follows: Sections 2 and 3 provide background information on spatial reference frame types and their use and on the Mesoamerican linguistic area, which we chose as the test case to probe the diffusion of reference frames. Sections 4 and 5 present the methods we used and the results we obtained. In Section 6, we discuss the Whorfian angle, and Section 7 concludes.

2 Referential Practice: Using Reference Frames

Reference frames are cognitive “coordinate systems” (metaphorically speaking, i.e., systems of axes) used to identify places (regions of space) and directions, often with respect to some reference entity or ‘ground’. Various classifications of reference frames have been proposed. Table 1 compares two of these. The one on the left is widely used especially in the psychological literature. It distinguishes between ‘egocentric’ (‘self’-centered, i.e., observer-centered) and two classes of ‘allocentric’ (‘other-centered) frames: ‘intrinsic’ (object-centered) and ‘geocentric’ (environment-centered) frames. The alternative classification on the right was proposed by Levinson (1996, 2003) on the basis of evidence from language typology. It distinguishes ‘relative’, ‘intrinsic’, and ‘absolute’ frame types. These two classifications are often misunderstood as terminological variants, but they in fact group frames quite differently. The psychological classification is based exclusively on what Danziger (2010) calls the ‘anchor’ of the frame: some entity or environmental feature defining the axes of the coordinate system. In egocentric representations, the anchor is the body of an observer. In intrinsic representations, the ground functions as anchor, and in geocentric ones, some environmental entity or feature does. In contrast, the typological classification also takes into account the origin point of the axis system. In intrinsic frames, the origin is the ground; in extrinsic ones, it is some other entity or environmental feature. Crossing the egocentric-allocentric dichotomy with the intrinsic-extrinsic dichotomy yields a four-way classification (Danziger 2010); but Levinson then collapses the egocentric intrinsic and allocentric intrinsic types to a single intrinsic super-type.

<table>
<thead>
<tr>
<th>Classification based on anchor (e.g. Carlson-Radvansky &amp; Irwin 1993; Wassmann &amp; Dasen 1998; Li &amp; Gleitman 2002; <em>inter alia</em>)</th>
<th>Classification by anchor and origin of axes (Levinson 1996, 2003; Pederson 1993; Danziger 2010; Bohnemeyer &amp; O’Meara 2012; <em>inter alia</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>egocentric extrinsic</strong> = relative (Levinson 1996)</td>
<td></td>
</tr>
<tr>
<td>anchor = body of an observer</td>
<td></td>
</tr>
<tr>
<td>ground ≠ anchor</td>
<td></td>
</tr>
<tr>
<td>axes projected (translated and/or reflected)</td>
<td></td>
</tr>
<tr>
<td><em>The ball is to the right of the chair.</em></td>
<td></td>
</tr>
<tr>
<td><strong>egocentric intrinsic</strong> (‘direct’ in Danziger 2010)</td>
<td></td>
</tr>
<tr>
<td>anchor = body of an observer</td>
<td></td>
</tr>
<tr>
<td>ground = anchor</td>
<td></td>
</tr>
<tr>
<td>axes extended (without projection or abstraction)</td>
<td></td>
</tr>
<tr>
<td><em>The ball is in front of me.</em></td>
<td></td>
</tr>
<tr>
<td><strong>allocentric intrinsic</strong></td>
<td></td>
</tr>
<tr>
<td>anchor ≠ body of an observer</td>
<td></td>
</tr>
<tr>
<td>ground = anchor</td>
<td></td>
</tr>
<tr>
<td>axes extended (without projection or abstraction)</td>
<td></td>
</tr>
<tr>
<td><em>The ball is in front of the chair.</em></td>
<td></td>
</tr>
<tr>
<td><strong>landmark-based</strong> (‘projected’ in Mishra et al. 2003; ‘head-anchored’ in Bohnemeyer &amp; O’Meara 2012)</td>
<td></td>
</tr>
<tr>
<td>anchor = environmental entity/feature</td>
<td></td>
</tr>
<tr>
<td>ground ≠ anchor</td>
<td></td>
</tr>
<tr>
<td>axes defined as vectors pointing toward/away from landmarks</td>
<td></td>
</tr>
<tr>
<td><em>The ball is mountainward of the chair.</em></td>
<td></td>
</tr>
<tr>
<td><strong>geomorphic</strong> (‘contextual’ in Jackendoff 1996: 17)</td>
<td></td>
</tr>
<tr>
<td>anchor = environmental entity/feature</td>
<td></td>
</tr>
<tr>
<td>ground ≠ anchor</td>
<td></td>
</tr>
<tr>
<td>axes projected</td>
<td></td>
</tr>
<tr>
<td><em>The ball is downriver of the chair.</em></td>
<td></td>
</tr>
<tr>
<td><strong>absolute</strong> (Levinson 1996; ‘geographical’ in Jackendoff 1996)</td>
<td></td>
</tr>
<tr>
<td>anchor = environmental entity/feature</td>
<td></td>
</tr>
<tr>
<td>ground ≠ anchor</td>
<td></td>
</tr>
<tr>
<td>axes abstracted from geomorphic or landmark-based system</td>
<td></td>
</tr>
<tr>
<td><em>The ball is downriver of the chair.</em></td>
<td></td>
</tr>
</tbody>
</table>

The following examples from languages of our sample illustrate some of the strategies. Example (1) is a description of the configuration found in the photo stimuli presented in Figure (1) in Yucatec Maya. This utterance combines three reference frames: the relator *tu-tséel* ‘at its side’ is interpreted (allocentrically) intrinsically, *te-xts’íik* ‘on the left’ relatively, and *te=chik’in* ‘in the west’ absolutely. The tendency to combine multiple different frame types in this manner is characteristic of this language (Bohnemeyer 2011).
The Tarascan (P’urhépecha) description of Figure 2 in (2) likewise features three frames. Tátsepani ‘behind’ is an (allocentric) intrinsic relator, whereas karháwa erákuteni ‘upward’ and pít’akuni ‘toward the wall’ are used here with two further subtypes of geocentric frames.

\[ \text{T-u=tséel, te=x-ts’iik te-estée-le=chik’in=o’, PREP-A3=side PREP:DET=F-left PREP:DET-HESIT-DET=west=D2} \]
\[ \text{hun-p’éel bóola yáan=i’, ch’uy-k’ah-a’n (…).} \]
\[ \text{‘On the (chair’s) side, on the left in the, uh, the west, there is a ball, it is suspended (…).’} \]

Figure 2. Ball & Chair 4.2

(2) Pelóta karháwa erákuteni, tátsepani, ximá pi-t’a-ku-ni.
‘The ball [is] in an upward direction, behind [the chair], there approaching the wall.’
(Capistrán-Garza 2011: 1022)

Karháwa erákuteni ‘upward’ is used in (2) with an instance of what we call a ‘geomorphic’ frame (Bohnemeyer 2012; Bohnemeyer & O’Meara 2012; O’Meara & Pérez-Báez 2011). The axes of this frame are derived from a gentle slope that dominates the local terrain of the field site, the town of Santa Fe de la Laguna in Michoacán, Mexico. Geomorphic frames differ from absolute ones in that their use is restricted to a
region proximate to the anchor. The same holds for ‘landmark-based’ frames such as illustrated by pit’akuni ‘toward the wall’. One axis of this frame is constituted by a vector pointing toward the anchor. Absolute frames can be understood as generalizing abstractions of geomorphic and landmark-based frames.

Lastly, (3) is part of a description of Figure 3 in East Highlands Mixe (cf. Romero-Méndez 2011) and illustrates the ‘direct’ frame type (Danziger 2010), in which the observer’s body serves as both anchor and ground.

![Figure 3. Ball & Chair 2.1](image)

(3) Mejts m-jën-tuujy ajá tam ējts n-jën-tuujy mejts m-jën-tuujy.

2SG 2POSS-front-LOC yeah like 1SG 1POSS-front-LOC 2SG 2POSS-front-LOC

‘In front of you, yes, like in front of me, in front of you.’

A central assumption of this study is that the use of particular reference frame types for particular referential purposes constitutes non-codified linguistic practice. To establish that this is the case, it must be shown that the use of particular reference frame types for particular referential purposes is:

- Learned rather than innate;
- Culturally transmitted and diffused;
- Not codified, i.e., not part of the grammar or lexicon of languages.

Let us address these points in turn. It has in fact been suggested that all major reference frame types are innately available to all human populations (Li & Gleitman 2002, Li et al 2011; cf., however, Levinson et al 2002, Haun et al 2011, and Bohnemeyer et al ms.. We address implications of the present study for this debate in Section 6). However, whether or not the knowledge of all reference frame types is innate, their use is indisputably not uniform across languages and cultures. In the study we present here, native speakers of Spanish – including speakers of rural Latin American varieties – made no more than marginal use of geocentric frames in the small-scale domain of the stimuli illustrated in Figures 1-3, whereas speakers of three Mesoamerican languages – Juchitán (Isthmus) Zapotec, Tarascan, and Tseltal Maya – made only marginal use of relative frames (cf. Figure 4 below). Moreover, most of our study populations made no use of the
Evidence that patterns of reference frame use are culturally shared comes from data collected at different times in different locations with different speakers of the same language using different instruments. Where such comparisons show intra-community consistency in the face of cross-community variation, this strongly suggests that population-specific usage profiles are cultural practices. Such comparisons are possible for two of the languages of our sample: Tseltal and Yucatec. In the case of Tseltal, Polian & Bohnemeyer (2011) report considerable variation across three different Tseltal-speaking communities, which does not include the village in which Brown & Levinson (1992, 1993) conducted their classic studies. All communities confirm to the no-more-than-marginal use of relative frames Brown and Levinson found, but none of them shows the prevalence of absolute or geomorphic frames. Instead, the dominant pattern is landmark-based frames for orienting entities and allocentric intrinsic frames for locative descriptions. This can be explained with reference to community adaptations to the local topography, which renders geomorphic descriptions more difficult to anchor compared to Brown and Levinson’s field site. Li et al (2011) report linguistic data consistent with the findings of Brown & Levinson and Polian & Bohnemeyer. In the case of Yucatec, Bohnemeyer & Stolz (2006), Bohnemeyer (2011), and Le Guen (2011) present a strikingly consistent picture of a language in which all major strategy types are used side by side, but in which absolute frames exhibit a strong demographic bias, being used almost exclusively by adult males. More direct support for the cultural transfer of reference frame use comes of course from the study we present here.

Lastly, as for codification, all spatial relators are at least potentially ambiguous or vague, permitting interpretations in multiple types of frames. Thus, the ‘front’/‘back’/‘left’/‘right’ terms used to designate horizontal regions without reference to environmental anchors are systematically compatible with allocentric intrinsic (‘object-centered’), egocentric intrinsic (‘direct’), and egocentric extrinsic (‘relative’) interpretations. For example, the ball may be said to be relatively ‘left of’ the chair in Figure 2 and intrinsically in Figure 3; a direct use would be ‘The ball is left of me’. Vertical relators such as ‘up’ and ‘down’, in addition to the aforementioned uses, also have absolute uses anchored to the gravitational vertical (Carlson-Radvansky & Irwin 1993, 1994). And geocentric relators are inherently compatible with both concrete environmental anchors and abstract absolute uses, as illustrated for Tseltal alan ‘down’ and ajk’ol ‘up’, which support both geomorphic uses anchored to the local mountain slope and absolute horizontal uses in which alan designates approximately 345° NNW and ajk’ol 165° SSE (in addition to intrinsic and gravitational vertical uses; cf Polian & Bohnemeyer 2011 for discussion). Bohnemeyer 2012 argues that the many-to-many relation between relators and reference frames is by no means accidental. Rather, it is the result of spatial relators not being specified in their lexical meaning for the anchors they occur with. Anchor selection is fundamentally contextual and therefore a pragmatic rather than a lexical-semantic property.

---

3 However, geomorphic uses in our Tseltal data are largely tied to the relators ajk’ol ‘up’ and alan ‘down’. These have both geomorphic and absolute uses, which are not easy to distinguish. See Polian & Bohnemeyer (2011) for discussion.
It is clear, then, that spatial relators are generally semantically compatible with a range of different frame types, and which of these they actually occur with depends on the practices of language use of individual speakers and their speech communities. Crucially, all linguistic varieties in our sample have the lexical and morphosyntactic resources necessary to make use of any of the frame types we distinguish in our analysis. Thus, the difference in frame use we observe across speakers and populations cannot be attributed to differences in lexicalization or grammaticalization, but are a matter of different practices of language use.

3 Mesoamerica as a Test Case

One of the questions we address in this study is whether practices of spatial reference may be shared among the speakers of a group of languages that form a linguistic area or sprachbund. Our test case is the Mesoamerican linguistic area. A sprachbund is a group of languages of distinct genealogical origins that have converged in grammar, lexicon, phonology, and/or (this is the hypothesis examined here) practices of language use as a result of sustained contact. There does not need to be a single feature that is shared among all the languages of the area; rather, the area may have the structure of a network of intersecting sub-networks of languages sharing different features locally. As a result, the precise boundaries of a sprachbund are often difficult to establish, and the membership of individual languages is often controversial and might perhaps best be considered a matter of degree rather than a categorical property (cf. Muysken 2008).

Every linguistic area is unique, and so is the evidence researchers rely on to establish or argue for its existence. The concept of ‘Mesoamerica’ was originally introduced by anthropologists in reference to a cultural area comprising the ancient civilizations of the Valley of Mexico, the Valley of Oaxaca, the Isthmus of Tehuantepec, the Mexican Gulf Coast, the Yucatan peninsula, the highlands of Guatemala, the coastal zones of Central America as far south as Nicaragua, and the regions in between these, as well as the modern descendants of these ancient peoples (Kirchhoff 1943; Carmack et al 1996). The archeological and historical evidence of intensive contact across the groups inhabiting this region is massive and incontrovertible. It extends back in time to several millennia prior to the European colonization.

However, a linguistic area can only be established on the basis of evidence for language contact – i.e., on the basis of linguistic evidence, although archeological and historical evidence can play a decisive role where the linguistic evidence is equivocal. The linguistic evidence for a Mesoamerican sprachbund was examined in detail in Kaufman (1973), Campbell (1979), Campbell et al (1986), and Smith-Stark (1994). Among the many features of phonology, morphosyntax, and lexicon considered in these works, perhaps the most persuasive are those that set most Mesoamerican languages apart from their geographic neighbors to the north and south, especially those neighbors that belong to the same language families:

- With only a few exceptions (including Tequistlatec, Huave, and some Oto-Manguean languages), Mesoamerican languages lack phonemic voicing contrasts in stops and fricatives, whereas some of their neighbors to the north (e.g.,
O’odham (or Pima); Raramuri (or Tarahumara)) and south (e.g., Sumu, Miskito; Chibchan) do have them.

- Except for the Mixean languages, verb-final constituent orders are absent from Mesoamerica, whereas both Uto-Aztecan and Yuman languages to the north and Chibchan and Misumalpan languages to the south are verb-final. Concomitantly, switch-reference marking, a feature that is typologically strongly correlated with verb-final order, is absent throughout Mesoamerica, but occurs in many of the northern and southern neighbors. Another correlated feature, adnominal possessors tend to follow the possessum in Mesoamerican languages, but precede it in Sumu and Miskito to the south(east) of the Mesoamerican area, though not in the Uto-Aztecan languages spoken north of the area.

- Mesoamerican languages tend to have few or no adpositions, relying instead on relational nouns and applicative morphology to express the corresponding meanings.

- Lastly, there is a large set of semantic calques that are shared across the Mesoamerican area. Aside from vigesimal numeral systems, this includes a set of more than 50 metaphors, among which feature prominently meronymic metaphors (e.g., ‘mouth’ of the house for the door; ‘excrement’ or ‘belly’ of the leg for the calf; ‘hair’ of a tree for the root; etc.).

The shared semantic calques are particularly interesting for present purposes, as they can be argued to constitute evidence of diffusion of semantic practices. However, these practices are in this case codified in the lexical meanings of particular lexical items. In contrast, the relation between spatial relators and the reference frames in which they are interpreted is not generally fixed in this manner, as argued in Section 2.

The criteria listed above, together with a large set of less widely shared properties, identify the following languages as belonging to the Mesoamerican area:

- The Mayan, Mixe-Zoquean, Oto-Manguean, and Totonac-Tepehuan language families;
- The Nawan (or Aztecan) branch or genus of the Uto-Aztecan language family;
- A number of linguistic isolates: Huave, Lenca, P’urhépecha (or Tarascan), Tequistlatec (or Oaxaca Chontal), and Xinca.

The language contact that brings about linguistic areas is a non-static phenomenon. In the case of the Mesoamerican area, it is likely that the nature of the contact relations between the participating languages has changed greatly over time as the influence of language groups throughout the region has changed. Most recently the influence of European languages has dominated the Mesoamerican area. It is the widespread presence of Spanish that allows us to use the Mesoamerican area as an ideal test case for the diffusion of linguistic practices via contact.

In a precursor to the present study, Eggleston (2012) compares the performance of Nicaraguan Spanish speakers from the same area as the Sumu-Mayangna speakers during a referential communication task to that of Barcelona Spanish speakers on the one hand and Sumu-Mayangna speakers on the other. The Europeans strongly preferred relative descriptions and used geocentric ones in under 5% of their uses, conforming to what has
been shown for speakers of Dutch (Pederson et al 1998) and English (Levinson 2003). The Nicaraguan Spanish speakers likewise showed a preference for relative frames, but a much more moderate one, intriguingly falling roughly midway between the Barcelonan Spanish speakers and the Sumu-Mayangna speakers in both their use of relative frames and that of geocentric ones. In the following section we present the populations whose data we use to explore such diffusion and the methods by which we test our predictions.

4 Methods and Predictions

4.1 Participants

We examined whether speakers of indigenous languages of Mexico and Nicaragua are influenced in their use of spatial reference frames by their use of Spanish as a second language. To this end, we compared the performance of speakers of eight indigenous languages of Mexico and Nicaragua in a picture matching referential communication task to that of speakers of three varieties of Spanish: Mexican, Nicaraguan, and Barcelonan Spanish. Table 2 lists the linguistic varieties involved, the researchers involved in data gathering and analysis, and the field sites where the data were collected.

<table>
<thead>
<tr>
<th>Linguistic variety</th>
<th>Genealogical affiliation</th>
<th>Membership in the Mesoamerican area</th>
<th>Field site(s)</th>
<th>Researcher(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Highlands Mixe Zapotec</td>
<td>Mixe-Zoquean</td>
<td>Yes</td>
<td>Ayutla, Oaxaca, Mexico</td>
<td>R. Romero Méndez</td>
</tr>
<tr>
<td>P'urhépecha (Tarascan)</td>
<td>isolate</td>
<td>Yes</td>
<td>Santa Fe de la Laguna, Michoacán, Mexico</td>
<td>A. Capistrán Garza</td>
</tr>
<tr>
<td>San Ildefonso Tultepec Otomi</td>
<td>Oto-Manguean</td>
<td>Yes</td>
<td>San Ildefonso Tultepec, Querétaro, Mexico</td>
<td>M. S. Hernández Gómez; N. Hernández Green; E. L. Palancar</td>
</tr>
<tr>
<td>Seri</td>
<td>Isolate</td>
<td>No</td>
<td>El Desemboque del Río San Ignacio, Sonora, Mexico</td>
<td>C. O’Meara</td>
</tr>
</tbody>
</table>

Some of these Nicaraguan Spanish speakers are bilingual in Miskito Coast Creole and/or indigenous languages other than Sumu, such as Miskito, another Misumalpan language. It is possible that these other languages have influenced their reference frame use.
Table 3 shows the localities of data collection and the distribution of sex and age.

Table 3. Participants by language, site, age, sex, and study (MA – Mesoamerican; NMA – non-Mesoamerican indigenous)

<table>
<thead>
<tr>
<th>Linguistic variety</th>
<th>Locality</th>
<th>Dyads</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt; 30 / ≥ 30</td>
<td>M / F</td>
</tr>
<tr>
<td>Tseltal (MA)</td>
<td>Chacoma, Chiapas</td>
<td>5</td>
<td>7/3</td>
<td>6/4</td>
</tr>
<tr>
<td></td>
<td>Yaxley, Quintana Roo</td>
<td>4</td>
<td>2/6</td>
<td>4/4</td>
</tr>
<tr>
<td></td>
<td>Felipe Carrillo Puerto,</td>
<td>1</td>
<td>0/2</td>
<td>1/1</td>
</tr>
<tr>
<td></td>
<td>Quintana Roo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>East Highlands Mixe</td>
<td>Ayutla, Oaxaca</td>
<td>5</td>
<td>3/7</td>
<td>3/7</td>
</tr>
<tr>
<td>(MA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otomi (MA)</td>
<td>San Ildefonso Tultepec,</td>
<td>4</td>
<td>0/8</td>
<td>1/7</td>
</tr>
<tr>
<td></td>
<td>Querétaro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isthmus Zapotec (MA)</td>
<td>La Ventosa, Oaxaca</td>
<td>6</td>
<td>4/8</td>
<td>3/9</td>
</tr>
<tr>
<td>Tarascan (MA)</td>
<td>Santa Fe de la Laguna,</td>
<td>5</td>
<td>4/6</td>
<td>4/6</td>
</tr>
<tr>
<td></td>
<td>Michoacán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seri (NMA)</td>
<td>El Desemboque, Sonora</td>
<td>5</td>
<td>1/9</td>
<td>2/8</td>
</tr>
</tbody>
</table>
We ran the task with five dyads of speakers per population – six in the case of European Spanish and Isthmus Zapotec. However, we were unable to obtain self-estimated literacy and education levels (cf. Section 5.2) from four of the 57 dyads and thus did not include the utterances produced by these four dyads in the analysis. Table 3 reflects only those participants whose productions were included in the analysis presented below.

<table>
<thead>
<tr>
<th>Language</th>
<th>Participants</th>
<th>dyads</th>
<th>self-estimate literacy</th>
<th>education levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sumu (NMA)</td>
<td>Rosita and Siuna, RAAN</td>
<td>5</td>
<td>2/8</td>
<td>5/5</td>
</tr>
<tr>
<td>Mexican Spanish</td>
<td>Chimalacatlán, Morelos</td>
<td>5</td>
<td>6/4</td>
<td>3/7</td>
</tr>
<tr>
<td>Nicaraguan Spanish</td>
<td>Rosita and Siuna, RAAN</td>
<td>4</td>
<td>0/8</td>
<td>2/6</td>
</tr>
<tr>
<td>European Spanish</td>
<td>Barcelona, Cataluña</td>
<td>4</td>
<td>2/6</td>
<td>1/7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>53</td>
<td>31/75</td>
<td>35/71</td>
</tr>
</tbody>
</table>

4.2 Procedure

To assess reference frame use in discourse, we conducted a referential communication task (Clark & Wilkes-Gibbs 1990) with members of 11 linguistic populations. The Ball & Chair (B&C) task is a photo-matching game consisting of four sets of 12 photos that show varying spatial configurations of a ball and chair. Figures 1-3 above are examples. A dyad of speakers sits side-by-side divided by a screen to prevent sharing their visual fields. Speakers each have identical sets in front of them. One speaker (designated the ‘director’ for that set) selects photos one by one and describes them with the goal of having the other speaker (designated the ‘matcher’) select the match from their own set. After each selection is made, the researcher records the selection, and the director places a marker on the photo to exclude it from further description. The matcher does not mark the photos they select, and therefore does not reduce the possible choices throughout the set. After all photos in the set are described, the researcher shows the participants their selected matches to confirm. Roles are reversed between sets.

To properly test for a possible effect of L2-Spanish use, we needed to guard against several possible confounds. Li & Gleitman (2002) hypothesize that reference frame selection may be driven by a variety of nonlinguistic factors. Previous research has associated a preference for relative frames in small-scale space with urbanized and highly literate and educated populations; e.g., speakers of Dutch and mainland Japanese (Pederson et al 1998; Majid et al 2004; Kita 2006). Empirical evidence of a literacy effect is presented in Danziger & Pederson (1998). In contrast, the use of geocentric frames may depend on the availability of suitable local landmarks and on a speech community that mostly inhabits a region in close proximity to these landmarks (cf. also Palmer 2001). The use of L2-Spanish covaries with these variables in our sample, as formal education and literacy are primarily in Spanish and the use of the indigenous languages is more

---

5 Región Autónoma del Atlántico Norte (Nicaragua).
6 However, primary school for Sumu children is currently entirely in Sumu. There is also some amount of bilingual education during early grades in several of the Mexican communities, including Isthmus (Juchitán) Zapotec and Yucatec. Isthmus Zapotec also has relatively high rates of Zapotec literacy.
prevalent in geographically more isolated communities. For this reason, we selected the Spanish-speaking communities in Mexico and Nicaragua to closely match the socioeconomic profile of the indigenous communities and the literacy and education levels typical of rural communities in Mexico and Central America.

To be able to isolate possible effects of L2 Spanish against these nonlinguistic variables, we asked the participants to assess their levels of literacy and education along with their frequency of L2-Spanish use (in the case of the speakers of the indigenous languages). We administered a survey that included a three-valued scale for frequency of Spanish use (none, occasional, frequent/regular) in various social contexts (the data from which we later collapsed for the purposes of the analysis), separate four-valued scales for the frequency of reading and writing (regularly, occasionally, rarely, never), and a three-valued scale for formal education level (no more than primary, any amount of secondary, any amount of post-secondary). Population density and topography for each field site were calculated from census data and maps (for Mexico: INEGI 2010; Hernandez Santana et al 2007, p.c.; for Nicaragua: INIDE 2005). Population density was assessed as the populations of each community divided by the area according to Google Earth. Population density and topography were assessed by community and as such apply to all the members of a community.

4.3 Coding
The directors’ locative descriptions were analyzed and coded for frame use based on the fine-grained six-way classification illustrated in Table 1 above. A number of further distinctions were added to the coding schema. We added a ‘topological’ (in the sense of Piaget & Inhelder 1956) category for frame-free descriptions. We furthermore distinguished between geocentric frames used with horizontal vs. with vertical relators. This decision was motivated by the privileged role the Earth’s field of gravity plays as an anchor of geocentric frames in the vertical across populations. We also singled out propositions that could not be assigned unambiguously to one frame type because the same proposition was true of the same stimulus picture in multiple frame types. Carlson-Radvansky & Irwin (1993) call such ambiguous representations ‘aligned’. We distinguished two types of ambiguity: alignment of intrinsic and relative frames, and alignment of intrinsic and geocentric frames with vertical relators.

While thus distinguishing a total of ten response categories in coding, we merged the three types of geocentric frames distinguished in Table 1 into a single undifferentiated geocentric category for the analyses described below in order to facilitate statistical modeling and because we often found it impossible to determine whether a given proposition was to be interpreted geomorphically or absolutely. This is because the same spatial relators are often compatible with both interpretations, and the two interpretations will often align in a given context in a similar fashion to the alignment types described above (see Polian & Bohnemeyer 2011 for discussion). Thus the number of dependent response variables in the analyses below is eight.

As illustrated by (1)-(2) above, the description of any given item may contain an indefinite number of spatial relators, each interpreted in a different frame type (in practice,
We thus assigned the picture descriptions eight binary response variables that recorded for each of the eight response types whether that type was instantiated by at least one relator in the description.

4.4 Predictions
Spanish is the dominant national language of both Mexico and Nicaragua. Every citizen of these countries nowadays has at least some amount of exposure to Spanish (even if they may be considered functionally monolingual in an indigenous language). The hypothesis that linguistic practices diffuse through language thus predicts that the use of reference frames among native speakers of the indigenous languages is influenced by their relative degree of exposure to Spanish.

To the extent that speaking a second language makes any difference at all for the use of spatial reference frames, the question arises whether it is the passive knowledge of the language – the speaker’s competence in the second language – or rather their active use that makes a difference for the speaker’s (linguistic and nonlinguistic) cognition – or perhaps in fact both of them. In this article, we restrict ourselves to testing the hypothesis that it is the use of the second language that may expose and habituate a speaker to cognitive practices not endemic in their culture.

Specifically, we predict that the frequency of use of Spanish as a second language is positively correlated with the frequency of use of relative frames and negatively with that of geocentric frames. In other words, the contact language Spanish should function as a conduit for the transmission of the European bias in favor of relative and against geocentric frames in small-scale space. In contrast, if non-codified linguistic practices do not “travel” together with elements of linguistic code in a contact situation, this predicts an absence of correlation between a speaker’s frequency of use of Spanish as an L2 language and their usage profile of spatial frames of reference. These are empirically testable predictions, which this article evaluates as a primary aim.

Furthermore, if linguistic practices such as the habitual use of particular reference frame types for particular purposes are indeed contact-diffused, they should also be sharable as areal features among genealogically unrelated languages. Since some, but not all, of the languages in our sample belong to a linguistic area – the Mesoamerican sprachbund (see Section 3) – this predicts that the speakers of Mesoamerican languages may be more similar to one another in their usage of spatial reference frames than to either the speakers of the three varieties of Spanish or the speakers of the non-Mesoamerican indigenous languages Seri and Sumu.

5 Results and Discussion

5.1 Distribution of the Responses
Since the participants switched roles between trials, we restricted the analysis presented here to the responses to the second and fourth set of the Ball & Chair pictures so as to keep the director constant and have a single set of demographic variables for each dyad. We furthermore confined the analysis to descriptions of the location of the ball with descriptions that employed more than three distinct types were rare).
respect to the chair in each picture\textsuperscript{7}. Figure \ref{fig:freq} shows the frequency of the eight response types discussed in §4.3 by linguistic variety.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{freq.png}
\caption{Frequency of response types by language in descriptions locating the ball vis-à-vis the chair (Linguistic varieties: BAR – European Spanish (Barcelona); JCH – Isthmus Zapotec (Juchitán de Zaragoza); MEX – Mexican Spanish; MIX – South Highlands Mixe; NIC – Nicaraguan Spanish; OTO – San Ildefonso Tulepect Otomí; SER – Seri; SUM – Sumu Mayangna; TAR – Tarascan; TSE – Tseltal; YUC – Yucatec Maya; Response types: AMB – ambiguous; TOP – topological (frame-free); INT – allocentric intrinsic; VERT – vertical geocentric; GEO – horizontal geocentric; DIR – egocentric intrinsic (direct); REL – egocentric extrinsic (relative))}
\end{figure}

The main findings that emerge from Figure 4 were anticipated in Section 2:

- No more than marginal use of geocentric frames in the three varieties of Spanish;
- No more than marginal use of relative frames in three of the Mesoamerican languages;
- Unrestricted occurrence of relative and geocentric frames in the remaining five languages.

Impressionistically, then, the participants differentiated themselves most strongly in their use of the relative and geocentric types. To test whether this was indeed so, we ran a Multi-Dimensional Scaling analysis (MDS). We assigned each of the 53 dyads a vector representing the frequencies with which they had used the eight response strategies. Interpreting these vectors as representing points in an octodimensional space, we constructed a 53x53 distance matrix, using the Manhattan Metric. Figure \ref{fig:mds} below shows a plot of the first two dimensions of a three-dimensional MDS model. We found that the

\footnote{Excluded from the analysis were descriptions of the disposition and location of ball within the picture, as well as the disposition and orientation of the chair.}
The first dimension of the MDS plot correlates strongly with both geocentric and relative reference frame uses (Spearman’s Rho = 0.85 for geocentric and -0.84 for relative; cf. Figures 6-7), confirming that these are the response types with the greatest amount of variation. The second dimension of the MDS plot correlates with the topological uses (Spearman’s Rho = -0.90; cf. Figure 8). Since topological uses are perspective-free, we ran regression models of just the use of relative and geocentric frames, the results of which we present in the next subsection.

**Figure 5.** MDS plot: the first two dimensions of the three-dimensional spatial model of the similarity matrix comparing the participants in terms of their use of the eight strategy types. The assessment is based exclusively on locative descriptions of Ball with respect to Chair. The dyads are labeled according to their native linguistic variety, using the same abbreviations as in Figure 4.

**Figure 6.** Correlation of MDS dimension 1 and geocentric descriptions
Figure 7. Correlation of MDS dimension 1 and relative descriptions

Figure 8. Correlation of MDS dimension 2 and topological descriptions

5.2 Impact of the Predictor Variables
To test our hypotheses about the contributions of independent variables on the use of a particular reference frame, we used generalized linear mixed-effects regression models (Gelman & Hill 2007, Jaeger 2008). Our models included fixed effects for our independent variables of L2 use of Spanish (L2-SPANISH USE), areal linguistic affiliation (LANGUAGE GROUP), education (EDUCATION), literacy as assessed as a maximum score of Reading or Writing (LITERACY), population density (POPULATION DENSITY) and local topography (TOPOGRAPHY). The LANGUAGE GROUP variable had three levels: Mesoamerican, Spanish, and non-Mesoamerican indigenous. We chose this grouping because models with an 11-valued L1 variable failed to converge, but also in order to test for a possible areal effect. We included random nested intercepts for participants and individual languages to avoid overfitting or lack of independence.
We regressed the probability of a given dyad of participants using a particular response type for a particular picture. The probabilities of using relative or geocentric frames are independent of one another – nothing prevented a participant from using both or neither in response to the same picture. We chose to run models for the two projective reference frame types that were strongly correlated with the first dimensions of our MDS plot, the relative and geocentric types (cf. Table 4 below). As mentioned in Section 5.1, these two dimensions account for the greatest amount of variance in responses. We ran two models for each of these dependent variables: a model that included \texttt{L2-SPANISH USE} and excluded Spanish speakers and one that conversely contained responses from all of our participants but omitted the \texttt{L2-SPANISH USE} variable. Stepwise reduction of the models comparing AIC goodness of fit was performed to present the simplest models to account for the data. \texttt{EDUCATION} made no significant contribution to any of the models and removing it as a variable from the models improved the AIC. These are the models that are presented below. \texttt{LANGUAGE GROUP}, \texttt{L2-SPANISH USE}, \texttt{POPULATION DENSITY}, and \texttt{TOPOGRAPHY} all made significant contributions in at least one model. As the \texttt{LANGUAGE GROUP} variable is categorical, a baseline must be specified. These models use the Non-Mesoamerican Indigenous Languages as a baseline. Table 4 summarizes the findings.\footnote{An anonymous reviewer asks how one can be confident that regression models based on a given number of observations support inferences about a relatively large number of independent variables. The principal answer to this question is that if the number of observations is insufficient given the number of variables and levels (values), the algorithm will fail to “converge” on a solution. Incidentally, our models are based on 1,272 observations (53 dyads x 2 sets x 12 items) in the case of the two large models and 960 (40 x 2 x 12) in the case of the smaller models that exclude the L1-Spanish speakers.}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Variables & 1 & 2 & 3 & 4 \\
\hline
Dependent & GEOMETRIC & & & \\
 & RELATIVE & & & \\
Indep. & L2-SPANISH USE incl. & & & \\
Results & LANGUAGE GROUP & *** & *** & *** \\
 & L2-SPANISH USE & *** & & \\
 & LITERACY & . & & \\
 & TOPOGRAPHY & * & & \\
 & POPULATION DENSITY & * & . & \\
\hline
\end{tabular}
\caption{Models using non-Mesoamerican Indigenous language type as a baseline (Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’)}
\end{table}

Table 5 differentiates the findings for \texttt{LANGUAGE GROUP} by individual group:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Variables & 1 & 2 & 3 & 4 \\
\hline
Dependent & GEOMETRIC & & & \\
\end{tabular}
\caption{Effects of \texttt{LANGUAGE GROUP} by group (Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’)}
\end{table}
5.3 Discussion
We now consider what the linear regression models suggest about each of our six independent variables, taking these up in turn.

5.3.1 Language Group
Language Group – i.e., whether the first language of the Director of a given dyad is Mesoamerican, non-Mesoamerican indigenous, or Spanish – is a significant independent factor in all models except for the geocentric model that excludes the L1-Spanish speakers (thus reducing the Language Group contrast to Mesoamerican vs. non-Mesoamerican). This is consistent with Figure 4, which shows that geocentric use was comparatively uniform across the speakers of the indigenous languages. The finding that language is a significant independent factor in reference frame use has important implications for the ongoing debate on the Linguistic Relativity Hypothesis, which we briefly address in Section 6.

We also found indirect evidence of a possible areal effect on reference frame use. As Table 5 shows, the speakers of the Mesoamerican languages differed significantly from the baseline, i.e., the speakers of the non-Mesoamerican indigenous languages, Seri and Sumu, in the use of relative frames. A comparison with Figure 4 suggests that the Mesoamericans used relative frames less frequently than the Seri and Sumu speakers and that this effect is primarily driven by the speakers of the three Mesoamerican languages who used relative frames only marginally, Juchitán Zapotec, Tarascan, and Tseltal. As it happens, the speakers of these three languages are also the most frequent users of geocentric frames among all 11 populations. However, these three languages are spoken hundreds of kilometers apart from one another and are not in systematic direct contact. It thus seems prudent to interpret Table 5 cautiously.

If there were to be no measurable areal effect in our data, we would be left with a situation that might appear almost paradoxical at first: The effect of the L2-Spanish use variable discussed in the next subsection suggests that the use of reference frames can be diffused through language contact. Why then should the languages of the Mesoamerican sprachbund not also share patterns of referential usage? But a moment’s consideration shows that there is actually no reason to believe that they do not. If the differences between the Mesoamericans and the speakers of the non-Mesoamerican indigenous languages are sufficiently small compared to the differences that set both of these groups apart from the Spanish speakers, our regression models might not detect the former, since they reach an acceptable fit with the data just by modeling the latter.

---

9 To say that two languages are not in contact with one another strictly speaking entails that they do not even share a single speaker. For languages such as the ones in question, with tens of thousands of speakers (hundreds of thousands in the case of Tarascan and Tseltal), this is difficult to verify.
What might be responsible for this comparative similarity in frame use across the Mesoamericans and their geographic neighbors? There are at least two conceivable factors involved here: First, language contact might extend beyond the Mesoamerican area to affect Seri and/or Sumu Mayangna as well.\textsuperscript{10} Since it is above all high geocentric, intrinsic, and topological scores that are shared among the speakers of the indigenous languages, one could hypothesize that a bias toward one or more of these strategies has been contact-diffused across the region.

Alternatively, it might be the case that it is specifically the use of relative frames that is contact-diffused – as our evidence from L2-Spanish use indeed suggests – whereas a bias toward geocentric frames represents an innate default. Higher geocentric scores would thus be shared across the speakers of the indigenous languages, not because of linguistic or cultural transmission, but because of biological transmission in combination with only limited contact with Spanish – assuming that it is contact with Spanish, along perhaps with education and literacy, that cause adoption of relative usage in the region. The possibility of a pan-simian geocentric bias was first raised by Haun et al (2006), based on evidence from studies with all species of Great Apes. Bohnemeyer et al. (ms.) present new evidence in support of Haun et al’s hypothesis, stemming from the results of a recall memory experiment across members of the 11 populations discussed here and also from a comparison of the behavior of five different Spanish-speaking populations: Barcelonan and Nicaraguan Spanish speakers and three different groups of Mexican Spanish speakers. It turns out that geocentric responses are preferred by the members of all those populations that do \textit{not} show a preference for relative frames in the Ball & Chair study. This includes several groups in which neither relative nor geocentric usage is prevalent in discourse, including the Yucatec speakers and the Spanish speakers from Rosita (Nicaragua) and San Miguel Balderas (Mexico). All of these groups appear to prefer geocentric solutions in the recall memory task. In Bohnemeyer et al (ms.), we interpret this to the effect of a hypothetical innate geocentric bias that may become restructured through a learned, culturally transmitted and diffused egocentric bias. Given that the geocentric bias is shared across populations that are culturally and linguistically unrelated, the hypothesis that it is innate does not seem too far-fetched. We suggest that geocentric and egocentric frames might represent different adaptations to the amount of control the individual exerts on their environment, with geocentric frames being a more advantageous strategy under low individual control and egocentric frames becoming more advantageous under higher individual control. Control might be reflected in the extent through which the individual restructured their environment through, for example, tool use.

5.3.2 L2-Spanish use
The significant L2-Spanish use effect in Model 4 of Table 4 (the relative model that excludes L1-Spanish speakers) indicates that the speakers of the indigenous languages in our sample are significantly more likely to use relative frames \textit{in their native languages} the more frequently they also use Spanish as a second language. We take this to be evidence that the use of relative frames diffuses through language contact. To our knowledge, this is the first quantitative demonstration of an apparent language contact

\textsuperscript{10} This possibility was suggested to us by Lev Michael [p. c.].
effect on practices of non-codified language use. It suggests that, along with sound patterns, syntactic constructions, lexical items, and idioms, languages may also share practices of language use through contact. Moreover, regardless of whether the frames are applied in the interpretation of linguistic or nonlinguistic representations, if reference frame use is taken to be primarily a cognitive practice (i.e., a ‘habitual’ style of ‘thought’), then linguistic transmission is implicated as a possible basis for the cultural transmission and diffusion of cognitive practices. We return to this issue in Section 6.

We found effects of L2-Spanish use exclusively when modeling the probability of use of relative frames. No effect showed up in the geocentric models. This of course makes sense, as there is no reason to expect that the use of geocentric frames is diffused through contact with Spanish.

5.3.3 The nonlinguistic variables
We found no effect of education or literacy in any of the models. Topography and population density made significant contributions toward predicting the use of relative frames in the smaller sample (Model 4), and population density was also a significant independent factor in geocentric use in that sample (Model 2). These two variables show a correlation at or above .6 in most models, meaning the models are unable to accurately discriminate between the effects of the two geographic variables. This is presumably the result of the small number of communities in our sample (12, including two different Yucatec-speaking communities) and the coarse grain of the two variables. Since these two variables are also the only community variables (with the exception of the Yucatec speakers), they effectively “suck up” as a package all the variance in the data that is in any way community-specific. Thus, it seems prudent to refrain from specific interpretations of these effects.

6 Beyond Language: The Linguistic Transmission Hypothesis

A growing controversy has arisen around the demonstration in Levinson (1996, 2003) and Pederson et al (1998) of a robust crosslinguistic alignment of reference frame use in language, recall memory, and spatial inferences (see also Le Guen 2011; Majid et al 2004; Mishra et al 2003; Vajpayee et al 2008; Wassmann & Dasen 1998; and several of the contributions to O’Meara & Pérez Báez (eds.) 2011). Pederson et al (1998) show that a bias for relative or absolute frames in discourse among the speakers of a language predicts a bias in the same direction in recall memory and placement inferences. They

---

11 However, alternative models run on the participants’ writing scores alone rather than the cumulative literacy variable showed an effect on the use of relative frames when all groups are considered, though not among the speakers of the indigenous varieties alone. This is consistent with previous findings pointing to literacy being a significant predictor of the ease with which participants discriminate between mirror images and identical copies of a figure rotated in a plane (Danziger & Pederson 1998; Danziger 2011). Literacy might directly habituate participants to discriminate between mirror images, but might also have a direct effect on reference frame use, and via it, indirectly on mirror image discrimination. Danziger (2011) hypothesizes that the discrimination of enantiomorphs requires the use of extrinsic reference frames.
suggest that language may be a causal factor in this alignment. Given that frame use is more varied across populations than within, communities must have some mechanism that allows their members to converge on the same biases. A population’s patterns of frame use form a cultural \textit{habitus} that, like all procedural cultural knowledge, can only be transferred across generations through observable behaviors such as speech and gesture. But Li & Gleitman (2002) argue against this view of the population-specific frame profile as a \textit{habitus} transferred through language. Levinson and colleagues view the cognitive ability to learn any frame type as innate, but the actual use of a particular frame type as learned, and its mastery as requiring habituation over significant periods of time (Haun et al 2011). In contrast, in Li & Gleitman’s account, all types of reference frames are innately available across populations, and the observed differences in preferred usage across populations are exclusively individual-level adaptations to variation in geography, literacy and education. These preferences are superficial and readily mutable in response to changes in the factors mentioned. Li & Gleitman assume the adaption to these factors to happen primarily ontogenetically in each individual, rather than phylogenetically in an entire community. Empirical evidence in support of this hypothesis is offered both by Li and Gleitman and in Li et al (2011), while Levinson et al (2002) and Haun et al (2011) provide counter evidence.

However, contrary to Li and Gleitman’s conjecture that the apparent role of language can be entirely attributed to covariation between language and nonlinguistic factors, we have shown above that language makes a significant independent contribution toward predicting frame use. It bears repeating in this connection that all the linguistic varieties in our sample have spatial relators that are compatible with relative, geocentric, and intrinsic frames. Thus, if our \textsc{language group} differ significantly from one another in frame use, and this variation cannot be explained entirely with reference to the nonlinguistic variables, then this suggests that reference frame use is a conventional property of the practices of language use of speech communities. These practices may very well reflect adaptations to nonlinguistic factors such as \textsc{topography} and \textsc{education}. But these adaptations appear to be at least to some extent transmitted and/or diffused across the members of the speech community, rather than to be purely a property of the behavior of the individual speaker.

Of course, that reference frame use is a property of cultural practices of language use does not mean that practices of reference frame use in discourse are causal forces in shaping practices of reference frame use in nonlinguistic cognition, as the neo-Whorfians maintain. However, in Bohnemeyer et al (ms.), we present the results of a recall memory study conducted with speakers of the same eleven linguistic varieties discussed here, with the upshot that \textsc{language group} here, too, makes an independent significant contribution toward accounting for the variation in frame choice. Whether this should be considered a language-on-thought effect in the sense of weak interpretations of the Linguistic Relativity Hypothesis is not entirely clear. On the assumption that reference frame selection is a property of language use rather than lexical or compositional semantics, the findings in Bohnemeyer et al (ms.) strictly speaking constitute evidence of an effect of language \textit{use} on thought, rather than of an effect of language-specific categories of the lexicon or grammar. And it is above all effects from lexical or grammatical categorization that much of the Whorfian debate has focused on.
In Bohnemeyer et al (ms.), we attempt to parcel out the role of language use in nonlinguistic cognition by suggesting the view of language as a transmission system for cultural practices of nonlinguistic cognition:

(4) **Linguistic Transmission Hypothesis (LTH):** “Using a language may facilitate the acquisition of cultural practices of non-linguistic cognition shared among the speakers of the language.”

Language is not the only such transmission system. In principle, the transmission or diffusion of cognitive practices may proceed through any observable behavior. Co-speech gesture has long been suggested to play a prominent role (Haviland 1993; Levinson (2003: 244-271; 280-325); Le Guen 2011); but agricultural, architectural, and religious practices may also matter (Bohnemeyer 2011).

### 7 Conclusions

We have provided quantitative evidence from a multi-population study to the effect that practices of language use may be diffused through language contact. More specifically, we have shown that even practices that do not themselves form a part of the lexicon or grammar of a language, such as the use of spatial frames of reference, nevertheless may co-diffuse together with elements of the grammar and lexicon in a contact situation. This is of course far from surprising, but it nevertheless has far-reaching potential implications. Contact diffusion of practices of language use might be as relatively mundane – from a structural linguistic and cognitive science perspective, though not from an ethnographic one – as sharing a speech act or speech event (Hymes 1972); say, the practice of uttering a verbal formula of a certain level of formality to greet or thank somebody or respond to a misfortune, etc. Yet on the other end of the continuum, diffusion of language use might introduce the members of a linguistic community to new styles of thinking or cognitive practices, as suggested by the Linguistic Transmission Hypothesis. The basis of the cultural transmission of any kind of procedural knowledge is the “infectiousness” of observable behavior – or rather, our social and cognitive readiness to be “infected” by (what we perceive to be) the practices of others. Language seems no exception to this principle.

### References


