

**Marcos de referencia en lenguas mesoamericanas:
un análisis multivariante tipológico**

**Jürgen Bohnemeyer^(a), Elena Benedicto^(b), Alejandra Capistrán Garza^(c), Katharine
T. Donelson^(a), Alyson Eggleston^(b), Néstor Hernández Green^(d), María de Jesús
Selene Hernández Gómez^(e), Jesse S. Lovegren^(a), Carolyn K. O'Meara^(e), Enrique
Palancar^(f), Gabriela Pérez Báez^(g), Gilles Polian^(h), Rodrigo Romero Méndez^(e), y
Randi E. Tucker^(a)**

(a) SUNY Buffalo, (b) Purdue University, (c) UAM, (d) CIESAS DF, (e) UNAM
(f) University of Surrey, (g) Smithsonian Institution, (h) CIESAS Sureste
b77@buffalo.edu

*Memorias del V Congreso de Idiomas Indígenas de Latinoamérica,
6-8 de octubre de 2011, Universidad de Texas en Austin*

Disponible en <http://www.ailla.utexas.org/site/events.html>

Marcos de referencia en lenguas mesoamericanas: un análisis multivariante tipológico *

Jürgen Bohnemeyer^(a), Elena Benedicto^(b), Alejandra Capistrán Garza^(c), Katharine T. Donelson^(a), Alyson Eggleston^(b), Néstor Hernández Green^(d), María de Jesús Selene Hernández Gómez^(e), Jesse S. Lovegren^(a), Carolyn K. O’Meara^(e), Enrique Palancar^(f), Gabriela Pérez Báez^(g), Gilles Polian^(h), Rodrigo Romero Méndez^(e), y Randi E. Tucker^(a)

(a) SUNY Buffalo, (b) Purdue University, (c) UAM, (d) CIESAS DF, (e) UNAM
(f) University of Surrey, (g) Smithsonian Institution, (h) CIESAS Sureste
b77@buffalo.edu

Palabras Claves: CILLA V, cultura, cognición, lenguaje, Marcos de Referencia, MesoSpace, Mesoamérica, multivariante.

1. Introducción

En los últimos tiempos, se ha observado un aumento en el interés por parte de los investigadores por los estudios sobre la relación entre lengua y cognición. Este interés se deriva de la propuesta de que algunos aspectos, no todos, de la cognición humana se comparten a través de la transmisión y difusión cultural. Las preguntas que quedan son, en primer lugar, ¿qué aspectos de la cognición son los que se comparten porque son innatos y cuáles son los que se comparten como parte de la cultura? En segundo lugar, y como principal propuesta de este trabajo, ¿cuál es el rol de la lengua en la cognición espacial? O más específicamente, ¿cuál es el rol de la lengua en la transmisión y difusión cultural de propiedades de la cognición?

La investigación que se ha producido en este sentido se ha concentrado en aquellos dominios cognitivos cuyas funciones están, suponemos, en principio, compartidas por todos los humanos. Los resultados que se han podido obtener hasta el momento incluyen las diferencias que se han descubierto entre hablantes de diferentes lenguas, en su actuación con respecto a la cognición espacial (Levinson, 2003; Pederson, *et al.*, 1998; Wassman y Dasen, 1998, entre otros) y también, por ejemplo a la percepción visual (Regier y Kay, 2009). Estos estudios son interesantes puesto que se ha observado una alineación entre los estilos cognitivos y lingüísticos de los hablantes al identificar objetos en un espacio local (Levinson, 2002; Majid *et al.*, 2004), al igual que entre la gestualidad y los estilos cognitivos relacionados con el espacio (Danziger, 2010; Haviland, 1993). En

* Queremos agradecer a nuestros profesores y consultores, los hablantes de las lenguas que estamos investigando; a nuestros colegas, los miembros del equipo de MesoSpace; a la National Science Foundation (NSF award #BCS-0723694 “Spatial Language and Cognition in Mesoamerica”), por darnos los recursos necesarios para realizar este estudio; al CIESAS y el Instituto Max Planck de Psicolingüística, por el apoyo logístico; a Marianne Gullberg, Steve Levinson y David Mark por sus consejos; a los organizadores del V Congreso de Idiomas Indígenas de Latinoamérica (CILLA-V).

otro estudio reciente (Haun, *et al.*, 2011) se investiga si es fácil o difícil para los niños usar estilos o estrategias cognitivas que no se usan frecuentemente en las comunidades. Los resultados sugieren que es difícil para los niños aprender y mantener el uso de estilos cognitivos que no son los preferidos en sus propias comunidades.

Por el contrario, también se ha llegado a la conclusión en algunos estudios importantes de que la facultad del lenguaje no afecta a la cognición espacial y, lo que es más, que las diferencias en actuación que se han observado entre hablantes de distintas lenguas pueden ser manipuladas modificando rasgos en el entorno medioambiental (Li y Gleitman, 2002). En este sentido, Li et al. 2011 afirman que debe ser fácil aprender un estilo que no es el preferido; según este análisis, no importa el grupo lingüístico o la comunidad a la que se pertenece (contrariamente a lo que se ve en Haun et al 2011).

El debate asociado con el lenguaje y su supuesto efecto sobre la cognición se halla en la base del presente estudio, el cual pretende contestar dos preguntas:

(1) ¿Qué relación puede establecerse entre la cultura y la cognición?

(2) ¿Cómo influye la lengua nativa en la cultura y la cognición de los hablantes?

Para poder contestar estas preguntas, es necesario disponer de un mecanismo sistemático para definir las estrategias de locación espacial. Los Marcos de Referencia (MdRs)¹ son sistemas de coordenadas que pueden servir para definir tanto zonas en el espacio como direcciones. Para clasificar los MdRs, hay dos conceptos que son de utilidad: el fondo y el ancla. La entidad de ‘fondo’ sirve como el punto de referencia en relación al cual se ubica el objeto que funciona como figura (Talmy, 2000); sin embargo, hay que tener en cuenta que esa entidad que sirve de fondo no siempre aparece representada en todos los MdRs (por ejemplo, los usados en las descripciones de orientación). El ancla, por su lado, es la entidad o pendiente del medioambiente que sirve de modelo para derivar los ejes del MdR (Danziger, 2010). Por ejemplo, en los MdRs de tipo egocéntrico, el cuerpo del observador es la entidad que sirve de ancla para el MdR. De igual manera, en los casos de MdRs geocéntricos se usa un elemento del medioambiente (p.ej., la inclinación de una ladera o la dirección de la corriente de un río) como ancla, mientras que en las descripciones que usan MdRs intrínsecos se usa el objeto de fondo mismo como ancla.

La variación en la distribución tipológica de los MdRs entre diferentes lenguas sugiere que la mayoría de las lenguas manifiestan una tendencia a preferir uno u otro marco de referencia y a evitar o rechazar los otros (Bohnmeyer y Levinson, *ms*; O’Meara y Pérez-Báez, 2011). Es importante notar que se observa una cierta variación en la categorización de los MdRs que se hace en la psicología y en la tipología semántica, como se ilustra en la Tabla 1 a continuación.

Tabla 1. *La clasificación de MdRs (como Bohnmeyer y Levinson, ms).*

¹ Las abreviaturas que se usan en este trabajo son: MdR = Marco de Referencia, MdRs = Marcos de Referencia, B&C = Ball and Chair, M&T = Men and Tree, L2 = segunda lengua, EMD = escalamiento multidimensional.

Clasificación en base al ancla (ej. Carlson-Radvansky & Irwin 1993; Wassmann & Dasen 1998; Li & Gleitman 2002; <i>inter alia</i>)		Clasificación por ancla y origen de ejes (Levinson 1996, 2003; Pederson 1993; Danziger 2010; Bohnmeyer & O'Meara, en prensa; <i>inter alia</i>)
egocéntrico	relativo (Levinson 1996) ancla = cuerpo del observador fondo ≠ ancla ejes proyectados (trasladados y/o reflejados) <i>La pelota está a la derecha de la silla.</i>	relativo
	egocéntrico intrínseco ('directo' según Danziger 2010) ancla = cuerpo del observador fondo = ancla ejes extendidos (sin proyección ni abstracción) <i>La pelota está frente a mí.</i>	intrínseco
alocéntrico	intrínseco centrado en el objeto (Carlson-Radvansky & Irwin 1993) ancla ≠ cuerpo del observador fondo = ancla ejes extendidos (sin proyección ni abstracción) <i>La pelota está en frente de la silla.</i>	
	geocéntrico tipo landmark ('proyectado' según Mishra et al. 2003; 'anclaje por punta (de vector)' según Bohnmeyer & O'Meara en prensa) ancla = elemento del entorno fondo ≠ ancla ejes definidos como vectores que apuntan hacia el ancla o alejándose del ancla <i>La pelota está en el dirección de la montaña a partir de la silla.</i>	
	geomórfico ('contextual' según Levinson 1996: 17) ancla = elemento entorno fondo ≠ ancla ejes proyectados <i>La pelota está rio debajo de la silla.</i>	
	absoluto (Levinson 1996; 'geográfico' según Jackendoff 1996) ancla = elemento del entorno fondo ≠ ancla ejes abstraídos a partir de un sistema geomórfico o tipo "landmark" <i>La pelota está rio debajo de la silla.</i>	absoluto

Por un lado, en la investigación que se hace en el campo de la psicología, se clasifican los MDRs en egocéntricos y aloecéntricos, y esta distinción se hace únicamente sobre la base del anclaje del marco en cuestión. Desde esta perspectiva, los MDRs intrínseco y geocéntrico forman la supra-categoría de marcos de anclaje aloecéntrico (cf. Carlson-Radvansky e Irwin, 1993; Li y Gleitman, 2002; Wassman y Dasen, 1998).

Por otro lado, el programa de investigación que surgió a partir del campo de la tipología semántica hace distinciones más tamizadas con respecto a los MDRs. Desde esta

perspectiva, los procesos cognitivos que se invocan para derivar los ejes de los MdRs a partir del ancla, juegan un papel secundario pero importante en la clasificación a la que finalmente se llega. Por ejemplo, para usar MdRs relativos, los hablantes deben trasponer o reflejar los ejes del ancla; sin embargo, para usar los MdRs intrínsecos los hablantes deben extender los ejes del ancla hacia afuera en el espacio. Esta es una distinción importante, pues permite distinguir entre MdRs relativos y MdRs intrínsecos egocéntricos, ambos con un anclaje egocéntrico. En el caso de los MdRs intrínsecos egocéntricos, el cuerpo del hablante sirve a la vez de ancla y de fondo, y, por tanto, los ejes se derivan por su extensión hacia afuera en el espacio, más que por reflejo o trasposición como sucedería en el caso de los MdRs relativos. La distinción entre MdRs relativos y MdRs intrínsecos egocéntricos acaba siendo una diferencia tipológicamente significativa, pues, mientras que todos los hablantes de todas las lenguas usan MdRs egocéntricos y MdRs intrínsecos, no todos los hablantes usan MdRs relativos (cf. Levinson, 1996, 2003; Pederson, 1993; Danziger, 2010). Por consiguiente, al identificar los mecanismos a partir de los cuales se generan los ejes, se pueden identificar distinciones significativas a nivel tipológico.

La investigación que presentamos aquí está motivada por las predicciones opuestas que hacen la propuesta innatista y la neo-whorfiana (véase la siguiente sección para una discusión más detallada), con respecto al lenguaje y la cognición espacial. Para poder evaluar la validez correspondiente de estas predicciones, se realizó un análisis estadístico multivariante, que pudiera llegar a medir y comparar las preferencias de MdRs de los hablantes de 11 lenguas y variantes lingüísticas. Se identificaron las preferencias de los hablantes por ciertos MdRs, en base a su comportamiento en una tarea de comunicación referencial. En total se incluyen en este estudio 6 lenguas mesoamericanas, 2 lenguas indígenas no-mesoamericanas, y 3 variedades del español. Si bien se han podido identificar varios factores interrelacionados como posibles factores de influencia en la preferencia de MdRs, los resultados estadísticos que aquí se presentan indican que el lenguaje continúa siendo el factor más relevante (pero véase la puntualización sobre el sentido de ‘lenguaje’ usado en este contexto, en la sección 2 a continuación).

Así pues, este trabajo está organizado en 5 secciones: la sección que sigue a continuación explora las diferencias en las propuestas entre innatistas y neo-whorfianos y elabora las predicciones que se derivan de ambas propuestas. La tercera sección identifica la metodología de las tareas y los grupos por lenguas que se han usado en este estudio; asimismo, se explican los métodos analíticos usados para generar el análisis descriptivo de los MdRs usados tanto por los hablantes de las lenguas de Mesoamérica que se han incluido en este estudio, como por los hablantes de las lenguas control no mesoamericanas. La cuarta sección presenta una discusión de los resultados y hallazgos que han surgido, mientras que la quinta sección concluye con los planes futuros para el proyecto en tanto que se expande para incluir análisis inferenciales, así como tareas y lenguas adicionales.

2. Hipótesis y predicciones de cada lado del debate

En esta sección destacamos dos posturas encontradas en cuanto a la relación entre lengua y cognición, esquemáticamente designadas como postura innatista o universalista por una parte y postura neo-whorfiana o relativista por otra parte. Levinson (1996) y

Pederson *et al.* (1998) observaron que el uso de MdRs en tareas de memoria varía según la lengua de los participantes. Para tratar esta co-variación, propusieron una explicación neo-whorfiana: que el lenguaje es un factor causal significativo en la selección de los MdRs. Li y Gleitman (2002) cuestionaron esta hipótesis y propusieron que la co-variación en el uso de MdRs en situaciones discursivas y en cognición no-lingüística podía explicarse en términos de factores externos al lenguaje, tales como la educación (formal), el nivel de alfabetización, la topografía local, la cohesión geográfica de la comunidad y la infraestructura; para ellos estos factores externos guiaban el uso de MdRs tanto en tareas lingüísticas como no-lingüísticas. Esta posición es la que se identifica como innatista o universalista. En este trabajo, presentamos un *test* específico para evaluar las predicciones de estas dos propuestas. Para ello, interpretamos las predicciones de estas dos posiciones, tal como aparecen en (i)-(ii):

(i) Hipótesis Neo-whorfiana:

El lenguaje es un factor más decisivo en la elección y uso que se hace de MdRs, que la educación (formal) y/o alfabetización.

(ii) Hipótesis Innatista:

La educación (formal) y la alfabetización son factores más decisivos en la elección y uso de MdRs, que el lenguaje.

De manera que, si el lenguaje acaba siendo un factor más fuerte que las variables no-lingüísticas mencionadas, ello daría apoyo a la posición neo-whorfiana (ya que, en tal caso, sería difícil explicar el efecto del lenguaje en términos de otras variables). Por el contrario, si resulta que alguna de esas otras variables no-lingüísticas acaba comportándose como un predictor más fuerte, ello daría apoyo a la propuesta innatista.

Hay que aclarar que ambas posturas comparten ciertos presupuestos acerca de lo que se entiende por lenguaje y por disponibilidad de MdR. En primer lugar, asumen que el conjunto de MdRs que hay disponible, en general, en el discurso lingüístico y en la cognición no-lingüística es el mismo; es decir, tanto las lenguas *per se* como la cognición no-lingüística escogen sus opciones del mismo conjunto total de MdRs. En segundo lugar, asumen también que cualesquiera que sean las restricciones o los *biases* que se den en una variedad lingüística en particular, éstos no se deben ni a la gramática ni al léxico de esa variedad lingüística. Es importante recalcar este último punto, pues se halla en el origen de ciertos malentendidos acerca de lo que se entiende por lenguaje en este contexto. Lenguaje sería pues, el conjunto de prácticas lingüísticas que tiene una comunidad de hablantes en particular, incluyendo la presencia o ausencia de segundas lenguas, entre otros factores.²

En el problema que nos interesa aquí, el de los marcos de referencia espaciales, estas dos posturas hacen predicciones diferentes, susceptibles de ser puestas a la prueba por los hechos, en particular a través de datos comparativos obtenidos en experimentos

² Esta noción de ‘lenguaje’ no estaría muy lejos de lo que algunas personas identificarían como ‘cultura’, o como la intersección de cultura/lenguaje, es decir, un sistema interpretativo de la realidad donde se haya la comunidad de hablantes, y que se transmite a través del lenguaje (en sentido más estricto). En particular, esta visión del lenguaje/cultura implica que se podría en principio cambiar las prácticas referenciales (re, uso de MdR), a través de capacitaciones u otro método, sin necesidad de cambiar la variedad dialectal que usan.

diseñados para este fin (véase la sección 3 sobre el experimento).

La postura innatista predice que las preferencias en el uso lingüístico y no lingüístico de los MdRs observadas para cada población se explican mayormente por factores externos a la lengua que moldean este uso. Estos factores incluyen el nivel de cohesión geográfica de esta población y la existencia de elementos ambientales cuyo conocimiento compartido incita a usarlos como puntos de referencia espaciales de manera eficiente. En este trabajo, sin embargo, vamos a considerar otro tipo de factores: el nivel de escolarización y de alfabetización.

Según la postura innatista, estos factores son susceptibles de influenciar en mayor medida el uso de MdRs en una población, en tanto que la escuela produce el tipo de cohesión interpersonal que puede afectar la selección de estrategias comunicativas espaciales entre la gama de recursos universalmente disponibles a los seres humanos. Más en concreto, se ha propuesto que las poblaciones más escolarizadas/alfabetizadas tenderán a ser espacialmente más *egocéntricas*³ (Li y Gleitman, 2002: 288), probablemente por el hecho de que aprender a leer y escribir implica el dominio de convenciones de orientación desde el punto de vista del que lee o escribe ("de izquierda a derecha", etc.) y el poder discriminar entre posibles signos que muestran una simetría en un eje vertical, como las letras "p" y "q" (véase también Levinson, 2003: 194).⁴ Se esperaría por lo tanto, siguiendo esta línea de razonamiento, que la escolarización/alfabetización pudiera cancelar una preferencia geocéntrica en una población, es decir, que podremos encontrar una comunidad lingüística donde los que no han ido a la escuela y/o no saben leer y escribir serán geocéntricos, mientras que el sector de la población que haya sido escolarizada/alfabetizada favorecerá estrategias egocéntricas relativas. En este sentido, el factor de escolarización/alfabetización resultaría más determinante que la lengua en sí, invalidando la tesis neo-whorfiana.

Por el contrario, la postura neo-whorfiana predice que la escolarización/alfabetización juega un papel secundario en el establecimiento de preferencias de selección de MdRs en cada población, en la medida en que la lengua compartida es el principal vehículo de transmisión de estas preferencias, más allá de las circunstancias extralingüísticas tanto geográficas como educativas. Si esta postura está en lo correcto, esperamos, por lo tanto, observar una mayor variación en el uso de MdRs entre hablantes dentro de una misma comunidad lingüística,⁵ y se predice una correlación menos fuerte entre el nivel de escolarización/alfabetización y el uso de MdRs. Se predice también que el dominio de una segunda lengua, en tanto factor lingüístico, puede ser determinante en este ámbito. En consecuencia, esperamos observar casos de influencia de

³ Li y Gleitman no distinguen entre los tipos egocéntricos y relativos. Los tratan como si fueran la misma cosa. Sin embargo, la variación tipológica se observa sólo en el uso de los marcos relativos, no en el de los marcos egocéntricos intrínsecos. Así que se debe entender la predicción de estos autores en el sentido de una correlación en el uso de marcos relativos y el nivel educativo.

⁴ Esto, por supuesto es únicamente válido para sistemas educativos occidentales. Quedaría por determinar qué podría suceder con otros tipos de sistemas educativos y, en especial, cuál sería el papel de un sistema educativo, en situaciones bilingües o multilingües, que transmitiera las preferencias de una 'comunidad lingüística' en particular y no de otra.

⁵ Entendiendo por 'comunidad lingüística' lo que hemos establecido en la página anterior: el conjunto de prácticas lingüísticas que tiene una comunidad de hablantes en particular, incluyendo la presencia o ausencia de segundas lenguas.

una segunda lengua, en el sentido en que los hablantes bilingües podrán mostrar preferencias de uso de MdRs acordes a la segunda lengua en contradicción con las preferencias asociadas a su primera lengua.

En síntesis, las predicciones de cada lado del debate en cuanto a los factores que determinan las preferencias de uso de los MdRs son las que se muestran en la Tabla 2.

Tabla 2. *Factores determinantes de preferencias de uso de MdRs.*

Determinantes más poderosos:	Postura innatista	Postura neo-whorfiana
Nivel de escolarización y de alfabetización:	sí	no
Lengua nativa y segunda lengua:	no	sí

En este trabajo, pues, vamos a comprobar estas predicciones mediante la observación de los niveles de influencia del uso de la primera y segunda lengua, los niveles de educación (formal) y la alfabetización en la elección y uso de MdRs por parte de los participantes en nuestro estudio.

3. Datos y métodos

3.1. La tarea de Ball & Chair

La tarea que utilizamos para estudiar el uso lingüístico de MdRs es la tarea de Ball & Chair (B&C), una tarea de comunicación referencial y que remplaza el uso de la tarea de Men & Tree (M&T) de Pederson *et al.* (1998). La tarea de B&C está compuesta de cuatro juegos de 12 fotos, y cada foto muestra una silla y una pelota en varias configuraciones espaciales (véase la Figura 1). Para cada prueba, dos hablantes observan copias de las mismas 12 fotos y tratan de encontrar las fotos que corresponden una a la otra. Los hablantes están sentados uno al lado del otro, con una pantalla o mampara entre ellos para bloquearles el acceso visual, así que tienen que resolver la tarea usando descripciones verbales explícitas. Este estímulo nuevo nos permite descubrir preferencias en la selección entre todos los tipos de MdRs en dominios a la escala de cuarto/habitación. De hecho, la tarea de M&T oculta usos intrínsecos por varias razones. Nuestro estudio se realizó a base de cinco parejas de participantes por lengua, salvo en el caso del español mexicano, donde hasta la fecha sólo tenemos los datos de tres parejas codificados.

Figura 1. *Dos fotos de Ball & Chair, mostrando un contraste intrínseco.*



Junto con los datos de la tarea de B&C, se recopilaron datos demográficos de los hablantes como factores de edad, sexo, nivel de educación, frecuencia del uso del español como segunda lengua (L2), y frecuencia en el uso de lectura y escritura. En este informe preliminar, los datos demográficos obtenidos sólo son aquellos datos reportados por los investigadores, pero para análisis futuros tenemos la intención de ir recopilando los datos directamente de los participantes.

3.2. Idiomas incluidos en el conjunto de datos

La investigación se basa en datos provenientes de seis lenguas nativas de Mesoamérica, los cuales se contrastan con los obtenidos en tres lenguas que sirven de control. Las lenguas estudiadas que forman parte del área lingüística mesoamericana (*Mesoamerican sprachbund*) abarcan tres familias y una lengua aislada:

Familia maya: tseltal (G. Pollian) y maya yucateco (J. Bohnemeyer).

Familia mixe-zoque: mixe de Ayutla (R. Romero).

Familia oto-mangue: zapoteco de Juchitán (G. Pérez) y otomí (N. Hernández, S. Hernández y E. Palancar).

Lenguas aisladas: purépecha (A. Capistrán).

Las lenguas de control son el seri (C. O'Meara), lengua aislada hablada en el noroeste de México, el sumu-mayangna (E. Benedicto, A. Eggleston, Mayangna Yulbarangyang Balna) perteneciente a la familia misumalpan hablado en Nicaragua, y variantes del español de México (R. Romero), de Nicaragua y de España (E. Benedicto, A. Eggleston).

3.3. Análisis de los datos

Para realizar el análisis, nos limitamos a las descripciones de los juegos 2 y 4 de la tarea de B&C; por cada uno de los 54 participantes y por cada tipo de marco se contó el número de fotos en las cuales se describió cada uno de los tipos de marcos. Después se clasificaron los marcos de referencia o las estrategias de referencia y se distinguieron ocho categorías:

- 1) Marco intrínseco centrado en un objeto.
- 2) Marco intrínseco “directo”, marco egocéntrico.
- 3) Marco relativo.
- 4) Ambigüedad entre marco intrínseco y relativo.
- 5) Marco geocéntrico.
- 6) Marco vertical absoluto.
- 7) Ambigüedad entre marco absoluto y vertical intrínseco.
- 8) Estrategia topológica (sin marco).

Después de codificar las respuestas de los participantes en términos de las categorías arriba mencionadas, se calculó un conjunto de ocho frecuencias para cada participante. Los conjuntos resultantes de este cálculo se pueden interpretar como las coordenadas de puntos en un espacio *octodimensional*. Las distancias entre estos puntos representan la

semejanza entre las respuestas de los participantes de B&C. Esta medida de semejanza se puede utilizar para comprobar las predicciones anteriormente mencionadas sobre el uso de MdRs. Esto es, según el bando innatista, los participantes deben agruparse sobre todo según sus niveles de alfabetización y educación, mientras que el bando neo-whorfiano predice que los participantes se agrupan sobre todo de acuerdo con su lengua nativa y su segunda lengua. Se calcularon las distancias entre los participantes en la métrica *city block* (o *Manhattan*), donde la distancia entre dos puntos se define como la suma del valor absoluto de las diferencias de sus coordenadas.

Para el análisis de la matriz de estas distancias (interpretadas como una representación de las semejanzas entre los participantes en términos del uso de MdRs) utilizamos varias herramientas metodológicas que nos permitieron obtener resultados confiables y precisos. La primera herramienta que usamos fue el algoritmo *Neighbor-net* (Huson y Bryant, 2006), que representa la distancia entre dos puntos como un conjunto de cambios alternativos entre dos nodos terminales de un grafo. Esto permite la representación bidimensional de las distancias con distorsión mínima. Este algoritmo fue diseñado para aplicarse a datos biológicos, porque permite observar de manera simultánea posibles relaciones filogenéticas entre organismos. Sin embargo, el mismo modelo matemático se puede aplicar para identificar posibles agrupaciones en cualquier grupo de datos que cuentan con un número alto o desconocido de “dimensiones”, de factores/variables que predicen su comportamiento.

Aunque este algoritmo se ha usado anteriormente en estudios de tipología sintáctica (Cysouw 2007; Bickel 2010), nuestro estudio es pionero en aplicar esta herramienta en la tipología semántica, además de ser el primer estudio experimental de este tipo. Asimismo, nuestro estudio es el primero en aplicar la estadística multivariante al análisis de la variación entre hablantes. Otras aplicaciones de estadística multivariante en la tipología semántica, por ejemplo, han examinado la semejanza en la codificación de los estímulos entre los hablantes de una muestra de lenguas (ver Levinson y Meira, 2003; Majid, Boster y Bowerman, 2008). Para realizar el análisis, se ingresó la matriz de las distancias en el programa *Splitstree4* (Bryant y Moulton, 2004).

La segunda herramienta que utilizamos fue el escalamiento multidimensional (EMD) (*Multidimensional Scaling – MDS* – por sus siglas en inglés). El escalamiento multidimensional es un conjunto de métodos que se utilizan para identificar los parámetros de variación más relevantes en datos que presentan un número alto o desconocido de tales parámetros (cf. Schiffman *et al.*, 1981). Nuestros datos fueron analizados utilizando el comando *cmdscale* en R (R Development Core Team, 2011). Por ello, el método de *Neighbor-net* y el escalamiento multidimensional son dos maneras distintas de visualizar la estructura de la misma matriz de correlaciones o semejanzas que describe el uso de MdRs por los participantes en nuestro estudio.

Los datos demográficos son producto de las estimaciones de nuestros investigadores. Por cada participante, se asignaron valores para edad, sexo, nivel de educación, uso del español, y saber (o frecuencia) leer y escribir. Las categorías son las siguientes:

Edad: Cuando se realizaron las grabaciones, el participante era (i) mayor o igual a 30 años o (ii) menor de 30 años.

Sexo: El participante es (i) un hombre o (ii) una mujer.

Nivel de educación: El participante (i) no tiene educación formal, (ii) ha cursado algún nivel de educación formal, (iii) ha completado algunos años de educación secundaria o más.

Uso de español: El participante (i) no sabe español y no puede hablarlo, (ii) sabe español y puede hablarlo, pero no lo usa regularmente/habitualmente, (iii) usa el español regularmente/habitualmente.

Saber leer y escribir: El participante (i) no puede leer o escribir, (ii) puede leer o escribir pero no lo hace habitualmente, o (iii) lee y escribe habitualmente.

En este análisis preliminar, se usan solamente estimaciones de los valores de las variables demográficas. Se planea recoger estos datos de los participantes directamente, en el futuro, por medio de un cuestionario.

4. Resultados

En la Figura 2 se ve el diagrama *Neighbor-net* de los participantes que obtuvimos. Cada nodo terminal representa a un participante. Las etiquetas codifican las lenguas nativas y los caminos entre los nodos terminales representan la semejanza entre los participantes en términos de los tipos de MdRs que usaron. Este *Neighbor-net* tiene dos polos salientes, aproximaciones de los cuales se muestran por los dos discos en la figura 2. El primer polo se caracteriza por altos valores relativos y el segundo por altos valores geocéntricos.

Figura 3. Gráfico EMD.

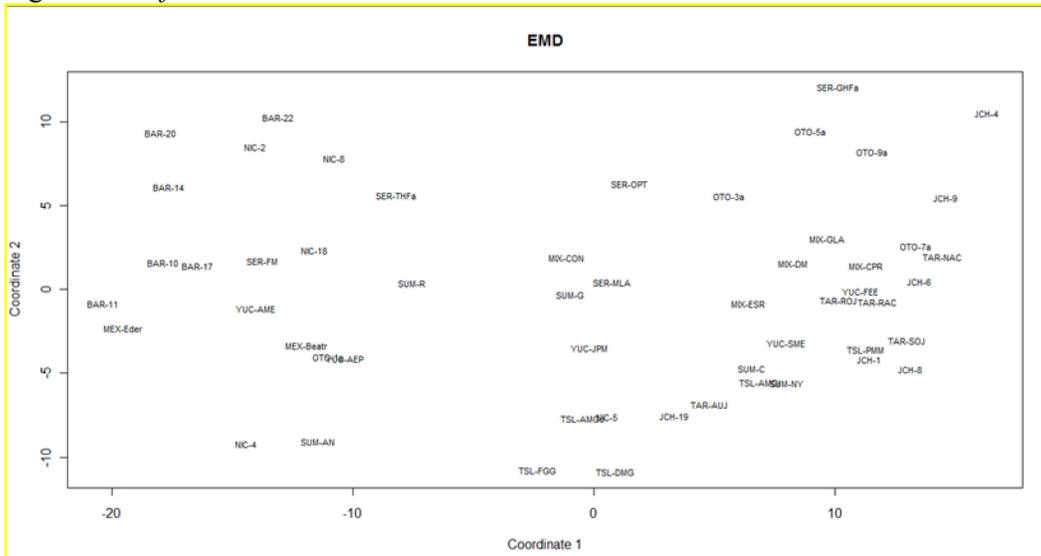
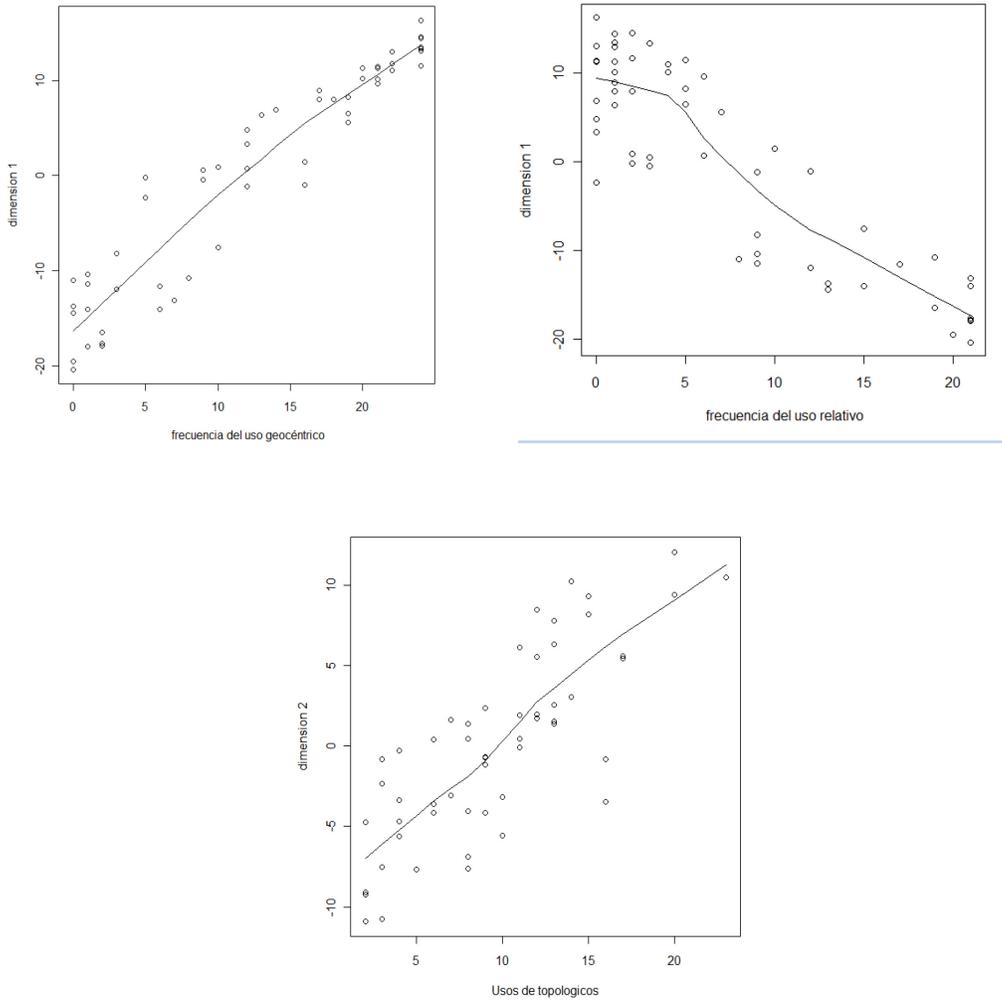
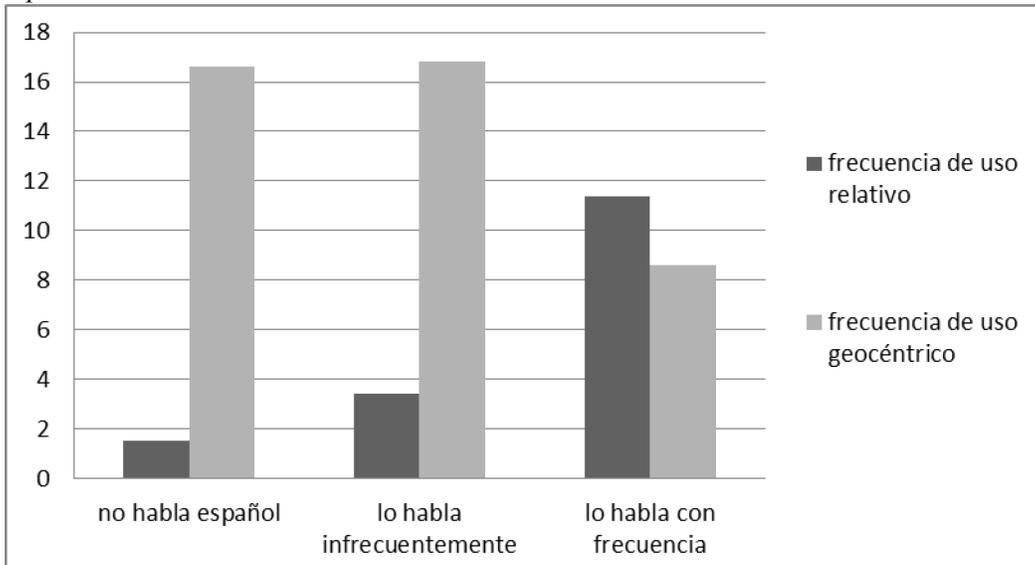


Figura 4. Correlaciones entre las dimensiones del gráfico EMD y la frecuencia de descripciones geocéntricas (izquierda), relativas (derecha), y topológicas (abajo).



A partir de este momento nos referiremos al polo del gráfico con altos valores relativos como la *zona relativa* y al polo con altos valores geocéntricos como la *zona geocéntrica*. La Figura 5 muestra el *Neighbor-net* con los nodos terminales etiquetados a través de las lenguas nativas de los participantes. Agrupamos las lenguas en tres categorías: lenguas mesoamericanas, lenguas indígenas no mesoamericanas, y las tres variedades del español. Se ve que los participantes se agrupan muy fuertemente por su lengua nativa: todos los hablantes de los dialectos del español se ubican en la zona relativa, con una sola excepción, una pareja de hablantes del español nicaragüense (cf. Figura 2) que usó los marcos relativos con una frecuencia comparativamente baja. En contraste, la zona geocéntrica está poblada por hablantes de lenguas indígenas, sobre todo

Figura 8. *Frecuencia de marcos relativos y geocéntricos por frecuencia del uso del español.*



Ahora la versión del gráfico que se ve en la Figura 9 tiene los nodos terminales etiquetados por los niveles educativos de los participantes. De esto también surge predecir el uso de los MDRs: entre los participantes con altos valores relativos, todos menos dos, tienen educación formal. Este resultado respalda al bando innatista. Sin embargo, otra vez no se observa una tendencia saliente en la zona geocéntrica. El gráfico en la Figura 10 sugiere que la frecuencia media del uso de los marcos relativos crece con la presencia de la educación secundaria.

Figura 9. El Neighbor-net codificado con colores por nivel de la educación.

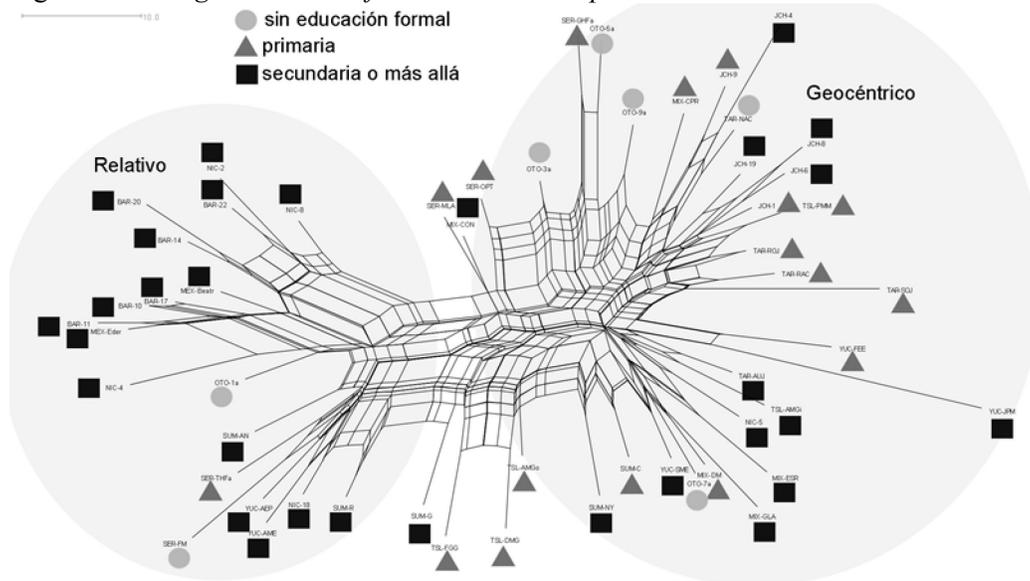
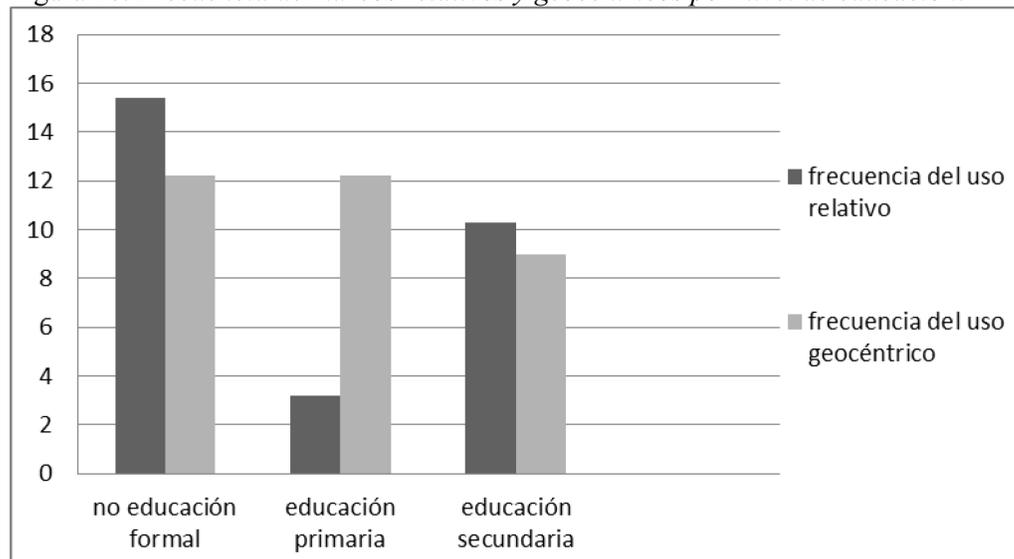
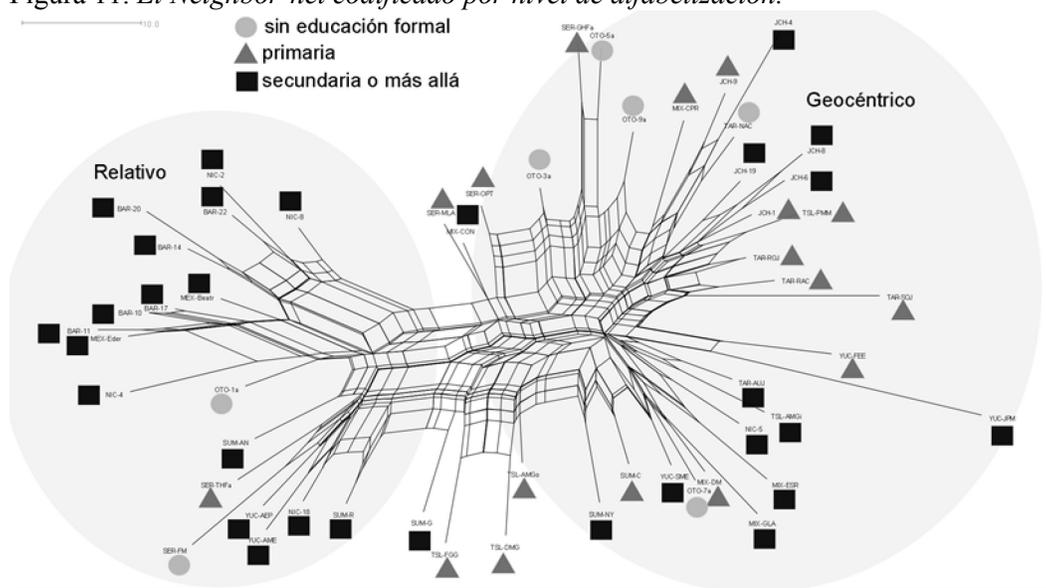


Figura 10. Frecuencia de marcos relativos y geocéntricos por nivel de educación.



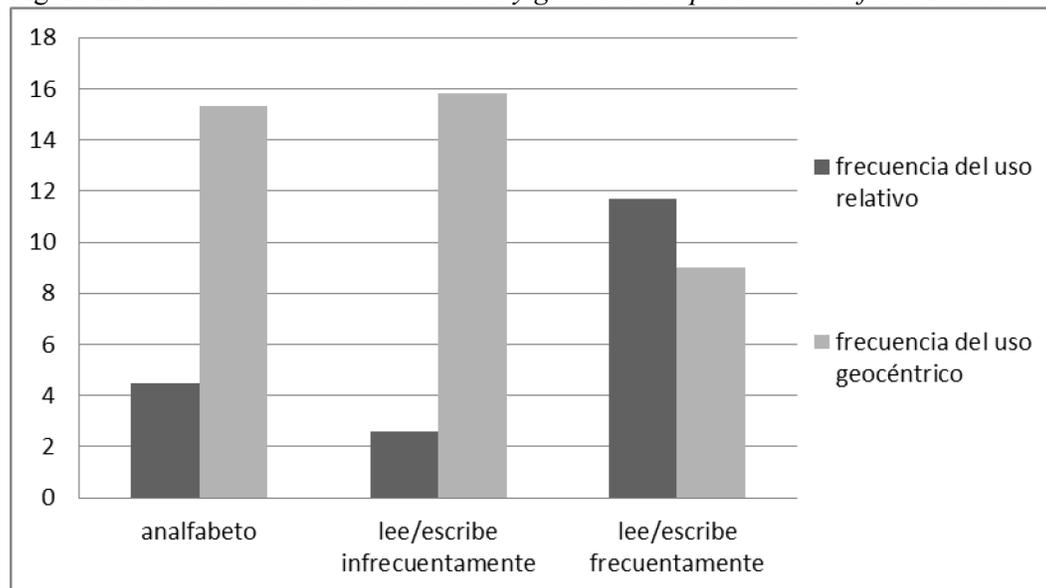
Y por último, también surge evidencia de un efecto de alfabetización conforme a las predicciones del bando innatista. En la Figura 11 vemos que entre los participantes con altos valores relativos, todos menos dos, están alfabetizados. Aquellos que son analfabetas son los mismos dos sin estudios formales. Y otra vez no se observa ninguna tendencia obvia en la zona geocéntrica.

Figura 11. *El Neighbor-net codificado por nivel de alfabetización.*



El gráfico en la Figura 12 confirma que la mayoría de los participantes que suelen leer y escribir con regularidad prefieren marcos relativos, mientras que la mayoría de los demás participantes prefieren marcos geocéntricos.

Figura 12. *Frecuencia de marcos relativos y geocéntricos por nivel de alfabetización.*



En resumen, hasta este punto parece que hay evidencia para todos los efectos predichos. Sin embargo, el efecto de la lengua nativa parece ser más fuerte que los demás, porque es el único que se manifiesta tanto en la zona relativa como en la geocéntrica. Además, los cuatro factores son obviamente interdependientes: ser hablante de una lengua indígena continúa estando correlacionado con niveles más bajos de educación formal y de alfabetización debido a la persistente marginalización socioeconómica. Los gráficos en las Figura 13 y 14 confirman estas correlaciones. Sin embargo, esta tendencia no se observa en todas las poblaciones indígenas. Los contraejemplos en nuestra muestra son los hablantes del zapoteco de Juchitán y del Sumu-Mayangna, que tienen altos niveles de educación formal.

Figura 13. *Nivel de educación por grupo de lengua nativa.*

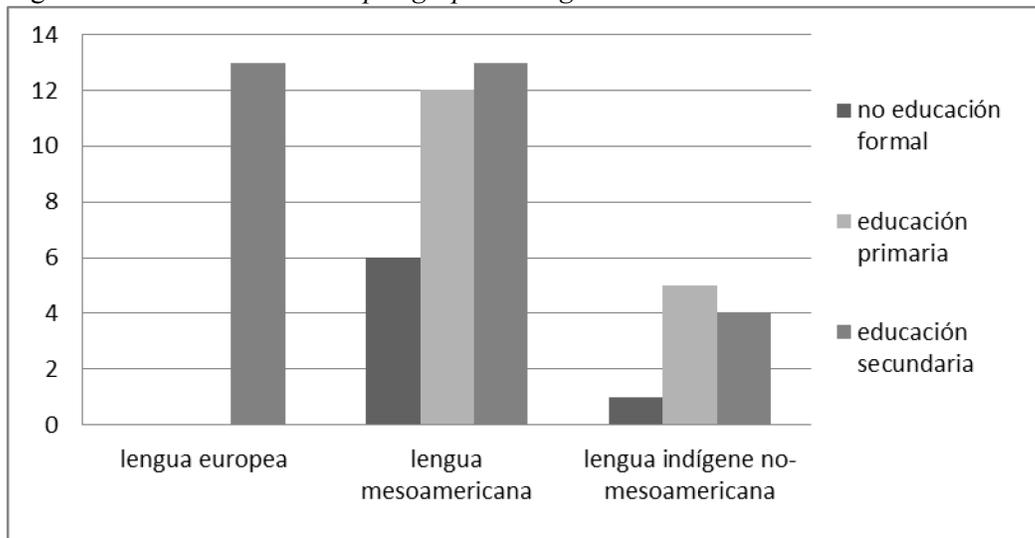
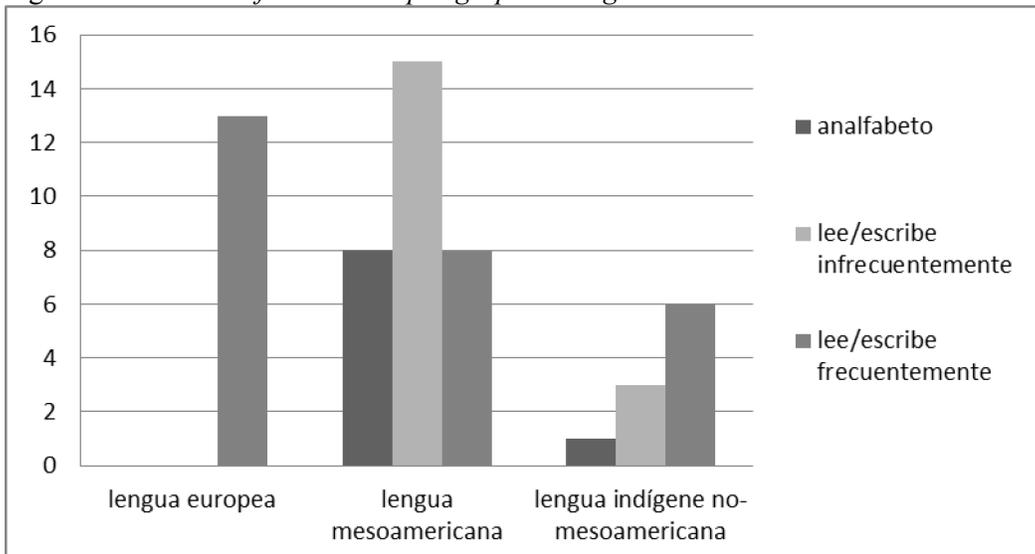


Figura 14. Nivel de alfabetización por grupo de lengua nativa.



Una manera de probar cuál de los cuatro factores es el más fuerte, es calcular para cada variable la distancia media entre los participantes que tienen el mismo valor. Esto nos da una medida bruta de la densidad de las agrupaciones. La idea es que las variables con las agrupaciones más densas son las que más predicen el comportamiento de los participantes. La Tabla 3 muestra los resultados de este cálculo. De este cálculo surge que la agrupación más densa se encuentra entre los hablantes nativos del español. En segundo lugar, los participantes que no hablan español en absoluto también se comportan muy similar en términos del uso de MdRs. En tercer lugar vemos los hablantes de lenguas indígenas no-mesoamericanas y sólo en cuarto lugar los participantes que carecen de educación formal. Aunque de carácter preliminar, esto nos parece sugestivo de un papel mayor para los factores lingüísticos.

Tabla 3. *Distancias medias para cada valor en la matriz de semejanzas.*

variable	valor	distancia media
edad	< 30	35.8
	≥ 30	39.2
sexo	masculino	35.0
	femenino	40.4
nivel de educación	no educación formal	28.9
	educación primaria	31.0
	educación secundaria	39.3
nivel de alfabetización	analfabeto	30.7
	lectura/escritura infrecuente	34.5
	lectura/escritura diaria	36.6
uso del español (como L1 o L2)	ausente	24.7
	uso infrecuente	35.8
	uso frecuente/diario	36.3
lengua nativa	mesoamericana	34.8
	indígena no-mesoamericana	26.4
	español	23.5

La evidencia que hemos presentado en esta sección sugiere que tanto los factores lingüísticos como la educación y la alfabetización influyen el uso de MdRs. Sin embargo, estos factores están interrelacionados y nuestro análisis multivariante sugiere que los factores lingüísticos son los más poderosos, respaldando a la posición neo-whorfiana. Entonces, aunque de manera preliminar, sí podemos ver que hay indicadores fuertes de que el lenguaje, tal como lo definimos anteriormente, es uno de los factores más significativos en la transferencia y la difusión de prácticas cognitivas culturales.

5. Direcciones futuras

Si bien los resultados que hemos presentado aquí son interesantes, hay que recordar que son, de momento, preliminares y que el trabajo continúa. Algunos miembros del equipo todavía se hallan en proceso de análisis y codificación de datos, tanto lingüísticos como no lingüísticos. Al mismo tiempo, el equipo de MesoSpace ha iniciado la codificación de los datos obtenidos en términos de información topológica y de urbanización, para poder diseñar un análisis de los posibles factores extra-lingüísticos que puedan ser relevantes. En este sentido, también estamos considerando las características de los modelos educativos existentes, tanto monolingües como bilingües, en los diferentes contextos culturales, y su posible influencia en la transmisión, mantenimiento o abandono de prácticas lingüísticas. Asimismo, el equipo de MesoSpace ha iniciado el proceso de diseñar un modelo multivariante que pueda realizar análisis inferenciales sobre los datos presentados aquí.

Por otro lado, y como resultado de la expansión del proyecto MesoSpace, se han empezado a recoger nuevos datos lingüísticos de lenguas de zonas geográficas tan dispares como Asia, África y América del Sur.

Y, finalmente, estamos preparándonos para implementar una nueva serie de estímulos y tareas tanto lingüísticas como no-lingüísticas que puedan iluminar la relación entre lenguaje y cultura que ha salido a la luz en este trabajo.

6. Referencias bibliográficas

- Bickel, Balthasar. (2010). Capturing particulars and universals in clause linkage: a multivariate analysis. En Isabelle Bril (Ed.), *Clause-hierarchy and clause-linking: the syntax and pragmatics interface*, Amsterdam: John Benjamins, pp. 51–101.
- Bohnemeyer, Jürgen & Stephen C. Levinson. (ms). Framing Whorf: A response to Li et al. 2011. *Cognition*.
- Bohnemeyer, Jürgen & Carolyn O'Meara. (en prensa). Vectors and frames of reference: Evidence from Seri and Yucatec. En Luna Filipović & Kasia M. Jaszczolt (Eds.), *Space and Time across Languages and Cultures*, Amsterdam: John Benjamins.
- Bryant, David & Vincent Moulton. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2): 255–265.
- Carlson-Radvansky, Laura A. & David E. Irwin. (1993). Frames of reference in vision and language: Where is above? *Cognition* 46: 223–244.
- Cysouw, Michael. (2007). New approaches to cluster analysis of typological indices. En Reinhard Köhler & Peter Grzбек (Eds.), *Exact Methods in the Study of Language and Text*, Berlin: Mouton de Gruyter, pp. 61–76.
- Danziger, Eve. (2010). Deixis, gesture, and cognition in spatial Frame of Reference typology. *Studies in Language* 34(1): 167–185.
- Haun, Daniel B., Christine Rapold, Gabriele Janzen, & Stephen C. Levinson. (2011). Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition* 119: 70-80.
- Haviland, John B. (1993). Anchoring, iconicity, and orientation in Guugu Yimithirr pointing gestures. *Journal of Linguistic Anthropology* 3(1): 3–45.
- Huson, Daniel H. & David Bryant. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2): 254–267.
- Jackendoff, Ray. (1996). *Languages of the Mind: Essays on Mental Representation*, Cambridge: MIT Press.
- Levinson, Stephen C. (1996). Frames of reference and Molyneux's Question: Crosslinguistic evidence. En Paul Bloom, Mary A. Peterson, Lynn Nadel, & Merrill F. Garrett (Eds.), *Language and space*, Cambridge, MA: MIT Press, pp. 109–169.
- Levinson, Stephen C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: Language affects spatial reasoning. *Cognition* 84(2), 155-188.
- Levinson, Stephen C. (2003). *Space in language and cognition*. Cambridge, UK: Cambridge University Press.
- Levinson, Stephen C. & 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.
- Li, Peggy & Lila Gleitman. (2002). Turning the tables: Language and spatial reasoning. *Cognition* 83: 265-294.

- Li, Peggy, Linda Abarbanell, Lila Gleitman, & Anna Papafragou. (2011). Spatial reasoning in Tenejapan Mayans. *Cognition* 120: 33–53.
- Majid, Asifa, Melissa Bowerman, Sotaro Kita, Daniel B. Haun, & Stephen C. Levinson. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3): 108–114.
- Majid, Asifa, James S. Boster, & Melissa Bowerman. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition* 109(2): 235–250.
- Mishra, R.C., Pierre R. Dasen & Shanta Niraula. (2003). Ecology, language, and performance on spatial cognitive tasks. *International Journal of Psychology* 38: 366–383.
- O’Meara, Carolyn & Gabriela Pérez Báez. (2011). Spatial frames of reference in Mesoamerican languages. *Language Sciences* 33: 837–852.
- Pederson, Eric. (1993). Geographic and manipulable space in two Tamil linguistic systems. En A. U. Frank & I. Campari (Eds.), *Spatial information theory*, Berlin: Springer, pp. 294–311.
- Pederson, Eric, Eve Danziger, David P. Wilkins, Stephen C. Levinson, Kita, S. & Gunter Senft. (1998). Semantic typology and spatial conceptualization. *Language* 74: 557–589.
- R Development Core Team. (2011). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- Regier, Terry & Paul Kay. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences* 13: 439–446.
- Talmy, Leonard. (2000). *Toward a Cognitive Semantics: Concept Structuring Systems*, vol. 2. Typology and Process in Concept Structuring. MIT Press, Cambridge.
- Terrill, Angela & Niclas Burenhult. (2008). Orientation as a strategy of spatial reference. *Studies in Language* 32(1): 93–116.
- Schiffman, Susan S., M. Lance Reynolds, & Forrest W. Young. (1981). *Introduction to multidimensional scaling: Theory, methods and applications*, New York: Academic Press.
- Wassmann, Jürg & Pierre R. Dasen. (1998). Balinese spatial orientation: Some empirical evidence for moderate linguistic relativity. *The Journal of the Royal Anthropological Institute* 4(1): 689–711.

7. Nota final

Jürgen Bohnemeyer
 Linguistics Department
 University at Buffalo
 609 Baldy Hall
 Buffalo, NY 14260
 USA
<http://www.acsu.buffalo.edu/~jb77/>