

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA
ESCUELA DE DOCTORADO

PROGRAMA DE DOCTORADO EN FILOLOGÍA:
ESTUDIOS LINGÜÍSTICOS Y LITERARIOS.



TESIS DOCTORAL

ANÁLISIS DEMOLINGÜÍSTICO DEL LÉXICO HISPÁNICO:
ESTUDIO APLICADO A LAS “NOCIONES ESPECÍFICAS” DEL *PLAN*
CURRICULAR DEL INSTITUTO CERVANTES

CARMEN CHACÓN GARCÍA

Licenciada en Filología Hispánica

DIRECTORA:

Dra. D.^a M.^a Antonieta Andión Herrero

2016

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA
ESCUELA DE DOCTORADO

PROGRAMA DE DOCTORADO EN FILOLOGÍA:
ESTUDIOS LINGÜÍSTICOS Y LITERARIOS.



TESIS DOCTORAL

ANÁLISIS DEMOLINGÜÍSTICO DEL LÉXICO HISPÁNICO:
ESTUDIO APLICADO A LAS “NOCIONES ESPECÍFICAS” DEL *PLAN*
CURRICULAR DEL INSTITUTO CERVANTES

CARMEN CHACÓN GARCÍA

Licenciada en Filología Hispánica

DIRECTORA:

Dra. D.^a M.^a Antonieta Andión Herrero

2016

AGRADECIMIENTOS

Esta tesis no habría sido posible sin la ayuda de M.^a Antonieta Andión Herrero. Cuando aún trabajaba como profesora en París y decidí buscar una universidad en Madrid para hacer mi tesis doctoral, la persona que me abrió las puertas y trabajó conmigo al otro lado del teléfono, en horario de fin de semana y sin ni siquiera conocerme, fue Marieta. De ella he aprendido la importancia del compromiso con la investigación, con la institución universitaria y, sobre todo, con los estudiantes. Marieta ha sido un pilar constante para el desarrollo de este trabajo. Con paciencia, ha contestado mis dudas y ha leído con dedicación y cariño cada una de estas páginas, así como los múltiples trabajos elaborados en estos años.

Durante más de cuatro años, he tenido la suerte de trabajar en el departamento de Lengua Española y Lingüística General de la UNED, gracias a una beca predoctoral para investigadores en formación (FPI-UNED). Agradezco a todo el departamento su acompañamiento durante estos años. En especial, me gustaría destacar el asesoramiento, el apoyo y el buen humor que me han ido brindando a lo largo de este tiempo José Amenós, Victoria Marrero, Angelita Estévez, José Ramón Carriazo y, especialmente, Alicia San Mateo, a la que le agradezco, además, el apoyo en la recta final y sus consejos sobre la edición de estas páginas. Tampoco puedo olvidar la calurosa acogida y las múltiples atenciones de Virginia Ovejero que, con deliciosas tazas de té, me ha hecho sentirme como en casa muchas mañanas.

Gracias a Álvaro García Santa-Cecilia por darme a conocer el *Plan curricular*, admitiendo mi colaboración en el Instituto Cervantes cuando todavía era estudiante en la universidad. Agradezco también a Hiroto Ueda y a Antonio Ruiz Tinoco su interés y disposición para este proyecto, dándome acceso a la base de datos de Varilex.

En estos años he realizado dos estancias de investigación: una, en Valencia; y otra, en París. En primer lugar, agradezco a Marta Albelda su apoyo y acogida en la Universitat de Valencia. Desde antes de empezar esta tesis, y también sin conocerme personalmente, Marta ya me había dedicado su tiempo para solicitar proyectos de investigación. En Valencia me contagió el entusiasmo por los corpus lingüísticos y me ofreció, además de su buen humor, pistas esenciales sobre el camino de la investigación.

La estancia en París no habría sido posible sin Carlota Piedehierro, con quien he ido compartiendo, azarosamente, espacios y momentos profesionales desde hace una década. Gracias por allanarme tantos caminos, también en la universidad *Sorbonne-Nouvelle*, donde pude conocer al apasionado hispanista, Eric Beaumatin. Agradezco también a Carmen Ballesteros su profesionalidad y acogida.

Echando la vista atrás, el aprendizaje más importante de esta tesis es la constatación de las inmensas posibilidades que abre la cooperación interdisciplinaria entre campos científicos diversos. En este sentido, me gustaría destacar la importancia que ha tenido en esta tesis la implicación de Ana y Pablo Chacón y de Oscar Perpiñán. Ana consiguió que superara los momentos de bloqueo de esta investigación, poniendo orden y rigor para plantear la metodología más adecuada. Pablo ha dedicado largos días a reflexionar sobre planteamientos de esta tesis, enseñándome con paciencia el funcionamiento de las bases de datos y obligándome a buscar siempre las soluciones «más inteligentes». El nivel de excelencia de Oscar me ha hecho constatar que los mejores profesionales son aquellos que, además de destacar por sus capacidades, se comprometen con objetivos que van más allá de sus propias necesidades, sin esperar nada a cambio. Gracias por descubrirme el universo «R» y por tu generosidad.

Quiero agradecer a mis amigos de la carrera, especialmente a Carlos y Laura, su apoyo constante durante estos años, la paciencia que han tenido para escuchar mis proyectos, la cantidad de veces que han venido a casa a cenar «para poder verme», las discusiones filológicas (y políticas). Mi cabeza ha aprendido a funcionar de otra manera desde que os conozco. Gracias también a María, que a pesar de vivir en Nueva York fue la primera que visualizó los mapas de esta tesis y me asesoró desde su óptica periodística; y a Diego, que desde Londres me dio las primeras pistas sobre aplicaciones dinámicas interactivas. Gracias a Lucía, Manu, Luismi, Alfonso y Cris. Tampoco me olvido de los momentos compartidos con Iván y Marga, que se han ocupado de que nuestra vida sea más fácil y, sobre todo, más feliz. Gracias a Cristina, por seguir más allá del tiempo. Gracias también a Guillermo y Marta, que comprenden lo que supone haber llegado hasta aquí. A Alberto y Javi, por sus risas; a Juanan y Ana, por regalarnos su casa; a Javi y Carlos, que se van pero siempre vuelven; a Andrea y David, que desde Montevideo me regalan nuevas palabras sobre las que investigar. Gracias a Alba y a Jose, siempre pendientes de que no descuidara la tesis; a Tomás y Belén, por querernos tanto. Gracias a mis amigos del Nido, especialmente a Vanessa, por los largos días de biblioteca.

También tengo muy presentes a mis amigos de París, a los que tanto extraño desde de que inicié esta investigación, y que me siguen acompañando y visitando, como si no hubiera dos mil kilómetros de distancia. Gracias especialmente, a Elena y Jean Pierre; a Javier, mi hermano parisino; a Carlota, a Bea y a Sonia; y a mi querida Judith.

Para Gonzalo no basta un «gracias». Aunque te has llevado la peor parte de este esfuerzo, siempre me has recordado la importancia de no traicionar los propios sueños. Gracias por ser mi espejo: para verme, solo tenía que mirarte. Y gracias porque en cuatro años de tesis todo se ha multiplicado menos las horas de sueño con Miguel y Marcos.

Por último, quiero dedicar esta tesis a toda mi familia, que siempre me ha acompañado y apoyado en esta andadura. Mis hermanos, mis sobrinos, mis suegros, mis primos, mis tíos. Mención aparte se merecen mis padres: sin ellos, sencillamente, no habría podido dar forma a estas páginas, ni sería lo que soy. Gracias a los dos por vuestra paciencia: no hay palabras que expresen el alcance de mi gratitud.

*Todo está en la palabra... Una idea entera se cambia
porque una palabra se trasladó de sitio, o porque otra se sentó
como una reinita adentro de una frase que no la esperaba y que le
obedeció... Tienen sombra, transparencia, peso, plumas, pelos,
tienen de todo lo que se les fue agregando de tanto rodar por el río,
de tanto transmigrar de patria, de tanto ser raíces.*

Pablo Neruda

ÍNDICE

INTRODUCCIÓN	17
CAPÍTULO 1. EL PCIC Y LA VARIACIÓN LÉXICA	25
1. El Plan Curricular del Instituto Cervantes: descripción general	25
1.1. Introducción	25
1.2. Antecedentes y bases del PCIC	27
1.3. Características generales del PCIC	30
1.4. Estructura y organización	31
1.5. Utilidad e importancia del PCIC	32
2. El PCIC: variedad lingüística y variación léxica	33
2.1. Unidad y diversidad del español	33
2.2. La variedad lingüística en la enseñanza de español	37
2.3. Norma lingüística del PCIC	41
2.4. Las <i>Nociones específicas</i> del PCIC y la variación léxica	44
2.5. Didáctica de la variedad léxica	49
3. Conclusión	54
CAPÍTULO 2. GEOGRAFÍA LINGÜÍSTICA Y LÉXICO HISPÁNICO	55
1. Dialectología y geografía lingüística	54
1.2. Dialectología: perspectiva histórica	56
1.3. Geografía lingüística	62
1.3.1. Geografía lingüística en Europa	63
1.3.2. Geografía lingüística en Hispanoamérica	67
1.3.3. La metodología de la geografía lingüística	69
1.3.4. Críticas y limitaciones de la geografía lingüística	71
2. El estudio de la variación léxica	72
2.1. Fuentes de consulta y documentación	75
2.1.1. Atlas lingüísticos	76
2.1.2. Diccionarios	77
2.1.3. Estudios especializados y bases de datos	82

2.1.4. Lingüística de corpus	86
2.2. Descripción general de la variación léxica	89
3. Conclusión	91

CAPÍTULO 3. PROYECTO INTERNACIONAL DE VARIACIÓN LÉXICA DEL ESPAÑOL EN EL MUNDO: VARILEX 93

1. Introducción	93
2. Origen de Varilex	94
3. Descripción general del proyecto	96
3.1. Fase preliminar e inicio del proyecto	97
3.2. Relevancia y aplicaciones del proyecto	99
3.3. Metodología	102
4. La base de datos Varilex como recurso de investigación	115
4.1. Características	115
4.2. Estructura	116
4.3. Métodos de búsqueda	118
4.4. Cartografía automática	118
4. Puntos fuertes y limitaciones de Varilex	119
5. Conclusión	124

CAPÍTULO 4. PROPUESTA DE ANÁLISIS DEMOLINGÜÍSTICO DE LA VARIACIÓN LÉXICA 127

1. Introducción	127
2. Origen de la investigación	128
3. Hipótesis y objetivos de investigación	129
4. Metodología	131
4.1. Tipo de estudio	131
4.2. Cálculo demolingüístico	133
4.3. Materiales de estudio	139
4.3.1. El Plan Curricular del Instituto Cervantes	139
4.3.2. El proyecto Varilex	139
4.3.3. Diccionario de americanismos	140
4.3.4. Otras obras consultadas	140
4.4. Datos analizados	142

4.4.1. Conceptos y variantes	142
4.4.2. Países encuestados y población	142
4.4.3. Informantes: edad y sexo	144
4.4.4. Encuestas	145
4.4.5. Identificación de variantes léxicas	146
4.5. Distribución en áreas dialectales	147
5. Limitaciones de esta investigación	148
6. Conclusiones	150
CAPÍTULO 5. RESULTADOS DEMOLINGÜÍSTICOS	153
1. Introducción	153
2. Presentación de resultados	154
3. Aplicación virtual para la representación cartográfica de los datos	156
4. Mapas estáticos de la representatividad léxica	161
5. Análisis de resultados	164
5.1. Sustitución de formas léxicas	165
5.2. Americanismos	166
5.3. Españolismos	168
5.4. Panhispanismos	171
5.5. La vitalidad de las voces	173
5.6. La convivencia de varias voces	175
5.7. La dispersión de las voces	177
6. Conclusión	180
CAPÍTULO 6. PROPUESTA DE REVISIÓN Y AMPLIACIÓN DE LAS NOCIONES ESPECÍFICAS DEL PCIC	181
1. Introducción	181
2. Cuestiones previas sobre las equivalencias léxicas	183
3. Acercamiento a la variedad léxica	188
4. Descripción general de nuestra propuesta de revisión y ampliación	189
4.1. Voces panhispánicas (Grupo I)	192
4.1.1. Propuesta específica de panhispanismos	193
4.1.2. Otros panhispanismos sujetos a valoración	197

4.1.2.1. Palabras panhispánicas ligeramente más representativas que descartamos para la ampliación	198
4.1.2.2. Palabras panhispánicas ligeramente más representativas propuestas para la ampliación	199
4.1.3. Voces panhispánicas excluidas	200
4.2. Voces americanas con mayor GR que los términos de referencia del PCIC (Grupo II.1)	200
4.2.1. Propuesta específica de americanismos	202
4.2.2. Otros americanismos sujetos a valoración	208
4.3. Voces americanas secundarias, con menor GR pero con alcance y extensión significativa (Grupo II.2)	212
4.3.1. Propuesta específica de voces americanas secundarias	213
4.3.2. Americanismos descartados para la ampliación	223
4.4. Casos particulares de ampliación	226
5. Marca [Esp]: voces preferentes del español de España o <i>españolismos</i>	227
5.1. Propuesta de la marca [Esp]	228
5.1.1. Españolismos puros	228
5.1.2. Españolismos de preferencia	232
5.2. Otros posibles españolismos sujetos a valoración	234
6. Conclusión	237
CONCLUSIONES	239
BIBLIOGRAFÍA	249
ANEXOS	269
I. Fuentes de consulta sobre la variación léxica americana	271
II. Conceptos de la base de datos de Varilex estudiados en esta investigación	275
III. Mapas estáticos de los conceptos estudiados	287

ÍNDICE DE FIGURAS

Figura 1: Organización y estructura del PCIC	107
Figura 2: La variedad lingüística en el PCIC	109
Figura 3: Propuesta de revisión y ampliación del PCIC	109
Figura 4: El concepto [COLD] según el Visual Theasurus	116
Figura 5: Niveles de colaboración de Varilex.....	102
Figura 6: Pregunta [A013] de Varilex en el formato de encuesta tradicional	103
Figura 7: Pregunta [A011] de Varilex en formato de encuesta web.....	103
Figura 8: Tabla interrelacionadas de la Base de Datos de Varilex	110
Figura 9: Distribución por edad de los informantes de Varilex.....	144
Figura 10: Distribución por sexo de los informantes de Varilex	144
Figura 11: Ejemplo de pregunta con ilustración de la encuesta de Varilex.....	145
Figura 12: Ejemplo de pregunta apoyada en ejemplos de la encuesta de Varilex	145
Figura 13: Cómo seleccionar un concepto de estudio en la aplicación virtual.....	158
Figura 14: Representación de las variantes más representativas del concepto.....	159
Figura 15: Resultados del GR de una variante léxica en distintos países.....	159
Figura 16: Consulta del mapa del concepto [COMPUTER] con la Tabla de Resultados	160
Figura 17: Propuesta de revisión y ampliación de las «Nociones específicas» del PCIC	190

ÍNDICE DE TABLAS

Tabla 1 Ejemplos de integración de la variedad en los inventarios del PCIC	42
Tabla 2 Diferencias léxicas entre España e Hispanoamérica.....	90
Tabla 3 Ciudades y países donde Varilex ha estudiado la variedad léxica.....	112
Tabla 4 Encuestas de Varilex con los correspondientes campos semánticos estudiados.....	120
Tabla 5 Distribución del concepto JACKET según el análisis demolingüístico de Ueda (1996)	134
Tabla 6 Diferencias del análisis demolingüístico de [BLACKBOARD] con GR general y GR ponderado ordenados de mayor a menor.....	136
Tabla 7 Diferencias del análisis demolingüístico de [BEER] con GR general y GR ponderado ordenados de mayor a menor.....	137
Tabla 8 Países y ciudades encuestadas con su población y abreviatura correspondiente	143
Tabla 9 Áreas dialectales analizadas.....	148
Tabla 10 Propuesta de ampliación de variantes panhispánicas en el PCIC.....	194
Tabla 11 Otras nociones panhispánicas, con GR ligeramente superior a las del PCIC.....	198
Tabla 12 Oposición <i>maletín – portafolios</i>	199
Tabla 13 Propuesta de ampliación de americanismos en el PCIC.....	204
Tabla 14 Valoración de otros americanismos con GR ligeramente superior.....	210
Tabla 15 Variantes americanas secundarias	215
Tabla 16 Distribución de la voz bocadito en el Diccionario de Americanismos y en Varilex	221
Tabla 17 Americanismos descartados para la ampliación	224
Tabla 18 Casos particulares de ampliación.....	226
Tabla 19 Españolismos puros del PCIC.....	230
Tabla 20 Distribución de españolismos de preferencia	233
Tabla 21 Otros españolismos según fuentes secundarias.	235

ÍNDICE DE MAPAS

Mapa 1 Concepto [BRASSIERE]	162
Mapa 2 Concepto [CELLOPHANE TAPE]	163
Mapa 3 Concepto [STAPLER]. Ejemplo de sustitución de formas léxicas	165
Mapa 4 Concepto [OVERALLS]. Ejemplo de americanismo	167
Mapa 5 Concepto [SWIMSUIT]. Ejemplo de españolismo	169
Mapa 6 Concepto [REFRIGERATOR]. Ejemplo de españolismo	170
Mapa 7 Concepto [BOSS]. Ejemplo de panhispanismo	172
Mapa 8 Concepto [TIDBITS]. Ejemplo de voz local con gran vitalidad en un territorio	174
Mapa 9 Concepto [EARRINGS]. Ejemplo de convivencia de varias formas léxicas	176
Mapa 10 Concepto [FOLDER]. Ejemplo de dispersión en la geografía hispánica	178
Mapa 11 Concepto [FLAT]. Ejemplo de dispersión en la geografía hispánica	179

INTRODUCCIÓN

Nuestra tesis, *Análisis demolingüístico del léxico hispánico: Estudio aplicado a las «Nociones específicas» del Plan Curricular del Instituto Cervantes*, nace de la conciencia de la dificultad que experimentan los profesores de español como segunda lengua o como lengua extranjera para determinar el léxico que deben incorporar en sus clases. La enorme extensión del mundo hispánico hace que la selección de un vocabulario representativo, general y, sobre todo, rentable, para los aprendices de español sea una tarea compleja.

El objetivo de nuestra tesis doctoral es, por un lado, definir el grado de representatividad de las «Nociones específicas» que propone el Instituto Cervantes para la enseñanza de español. Por otro lado, queremos ofrecer una propuesta de ampliación, fundamentada y rigurosa, que dé cuenta de las distintas alternativas léxicas válidas para el conjunto de la geografía hispánica.

Se trata, por tanto, de una tesis sobre la variación léxica del español actual, aunque abundan en ella diversas consideraciones sobre la geografía lingüística, el procesamiento estadístico de los datos, la representación cartográfica de las variantes o los mecanismos de incorporación léxica. Así, nuestra aportación pretende actualizar las nociones propuestas por la obra de referencia para la enseñanza de español, el *Plan Curricular del Instituto Cervantes* (en adelante, *PCIC*), incorporando variantes representativas de otras variedades de la lengua española.

Desde la publicación del *Marco Común Europeo de Referencia* (en adelante, *MCER*), se aboga por la necesidad de introducir las variedades lingüísticas en la enseñanza. Por su parte, el *PCIC* reconoce la validez de varias normas cultas del español e incorpora anotaciones y especificaciones de las particularidades lingüísticas que coinciden con amplias zonas dialectales del mundo hispánico. Para seleccionar el material lingüístico, el Instituto Cervantes toma preferentemente la norma centro–norte peninsular española. Esta elección se sustenta en la proyección de esta variedad como modelo estandarizado prestigioso de la lengua así como en la cantidad de rasgos comunes que comparte con las restantes normas cultas del español.

Las especificaciones del *PCIC* relativas a otras variedades de español son necesarias, sobre todo, cuando las realizaciones lingüísticas del español peninsular difieren de la norma culta de otras variedades geolectales. Estas realizaciones no coincidentes deben contar con un área de extensión y uso suficientemente amplia como para que sean significativas en la enseñanza de español como segunda lengua o lengua extranjera. Para el docente de español, que pertenece inevitablemente a una variedad concreta, la existencia de un documento lingüístico que especifique los rasgos dialectales de otras variedades tiene una importancia fundamental.

Si nos centramos en el plano léxico, la descripción y sistematización del vocabulario del español es un asunto particularmente complicado por el carácter variable de este componente, así como por la falta de exhaustividad de las investigaciones lexicográficas.

El problema que encuentra el profesor de español para tratar la variación léxica es doble: por un lado, las variaciones no responden a reglas categóricas sino a diferencias de frecuencia o uso difíciles de sistematizar; por otro, apenas existen materiales de referencia que le permitan integrar la variedad léxica en el currículo de manera significativa. Este vacío bibliográfico exige proponer inventarios que presenten de forma sistemática la distribución de las variantes de español por campos y niveles de dominio.

En lo que se refiere a la descripción lingüística del *PCIC*, aunque en otras secciones del documento reconocemos anotaciones sobre rasgos dialectales diversos, como por ejemplo en el inventario de «Gramática», el repertorio léxico que propone el Instituto Cervantes presenta exclusivamente el español peninsular, acusando la ausencia de voces relativas a otras variedades de español. Así, nuestra investigación aporta una propuesta fundamentada para integrar en este inventario otras variedades geolectales altamente representativas y responde a un asunto prioritario de las investigaciones lingüísticas relacionadas con la didáctica del léxico del español.

Los datos que utilizaremos para nuestro análisis de la representatividad proceden de las investigaciones del grupo Varilex, que ha estudiado la variación léxica del español en el mundo desde los años 90. Para ello, partimos de una base de datos relacional, elaborada expresamente para nuestra tesis, que nos permite analizar la distribución de las variantes léxicas registradas en el dominio geográfico del español mediante cálculos demolingüísticos que aportan el Grado de Representatividad (GR) de las variantes léxicas de cada concepto. En este sentido, nuestro trabajo se relaciona con la geografía lingüística, pues analiza e interpreta de manera exhaustiva la variedad léxica de nuestra lengua en el conjunto del territorio hispánico.

A pesar de que los estudios lingüísticos sobre léxico se centran tradicionalmente en exponer las diferencias de vocabulario, nuestro análisis parte de la hipótesis de que existen variantes léxicas que son comunes a una mayoría de hispanohablantes. Estas semejanzas suponen el acercamiento a una *koiné* hispánica que desdibuja las particularidades y resulta imprescindible en la enseñanza de español, puesto que garantiza los criterios de extensión, frecuencia y rentabilidad.

Esta investigación pretende integrar las variantes léxicas más representativas del mundo hispánico en el *PCIC*. Nuestro estudio de la variación se imbrica, por tanto, en el ámbito de la lingüística aplicada, concretamente, en la enseñanza de Español como Lengua Extranjera o Segunda Lengua (LE/L2). El análisis demolingüístico que presentamos en este trabajo tiene una aplicación directa y

relevante para la enseñanza de español y aporta una valiosa herramienta para estudiar y sistematizar la variación léxica, según criterios de rentabilidad, disponibilidad y frecuencia nociofuncional de las palabras.

Seleccionar aquellas variantes que son verdaderamente representativas supone que, además de incorporar voces panhispánicas comunes a la geografía hispánica, también propongamos una serie de voces americanas que, aunque no existen en la variedad peninsular, cuentan con un área de validez y uso suficientemente amplia que las hace fundamentales para la enseñanza de español.

Por otra parte, aunque ofrecemos algunas reflexiones sobre la enseñanza y el aprendizaje del vocabulario, nuestra tesis no constituye un tratado didáctico sobre cómo enseñar el léxico en clase de ELE. Tampoco pretendemos valorar qué variedad de español debe presentar el profesor en el aula, puesto que cualquier variedad de español en su registro culto oral y escrito podría ser la variedad preferente de un curso de español como segunda lengua o lengua extranjera (Andión, 2003).

En definitiva, la presente investigación ofrece resultados empíricos sobre la distribución de las voces a lo largo del territorio hispánico que contribuyen a desarrollar una enseñanza rigurosa del vocabulario, según criterios sólidos y fuentes de consulta contrastadas sobre las variantes léxicas del español.

En el primer capítulo de esta tesis describimos la obra sobre la que articulamos nuestra investigación, el *Plan Curricular del Instituto Cervantes* (Instituto Cervantes, 2006). Pasada ya una década desde su publicación, consideramos que es conveniente revisar el componente léxico de este material. Comenzamos nuestra exposición con una descripción general de la obra, para comprender su sentido e importancia en la enseñanza de español como segunda lengua o lengua extranjera. El *PCIC* es una obra esencial puesto que fija y desarrolla los componentes de aprendizaje que necesita el estudiante de español para lograr objetivos comunicativos en los diferentes niveles de dominio. También nos ocuparemos de sus antecedentes, del proceso de elaboración, de las utilidades y de los destinatarios de este material. Se trata de presentar las características generales de

la obra, así como su organización y estructura. En particular, trataremos el tema de la norma lingüística y la integración de las variedades dialectales en este documento, para centrarnos después en el componente léxico que propone el inventario «Nociones específicas». Consideramos que la presentación y descripción del *PCIC* que abre nuestra tesis es un paso previo necesario para ocuparnos después del análisis de la representatividad de las nociones que presenta el documento.

Los capítulos 2 y 3 nos permiten introducir el marco de estudio de esta tesis y constituyen la base teórica de nuestro análisis geolectal de la variación. En el capítulo 2, ofrecemos una revisión crítica sobre de las principales cuestiones que se han formulado desde la dialectología en las últimas décadas, centrándonos particularmente en la geografía lingüística. Analizaremos también los métodos empleados para el estudio del léxico en España y América, exponiendo sus ventajas e inconvenientes. El marco teórico ofrece, además, información general sobre las fuentes de consulta disponibles en la actualidad para el estudio de la variación léxica hispánica, con sus características específicas.

Una vez concluida nuestra revisión crítica sobre las investigaciones léxicas, el capítulo 3 de la tesis se centra específicamente en una de ellas: el proyecto Varilex (*Variación Léxica del español en el mundo*). Los datos de Varilex son la base de nuestro trabajo. Sin ellos, no habría sido posible plantear nuestro estudio demolingüístico de la variación. En este capítulo analizamos críticamente las características generales, el procesamiento, registro y análisis de los datos, así como la fundamentación metodológica de esta investigación. También presentamos otras fuentes de referencia sobre variación léxica que se tienen en cuenta en nuestro análisis de la variación. Este capítulo es esencial para entender de dónde obtenemos los datos de nuestro estudio.

La parte central de la tesis está constituida por los capítulos 4 y 5, donde presentamos formalmente nuestra investigación y ofrecemos un análisis de la representatividad de las variantes léxicas. El capítulo 4 comienza explicando el origen y justificación de nuestro trabajo. Exponemos a continuación las hipótesis de

investigación que orientan nuestro estudio así como los objetivos generales y específicos que pretendemos alcanzar. En estas páginas explicamos la metodología con la que medimos el concepto de representatividad. Presentamos, así, nuestro estudio descriptivo de las voces, que está basado en cálculos demolingüísticos y pretende reflejar el panorama global sincrónico de la variación léxica del español. También definimos las actividades de investigación así como los materiales esenciales con los que hemos trabajado. Por último, explicamos en detalle los conceptos y variantes estudiadas así como los países, los informantes, las encuestas, los criterios de identificación de las voces y las áreas dialectales que tendremos en cuenta para valorar la incorporación de nociones en el *PCIC*. El capítulo cierra con las limitaciones generales de nuestro trabajo.

El capítulo 5 tiene como objetivo exponer los resultados demolingüísticos del análisis de 182 nociones del *PCIC*. Para ello, hemos diseñado una aplicación virtual dinámica que muestra la distribución léxica de cada concepto a través de mapas e incluye una tabla con los resultados del análisis de la representatividad. Siguiendo la tradición de los atlas lingüísticos, consideramos que la representación cartográfica aporta una visión panorámica de la distribución de las variantes léxicas, que permite estudiar e interpretar los resultados demolingüísticos y extraer conclusiones a partir de una imagen global de la lengua. Los mapas de esta aplicación, además de mostrar nuestros resultados, podrían ser útiles para desarrollar hipótesis o investigaciones ulteriores sobre las inferencias causales, el desarrollo histórico de los fenómenos de variación u otras cuestiones relacionadas con la distribución léxica. En todo caso, nuestro objetivo se centra en valorar la representatividad de las nociones del *PCIC* y de sus correspondientes equivalencias. En este capítulo, presentamos el sistema de representación cartográfica de datos que hemos utilizado en nuestra investigación, gracias al entorno R y el paquete Shiny, que nos permite una visualización dinámica, fácil y práctica para el usuario. También presentamos nuestro diseño de mapas estáticos para mostrar los resultados en formato papel. Por último, incluimos una síntesis de las aportaciones más significativas de nuestros mapas lingüísticos, de los que se desprenden informaciones relacionadas con la distribución léxica, la

existencia de panhispanismos, americanismos y españolismos, así como la vitalidad, convivencia y dispersión de las voces.

La última parte de la tesis corresponde al capítulo 6 y constituye la culminación de nuestro estudio. En este apartado presentamos nuestra propuesta de ampliación de las «Nociones específicas» del *PCIC* con voces representativas del mundo hispánico. Basándonos en los conceptos teóricos de los anteriores capítulos y teniendo en cuenta nuestros resultados demolingüísticos, así como la información de otras fuentes como el *Diccionario de americanismos* o las consultas del CORPES, configuramos una propuesta rigurosa de ampliación del repertorio léxico del *PCIC*. Esta revisión y ampliación de las «Nociones específicas» se presenta en tres direcciones: en primer lugar, proponemos variantes léxicas panhispánicas, que son comunes a España y América y pueden considerarse representativas del español global actual; por otro lado, incluimos variantes americanas que cuentan con un alto grado de representatividad en la geografía hispánica, aunque no sean propias del español peninsular; por último, identificamos aquellas palabras del *PCIC* que existen únicamente en España e introducimos la marca de españolismo, con el fin de hacer de este material un documento riguroso para la enseñanza de español. Las propuestas específicas de ampliación se presentan a través de diferentes tablas que vamos intercalando en nuestra discusión y valoración crítica de los resultados.

En su conjunto, nuestra tesis pretende revisar oposiciones diatópicas del léxico hispánico, tratando de establecer una metodología propia de aproximación al fenómeno de la representatividad. Ante la imposibilidad de presentar una taxonomía lingüística del léxico, nuestro objetivo consiste en proponer una ampliación del vocabulario del *PCIC* que se basa en un análisis demolingüístico y permite incorporar elementos fundamentales de otras variedades geolectales, tratando con rigor la distribución de las variantes léxicas en la enseñanza de español como segunda lengua o lengua extranjera.

CAPÍTULO 1

EL PCIC Y LA VARIACIÓN LÉXICA

1. El Plan Curricular del Instituto Cervantes: descripción general

1.1. Introducción

La presente investigación se articula en torno al *Plan Curricular del Instituto Cervantes* (en adelante, *PCIC*). El objetivo último de la tesis es revisar y proponer una ampliación del repertorio léxico de este material, incorporando nociones altamente representativas de toda la geografía hispánica. Para comenzar, en este primer capítulo vamos a presentar una descripción general de esta obra de referencia. Se trata de comprender el sentido, la importancia y el enfoque de este documento fundamental en el ámbito de la lingüística aplicada a la enseñanza de idiomas.

El *PCIC* es el material que fija y desarrolla los componentes de aprendizaje que necesita un estudiante de español para lograr objetivos comunicativos. A partir de los niveles de referencia establecidos por el Consejo de Europa, el Instituto Cervantes especifica el material de enseñanza y aprendizaje para la didáctica de Español como Lengua Extranjera (en adelante, *ELE*). Sobre la base de la variedad peninsular centro–norte española, el documento incluye elementos lingüísticos y extralingüísticos que especifican, por ejemplo, las unidades gramaticales, las

estrategias pragmáticas o el vocabulario que deben manejar los estudiantes de español en los diferentes niveles de aprendizaje.

Desde una línea avanzada en la investigación sobre planificación curricular de lenguas extranjeras, el *PCIC* ofrece objetivos generales y contenidos de enseñanza y aprendizaje en cada nivel del *Marco Común Europeo de Referencia* (en adelante, *MCER*): A1–A2, B1–B2, C1–C2. De esta manera, el Equipo Coordinador del Instituto Cervantes y un comité de especialistas detallan los elementos gramaticales, nocionales, pragmático–discursivos, culturales y de aprendizaje que deben conocer los estudiantes en cada nivel.

Para los esquemas de cada inventario así como para la selección, el tratamiento y la secuenciación de los contenidos en niveles de dominio, colaboraron unos 250 profesores del Instituto Cervantes y decenas de especialistas. Un grupo de expertos revisó y depuró el documento siguiendo las más recientes investigaciones sobre adquisición lingüística y a través de una pluralidad de enfoques teóricos. La base fundamental para la selección del material fue el conocimiento especializado, la experiencia en la práctica docente y la revisión de los sucesivos borradores por parte de profesionales de ELE.

El camino para la elaboración del *PCIC* podría haber sido el inverso. En lugar de establecer niveles de aprendizaje con un comité de expertos, estos niveles se podrían haber fijado partiendo de muestras de lengua de estudiantes de español, que informarían de los elementos que saben utilizar los estudiantes en cada etapa de aprendizaje. Esta vía, mucho más laboriosa, permitiría situar con mayor precisión aquello que los alumnos de un determinado nivel saben hacer, decir o comprender. De hecho, la *Guía para la elaboración de descripciones de niveles de referencia para las lenguas nacionales y regionales* del Departamento de Política Lingüística del Consejo de Europa recomendaba para la descripción de los niveles de referencia el uso de inventarios de frecuencia estadística y otros datos lingüísticos producidos en situaciones de evaluación. La falta de un corpus representativo, fiable y digitalizado de muestras de lengua de estudiantes de español que recogiera

exámenes, actividades, materiales de clase, etc., hizo que este camino fuera imposible para el Instituto Cervantes.

En todo caso, la «Introducción general» del *PCIC* invita, tanto a investigadores como a profesores, a revisar el documento para que pueda «contrastarse, corregirse y mejorarse en sucesivas ediciones, en la medida en que puedan introducirse elementos más sofisticados de selección y control» (Instituto Cervantes, 2006). Nuestra tesis es precisamente una propuesta de ampliación del léxico del *PCIC*, a través de mecanismos de selección y control que permiten contrastar y mejorar el vocabulario incluido en las «Nociones específicas».

1.2. Antecedentes y bases del *PCIC*

Si el *PCIC* define y desarrolla los niveles de referencia europeos para el español, debemos explicar cómo surge la idea de establecer niveles de referencia y en qué contexto se plantea su aplicación a la lengua española. Existen dos publicaciones técnicas del Consejo de Europa que constituyen las principales bases del *PCIC*:

—Las descripciones de niveles de competencia lingüística desarrolladas a partir del *Threshold Level* en los años 70.

—El *MCER* como documento de referencia en el año 2001.

Desde su creación en 1949, el Consejo de Europa ha buscado diversas formas de acercamiento y de convergencia entre los estados europeos, entendiendo la diversidad cultural como fuente de riqueza. En el ámbito de la política lingüística, ha desarrollado diferentes proyectos y conferencias intergubernamentales para promover la comunicación, el intercambio entre ciudadanos y el plurilingüismo, teniendo como objetivo último desarrollar un enfoque europeo de la enseñanza de lenguas.

En 1975 se publica el *Threshold Level (T-Level)*, ante la necesidad de definir unos niveles de enseñanza y aprendizaje comunes para los ciudadanos europeos.

Basado en situaciones reales de comunicación, el *T-Level* fue el primer gran logro del Consejo de Europa en lo que respecta a la enseñanza de idiomas. El enfoque nociofuncional centrado en el uso lingüístico (una conducta observable) y no en un sistema formal de lengua (un conjunto de estructuras sintácticas, fonológicas y de campos léxicos) tuvo una gran repercusión. De este modo, el *T-Level* describe un grado de dominio general que permite a los hablantes relacionarse con eficacia satisfaciendo unas necesidades comunicativas básicas similares para todos los hablantes europeos (Slagter, 1979). Este documento incluye ejemplificaciones en inglés pero tiene la vocación de señalar las funciones lingüísticas esenciales para ciudadanos adultos de cualquier lengua. Años después, el *T-Level* se amplió con la publicación de otras series, constituyendo la base del segundo gran documento del Consejo de Europa: el *MCER*.

En el año 2001 nació el *MCER* para «proporcionar una base común para la elaboración de programas de lenguas, orientaciones curriculares, exámenes, manuales y materiales de enseñanza en Europa» (Consejo de Europa, 2001). Este documento supuso la culminación de la política lingüística europea de convergencia. El Consejo de Europa encomendaba a cada lengua europea que desarrollara un material lingüístico específico, con las competencias comunicativas y generales que caracterizan a los estudiantes de cada nivel, para avanzar en la colaboración internacional entre países en el campo de la investigación sobre lenguas y educación.

A pesar de las críticas que haya podido suscitar (Figueras, 2008), el *MCER* ha permitido avances fundamentales en el ámbito de la adquisición lingüística de segundas lenguas, garantizando la equivalencia de las certificaciones y competencias evaluadas en Europa y fomentando una enseñanza de calidad. Su influencia, de hecho, ha trascendido el ámbito europeo, convirtiéndose en un documento de suma importancia para ministerios de educación, entidades certificativas, creadores de materiales o instituciones de diferentes continentes.

Una de las aportaciones esenciales del *MCER* es la definición de las cuatro competencias generales que influyen en el aprendizaje: *saber* (conocimiento declarativo), *saber hacer* (conocimiento procedimental), *saber ser* (competencia existencial) y *saber aprender* (que moviliza las competencias anteriores para permitir el aprendizaje). También presenta un análisis detallado de las competencias comunicativas lingüísticas, sociolingüísticas y pragmáticas, e incorpora conceptos esenciales sobre la educación, como el *aprendizaje a lo largo de la vida* o la importancia de la *autonomía* de los estudiantes.

Además, el *MCER* incluye:

- Una *escala global* con seis niveles de competencias: A1–A2, B1–B2, C1–C2.
- Unos *descriptores generales* de cada nivel, que definen lo que un estudiante es capaz de hacer mediante el uso de la lengua.
- *Actividades* (de producción, recepción e interacción) que ponen en marcha las competencias lingüísticas en cada nivel.

Las *escalas de descriptores* en 6 niveles de aprendizaje son seguramente el componente clave de este documento. Responden a la pregunta de qué es capaz de hacer un alumno en cada nivel, ofreciendo descripciones generales que se deben adaptar a cada lengua con unos contenidos gramaticales, léxicos, pragmáticos o socioculturales específicos. La obra sobre la que articulamos nuestra tesis, el *PCIC*, es el resultado de la especificación del *MCER* para el español.

Además del *MCER* y de las series del *T-Level*, para elaborar el *PCIC* el Instituto Cervantes también tuvo en cuenta su propio plan curricular de 1994, que definía de manera general cuatro niveles de aprendizaje (García Santa-Cecilia, 2007). También se consultaron otros documentos de relevancia internacional, como los descriptores de *ALTE Can Do Project* (1992–2002), el sistema de evaluación diagnóstico de lenguas *Dialang* (1996–2002), el *Portfolio europeo de las lenguas* (2001) y las primeras adaptaciones de los niveles de referencia para el nivel B2 del francés, publicadas en 2004 (*Niveau B2 pour le français: un référentiel*). También se

utilizaron diferentes diccionarios y bibliografía especializada, y se tuvieron en cuenta las investigaciones recientes en sociolingüística, pragmática, psicolingüística, teoría de la educación o análisis del discurso.

1.3. Características generales del *PCIC*

Como ya hemos comentado, el *PCIC* desarrolla y especifica los niveles de referencia europeos del *MCER* para el español. Presentamos a continuación las características esenciales del documento base de nuestra tesis:

1) Nace de una reflexión profunda del español y es forzosamente diferente a las especificaciones de otros idiomas.

2) Supone un nexo de unión entre la teoría y la práctica en la didáctica de español como segunda lengua o lengua extranjera.

3) Presenta una coherencia interna entre los distintos niveles y temas, con un sistema unificado de epígrafes y referencias cruzadas que los interrelacionan.

4) Concreta de manera exhaustiva las etapas de aprendizaje aunque exige un ajuste de los centros educativos y del profesor; no puede llevarse directamente al aula sino que debe adaptarse al contexto de aprendizaje.

5) Está basado en los descriptores del *MCER* y en especificaciones del *T-Level*, incluyendo los inventarios mínimos recomendados por el Consejo de Europa («Funciones», «Nociones» y «Gramática»), además de otros contenidos relacionados con el discurso, la dimensión cultural, la fonética, la ortográfica o las estrategias de aprendizaje.

6) Constituye un currículo nociofuncional que, a modo de marco articulador del gran sistema de la lengua, integra de manera natural la gramática y la comunicación.

En conclusión, el *PCIC* no es una traslación del *MCER* ni una traducción del *T-Level*. Tampoco es un tratado lingüístico normativo ni un programa de lengua para llevar directamente a la clase. El *PCIC* es un material específico para el

español que constituye un primer nivel de concreción lingüística y exige otros niveles de ajuste y adaptación por parte de sus usuarios para dotarlo de pleno sentido.

1.4. Estructura y organización

Comentamos ahora la estructura y organización del *PCIC* para enmarcar y situar el inventario de «Nociones específicas», sobre el que centraremos nuestra investigación. El *PCIC* presenta una triple dimensión en lo que respecta al sujeto de aprendizaje. Por una parte, fiel a su óptica comunicativa, considera al estudiante un *agente social*, conocedor de elementos lingüísticos y extralingüísticos que le permiten desenvolverse en la interacción. El estudiante se concibe, además, como un *hablante intercultural* que identifica aspectos relevantes de la nueva cultura y establece puentes entre su cultura de origen y la cultura hispánica. Por último, el estudiante se considera un *aprendiente autónomo* que, tal y como definía el *MCER*, es responsable de su aprendizaje y va avanzando más allá del propio currículo, desde una perspectiva activa que facilita el aprendizaje a lo largo de la vida.

En relación al objeto de aprendizaje, el *PCIC* presenta tres secciones, organizadas según los niveles del *MCER* (A1–A2, B1–B2 y C1–C2). En cada una de ellas se presentan los objetivos generales del nivel, más doce inventarios o temas. Estos inventarios cubren cinco componentes de lengua distintos, como vemos a continuación:

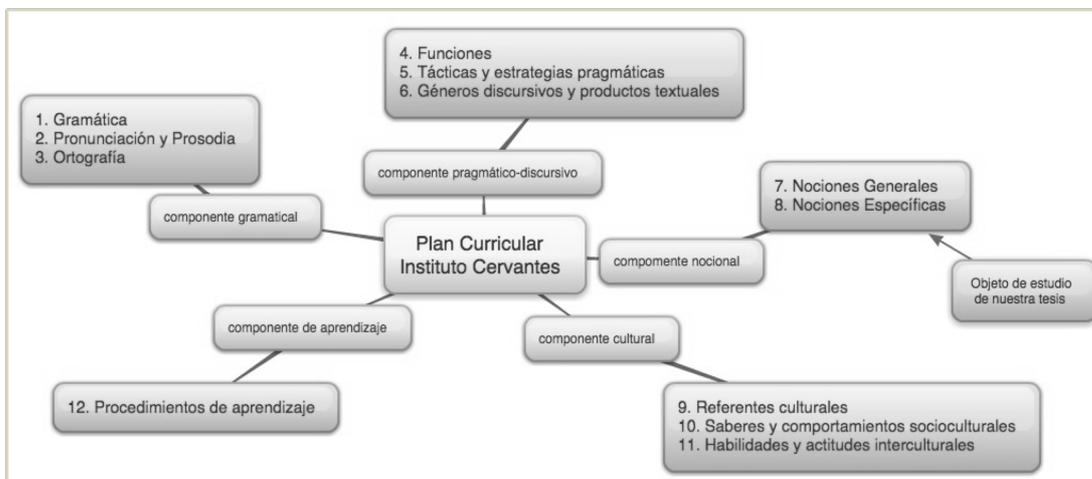


Figura 1: Organización y estructura del PCIC

En esta tesis revisaremos el componente nocional del *PCIC*. En particular estudiaremos el inventario de «Nociones específicas», que fija y describe las nociones léxicas que debe conocer el estudiante de español en los distintos niveles de referencia.

1.5. Utilidad e importancia del *PCIC*

La importancia que ha adquirido el *PCIC* en los últimos años ha aumentado de forma exponencial hasta convertirse en la obra de referencia para la enseñanza de español en todo el mundo. En este sentido, nuestra propuesta de ampliación y mejora de este material adquiere un extraordinario interés en el campo de la didáctica de ELE. La secuenciación del aprendizaje que propone el Instituto Cervantes así como la distribución de los componentes lingüísticos en niveles de dominio es fundamental para autores de materiales, editoriales, estudiantes, entidades encargadas de la certificación de los niveles de aprendizaje y, por ende, para cualquier institución u organismo dedicado a la difusión del español.

Un material con una descripción lingüística tan profusa ofrece múltiples aplicaciones para los profesores de ELE, entre las que destacaríamos:

- Facilitar la primera toma de contacto con los estudiantes.
- Programar con rigurosidad el currículo de un curso de español.

- Secuenciar el aprendizaje de manera sistemática.
- Definir lo que se va a aprender y consensuar el aprendizaje.
- Desarrollar materiales significativos, diseñar tareas y preparar las clases.
- Favorecer la integración de destrezas.
- Evaluar críticamente los materiales que llevamos al aula.
- Determinar qué vamos a evaluar y diseñar pruebas de evaluación.
- Ejercer funciones de autoevaluación.
- Revisar críticamente el propio *PCIC*.

En palabras de César Antonio Molina: «[U]na obra de esta naturaleza se va afianzando con el tiempo a partir de las observaciones que hacen quienes la utilizan» (Navarro Giner y Navarro Serrano, 2008). Por ello, este documento nunca llegará a ser definitivo o terminal. En este sentido, el *PCIC* es un proyecto colectivo que responde a un planteamiento abierto, consensuado y que puede ser mejorado en ediciones sucesivas. Tanto lingüistas como profesores debemos reformularlo, complementarlo, revisarlo y enriquecerlo con experiencias e investigaciones.

Nuestra tesis surge de las necesidades y problemas teórico-prácticos que tienen los profesores de ELE para integrar con rigor la variedad léxica en el currículo. Para ello proponemos una revisión y ampliación del inventario número 8, que se corresponde con las «Nociones específicas» de aprendizaje, con el objetivo de ofrecer un repertorio léxico del español más completo gracias a una investigación demolingüística.

2. El *PCIC*: variedad lingüística y variación léxica

2.1. Unidad y diversidad del español

En nuestra propuesta de ampliación del léxico del *PCIC* se concibe la lengua española como un «complejo dialectal» (García de Diego, 1950), lo que produce diversas consideraciones ideológicas relacionadas con la estandarización, la fragmentación o la cohesión del idioma. La inmensa amplitud geográfica del español

constituye un terreno fértil para la variación, lo que ha dado lugar a diferentes actitudes respecto a la diversidad. En todo caso, considerar otras variedades del español es absolutamente necesario para que los estudiantes de ELE puedan desarrollar su competencia intercultural y se relacionen con hablantes de cualquier procedencia.

En español todos «hablamos alguna variedad de una lengua abstracta común y compartimos propiedades de varios dialectos. Somos únicos y mestizos; nacionales, internacionales y locales» (Demonte, 2003: 1). Los hablantes de español somos conscientes de las múltiples variedades lingüísticas que enriquecen nuestro idioma. A pesar de que en algunas épocas la heterogeneidad del español se entendía como «jerigonza bárbara» (Iglesias y Sopova, 2012), la diversidad lingüística de nuestra lengua ha sido el motor de excepcionales investigaciones y ha suscitado gran interés por la doble dimensión que comporta: unitaria y heterogénea.

Tal diversidad plantea la necesidad de definir una variante común a las diferentes áreas geográficas del español que denominamos *lengua estándar*:

La lengua de intercambio de una comunidad lingüística, legitimada e institucionalizada históricamente, con carácter suprarregional, que está por encima de la(s) lengua(s) coloquial(es) y los dialectos y es normalizada y transmitida de acuerdo con las normas del uso oral y escrito correcto. Al ser el medio de intercomprensión más amplio y extendido, la LE [lengua estándar] se transmite en las escuelas y favorece el ascenso social; frente a los dialectos y sociolectos, [es] el medio de comunicación más abstracto y de mayor extensión social (Lewandowski 1982: 201).

En función del prestigio y de la historia de la lengua nace una lengua estándar como convención que garantiza la extensión de las lenguas y su supervivencia. Sin embargo, «a totally uniform, regionally neutral, and unarguably prestigious variety does not yet exist worldwide» (Crystal, 1994: 113). En la misma línea, Lippi-Green (1997) afirma que el estándar se concibe como un ideal en la mente de los hablantes pero no como una realidad lingüística. En este sentido la lengua estándar es un constructo abstracto de la lengua común en el que se desdibujan o neutralizan los rasgos regionales en favor de los elementos lingüísticos que todos los hablantes comparten. Así, los rasgos del español estándar

no configuran un sistema exhaustivo y homogéneo sino que surgen por contraste y debilitación de los rasgos específicos de las múltiples variedades (Demonte, 2003).

Ningún hablante puede hablar la lengua estándar. Todos utilizamos variedades subestándar¹. En el caso del inglés, muchos estudios se centran en la descripción de los rasgos de comunicación que lo caracterizan como lengua franca (Pennycook, 2008; Rubdy y Saraceni, 2006; Kirkpatrick, 2006; Jenkins, 2006). En el caso del español a nivel mundial, la realización lingüística que más se aproxima a una lengua estándar internacional es el denominado *español neutro*², que está presente sobre todo en los medios de comunicación para dirigirse a una audiencia globalizada (Andión Herrero, 2008; Bravo, 2008). Este tipo de español garantiza la comunicación internacional, sin escoger los términos más frecuentes sino los menos marcados, es decir, aquellos que pertenecen al español puesto que todo el mundo los comprende, aunque no sean los más utilizados (López Morales, 2010).

Es importante no identificar lengua estándar con norma lingüística. En el plano normativo emitimos juicios de valor desde una visión conservadora que señala aquello que se debe decir. Mientras que el *estándar* es por su naturaleza una imprecisión, la *norma* es precisamente la especificación de lo correcto, «lo que se dice allí donde se dice y cabe decirlo», y lo incorrecto, «lo que no se dice allí ni cabe decirlo» (Coseriu, 1992: 14).

En este trabajo consideramos la *norma* como una descripción objetiva de *cómo* se dicen las cosas y no una entidad estática u homogénea sobre la que hacer prescripciones del tipo *correcto / incorrecto* (Coseriu, 1967). La lengua es, en realidad, un elemento flexible «en tensión permanente entre una fuerza centrípeta conservadora y otra centrífuga innovadora» (Mangado Martínez, 2006: 42). Si

¹ Tengamos en cuenta que si definimos lengua estándar como conjunto flexible de rasgos lingüísticos compartidos tenemos que afirmar que pueden existir diferentes lenguas estándar del español en función de los países para los que busquemos rasgos neutros comunes (Lapesa, 1992; Lope Blanch, 1997; Moreno de Alba, 2001).

² Interesante lectura sobre las características y limitaciones de esta variedad de español en Llorente Pinto (2013).

entendemos la norma como conjunto de realizaciones «vivas» del idioma, podemos afirmar que en español existirán pluralidad de normas diafásicas, diastráticas y diatópicas, dependiendo de las realizaciones particulares, sociales y geográficas de los hablantes.

En este sentido, el español puede definirse como un idioma plurinormativo o pluricéntrico, concepto acuñado por Kloss en 1978 (Torres Torres, 2013) y posteriormente por Clyne (1992), que fue desarrollado particularmente para la lengua española por Thompson (1992). Estas múltiples modalidades de referencia, con diferentes centros de prestigio interrelacionados, contribuyen a consolidar una lengua común como factor de integración favoreciendo una competencia plurilectal (Rivarola, 2006). Este código, tan evidente en la era de la globalización, constituye un producto ecléctico que favorece la comunicación masiva entre hispanohablantes, sin que por ello se debiliten las variedades lingüísticas regionales.

La constatación de las múltiples normas del español hace que algunos investigadores hayan señalado una posible fragmentación del idioma (Rona, 1958; Entwistle, 1969; Bartos, 1971). Lejos de esta concepción, y en línea con Rosenblat (1971), Criado de Val (1972), Llorente Maldonado (1986) y Alvar (1996), entre otros, defendemos en esta tesis que el español no se resiente por su diversidad sino todo lo contrario, se enriquece y fortalece en torno a una base común. En este sentido, nuestra ampliación dialectal del *PCIC* responde a una visión suprarregional integradora de las variedades lingüísticas que pretende aportar reconocimiento y democratización a la heterogeneidad del idioma y puede favorecer los vínculos políticos, intelectuales y económicos entre España y América, al tiempo que contribuye a consolidar la cultura hispánica.

Sin olvidar que las múltiples variedades lingüísticas que se distribuyen a lo largo de los países hispanohablantes dan forma a una lengua común que es, en su mayoría, compartida, la fuerte variabilidad exige herramientas de análisis que permitan tratar con rigor la diversidad. Una lengua con la extensión geográfica del

español exige un tratamiento exhaustivo de la variación, para el que se hacen imprescindibles tanto investigaciones como materiales lexicográficos sólidos que permitan integrar de forma sistemática este aspecto genuino e idiosincrático de nuestra lengua. En el capítulo 2 de esta tesis presentamos un estado de la cuestión de las investigaciones sobre variación lingüística que han desarrollado históricamente la dialectología y la geografía lingüística y nos centraremos especialmente en la variación léxica por ser el tema central de nuestro estudio.

2.2. La variedad lingüística en la enseñanza de español

Si toda enseñanza debe integrar y respetar la diversidad como forma de riqueza y práctica democrática (Pérez Gómez, 2002), la didáctica de lenguas integrará también las variedades dialectales en el currículo, con el objetivo de garantizar la comunicación de los estudiantes y de presentar la gran riqueza dialectal del español. Partiendo de la base de que no existe en el mundo hispánico una única variedad regionalmente neutra y prestigiosa (Vázquez, 2008), el profesor debe posicionarse en torno al tema de la diversidad lingüística.

El *MCER* atribuye una gran importancia a la integración de la variedad como elemento fundamental en la formación del estudiante. La lengua se considera así un fenómeno sociocultural que facilita las relaciones humanas y exige el conocimiento de marcadores lingüísticos, normas de cortesía, expresiones de sabiduría popular, registros, dialectos, acentos, etc. Para garantizar la comunicación el estudiante deberá conocer las variedades de la lengua meta. La variedad se escogerá en función de las circunstancias sociales y personales de enseñanza, en función de la proximidad geográfica y dependiendo de las necesidades de uso de los estudiantes. El *MCER* destaca que:

[E]l reconocimiento de dichos rasgos dialectales ofrece, pues, claves significativas respecto a las características del interlocutor [...]. Con el paso del tiempo, los alumnos también entrarán en contacto con hablantes de varias procedencias; antes de que puedan adoptar ellos mismos las formas dialectales, deberían ser conscientes de sus connotaciones sociales y de la necesidad de coherencia y constancia (Consejo de Europa, 2001: 131).

Así el profesor de español, además de la variedad principal que emplee en el aula, integrará en el curso otras variedades dialectales. Andión Herrero (2007) propone la distinción entre *variedad preferente*, que se corresponde con el modelo principal de lengua, y *variedad periférica*. Así, en función de los destinatarios, del contexto de enseñanza o de los objetivos del curso, el profesor escogerá una variedad que actúa como preferente, a la que pertenecen la mayoría de textos orales y escritos que se introducen en el aula, las actividades de aprendizaje, los contenidos lingüísticos, etc. Cualquier variedad del español podrá ejercer la función de variedad preferente: «Hay que descender a la realidad particular de cada lengua que se enseña y, además, tener presentes siempre a los destinatarios del aprendizaje» (Vila Pujol, 2009: 208).

El resto de variedades distintas a la variedad preferente que se incorporen en el curso de español se consideran variedades periféricas y son fundamentales para que el estudiante reconozca el habla de otras zonas lingüísticas y desarrolle estrategias de comunicación para interrelacionarse. Si tenemos en cuenta el criterio de rentabilidad en el aprendizaje, que pretende aportar a los estudiantes aquellas variedades que les permitan comunicarse en el mayor número de contextos posibles, el profesor de ELE debería elegir, en principio, las variantes diatópicas que están más extendidas en la geografía del español. En línea con esta búsqueda de elementos lingüísticos compartidos, hay que apostar por una norma híbrida para garantizar la competencia comunicativa. También Moreno Fernández (1997) destaca la importancia de los usos comunes del mundo hispánico.

En este sentido, y teniendo en cuenta que la elección de la variedad preferente estará condicionada por el contexto de aprendizaje, consideramos que es preferible abandonar las voces locales en favor de otras más rentables (Chacón García, 2014).

De manera general podemos afirmar que para el estudiante de español la variedad más rentable será la que garantice mayor neutralidad y extensión³. Es la variedad lingüística común o compartida por todos los hispanohablantes la que tiene que aparecer ineludiblemente en el aula de ELE, sea cual sea la variedad preferente del curso.

El *PCIC* señala la relevancia de conocer e introducir las variedades dialectales en el aprendizaje. Desde una posición abierta y flexible, invita al reconocimiento ecuaníme de las normas lingüísticas, anotando y comentando ciertos usos lingüísticos particulares de otras zonas hispánicas:

El español tiene la cualidad de ser una lengua que cuenta con varias normas cultas que pertenecen a diferentes localizaciones geográficas; la correspondiente a la norma centro–norte peninsular española es sólo una de ellas. De ahí que, como no podía ser menos, sean anotadas y comentadas especificaciones de considerable extensión en las que la norma central descrita no coincide con amplias zonas lingüísticas del mundo hispánico. Esta flexibilidad en el inventario compensa la restricción que supone describir preferentemente una de las muchas variedades del español y enriquece la representatividad del corpus. (Instituto Cervantes, 2006).

El Instituto Cervantes incorpora así una clara vocación integradora de las variedades, que sigue la línea actual de política panhispánica⁴. De hecho, el *PCIC* pretende incluir en sus inventarios aspectos lingüísticos propios de otras variedades del español siempre que tengan una alta frecuencia y sean de considerable extensión. Para que un estudiante pueda comunicarse con cualquier hablante de español es fundamental tener en cuenta el componente diatópico de la variación:

³ En este sentido, Andión Herrero (2008) expone dos criterios ineludibles para seleccionar la lengua *estándar* del aula de ELE son:

- *Principio de comunidad*: los elementos deben ser troncales o comunes a los hablantes.
- *Principio de neutralidad*: los elementos deben tener un grado de independencia más o menos alto de contextos específicos.

⁴ Las políticas lingüísticas panhispánicas actuales se constatan en las siguientes publicaciones: *Diccionario panhispánico de dudas* (2005), el *Diccionario del estudiante* (2005), la *Nueva gramática de la lengua española* (2009), el *Diccionario de americanismos* (2010), la *Ortografía de la lengua española* (2010), *El buen uso del español* (2013) y el *Diccionario de la lengua española* (2014), todas con un marcado carácter panhispánico.

El respeto a las variedades lingüísticas, el aprender a negociar significados, no solo desarrolla las destrezas lingüísticas en la clase, sino que también refleja una necesidad comunicativa de todo el mundo de habla hispana. Aprender a adaptarse a distintas variedades lingüísticas, apreciar la riqueza de las variedades dialectales, y explorar maneras de descubrir y comprender diferencias lingüísticas prepara al estudiante para comunicarse en cualquier comunidad de habla hispana (Villa, 1996: 197).

Es importante destacar que no se trata solo de escoger *una* variedad lingüística que vertebré el currículo —cuestión que recogen, entre otros, Beaven y Garrido (2000), Flórez Márquez (2000), Rivarola (2006), Andión Herrero (2008), Vázquez (2008)— sino que la variedad dialectal deberá integrarse transversalmente en el currículo. Tanto el docente como sus materiales de referencia deben estar preparados para «transcurrir transversalmente a otras variedades» (Martín Perís, 2001: 134), aprendiendo a secuenciar la variedad lingüística en el aula y proponiendo una suma de variedades que garantice el éxito de la comunicación de los estudiantes. En este sentido, el docente de ELE debería ser conocedor de una norma panhispánica estándar porque «exponerles a toda la riqueza y a la variedad lingüística y cultural del mundo de habla hispana es, a nuestro parecer, la única alternativa viable» (Beaven y Garrido, 2000).

En relación con el diseño y la publicación de materiales de enseñanza de ELE, podemos afirmar que la integración de las variedades dialectales no ha sido una verdadera prioridad (Flórez Márquez, 2000). Aunque se aprecian algunos esfuerzos por incorporar variedades dialectales en los manuales de español, la mayor parte de los recursos didácticos se redactan siguiendo la norma peninsular y, cuando introducen variedades dialectales, las tratan de manera anecdótica o imprecisa (García Fernández, 2010).

Otra cuestión importante es cuándo debemos introducir la variedad en los diferentes niveles de aprendizaje. Según Blanco (2000) y Andión y Casado (2014), los elementos deben introducirse desde estadios iniciales del aprendizaje, mientras que Mendizábal (1997) o England (1998) defienden su incorporación en niveles intermedios o avanzados. Vila Pujol (2009) defiende el

criterio de máxima rentabilidad por lo que la integración de las variedades más restringidas se hará en función de las necesidades específicas de los estudiantes.

2.3. Norma lingüística del *PCIC*

Ante la enorme extensión geográfica de la lengua española y debido a la diversidad de normas que confluyen en el mundo hispánico, el Instituto Cervantes debe escoger una variedad lingüística para definir los niveles de referencia del español. Así, en la «Introducción» se explica:

Estos presupuestos han sido considerados a la hora de seleccionar el material lingüístico que aparece en el *Plan curricular del Instituto Cervantes* y sus *Niveles de referencia para el español* y que corresponde preferentemente a la norma culta de la variedad centro–norte peninsular española. La selección de esta variedad se sustenta en los rasgos comunes que comparte con las restantes normas cultas del mundo hispánico y en su proyección dentro del modelo estandarizado prestigioso de la lengua para la propia comunidad hispánica, a lo que habría que añadir la propia adscripción de la institución encargada de elaborar el Plan curricular (Instituto Cervantes, 2006).

Así, aunque la variedad culta peninsular es solo *una* de las múltiples variedades dialectales de la lengua, la elección se justifica tanto por la proyección internacional del español peninsular como por la adscripción del Instituto Cervantes a la norma española.

Sin embargo, si todas las variedades del español son, por derecho propio, susceptibles de ser la variedad preferente de un curso de ELE (Andión Herrero, 2003), podemos preguntarnos qué anotaciones podría incluir el *PCIC* que permitieran a los profesores de español integrar, tratar y secuenciar la variedad geolectal en el currículo. Esta cuestión, ineludible para la enseñanza de español, se ha considerado en la mayoría de inventarios del *PCIC*. Veamos algunos ejemplos a continuación:

Tabla 1*Ejemplos de integración de la variedad en los inventarios del PCIC*

Inventario	Apartado	Anotaciones
Objetivos generales	§2. El alumno como hablante intercultural.	Incrementar el grado de sensibilidad cultural y estimular la actitud de apertura hacia otras culturas, en particular hacia las de España e Hispanoamérica.
Gramática	A2 §5. Los posesivos.	[Hispanoamérica] Tendencia a expresar la posesión mediante estructuras pospuestas. Posesivos pospuestos analíticos y sintéticos: <i>el libro tuyo</i> (tu libro) <i>la escuela de nosotros</i> (nuestra escuela)
Gramática	B1 §1.2. El género de los sustantivos.	Términos con género distinto en las variedades del español: <i>la bombilla</i> [España] / <i>el bombillo</i> [México, América Central, Las Antillas, Colombia y Venezuela] <i>el bolso, la costumbre</i> [España] / <i>la bolsa, el costumbre</i> [parte de México, Costa Rica y Colombia] <i>el salón</i> [España] / <i>la sala</i> [Hispanoamérica]
Gramática	C2 §2.2. El género de los adjetivos.	[Hispanoamérica] Tendencia a distinguir el género de ciertos adjetivos que no tienen marca de género en el español estándar y a crear formas masculinas en –o: <i>bromisto, cuentisto</i>
Pronunciación y Prosodia	C1C2 §5.2.4. Consonantes alveolares.	[Hispanoamérica] /s/ predorsal o dorsoalveolar, articulada con el ápice de la lengua hacia abajo [excepciones: Antioquia (Colombia), Costa Rica (valle central), La Habana (Cuba), Arceibo y Guaynabo (Puerto Rico)]; apicoalveolar en el centro de Bolivia; apicodental en el norte de México, la sierra de Ecuador y sierra de Perú y norte de Argentina. Variantes ciceantes aisladas en El Salvador, Nicaragua, Honduras, norte de Venezuela, y en algunas localidades de México, Colombia y Puerto Rico.
Ortografía	C1.C2 §1.3.4. Letra h.	[Hispanoamérica] Pronunciación ante los diptongos <i>ua</i> , <i>ue</i> , <i>ui</i> en posición inicial como una versión suave del fonema velar sonoro /g/. Incidencia en la escritura <i>hua</i> / <i>gua</i> , <i>hue</i> / <i>güe</i> , <i>hui</i> / <i>güi</i> – <i>huaca</i> / <i>guaca</i>

Tácticas y estrategias pragmáticas	y B1.B2 §1.2.5. Controladores del contacto.	[Zonas voseantes de Hispanoamérica] <i>¿sabés?, ¿entendés?</i>
Tácticas y estrategias pragmáticas	y C1.C2 §1.7.2. La negación con refuerzo.	[Hispanoamérica] Frases hechas <i>¡ni modo!</i> [México y América Central] <i>¡qué esperanza!</i> [México, Puerto Rico, Ecuador, Bolivia, Río de la Plata] <i>¡qué capaz!</i> [México, Guatemala, Ecuador]
Nociones generales	B1B2 §1.6. Certeza, incertidumbre.	<i>tal vez, a lo mejor</i> [México, Cuba y Venezuela] <i>a la mejor</i> [Hispanoamérica] <i>de repente</i> <i>con seguridad</i> [Río de la Plata, Colombia y Chile] <i>a la fija</i> <i>por supuesto</i> [América Central, Cuba, Colombia, Perú, Uruguay, Argentina y Chile] <i>de juro</i>
Referentes culturales	§1.4. Organización territorial y administrativa.	Bebidas típicas de las demarcaciones territoriales <i>horchata, sidra</i> (España); <i>chicha, mazato de yuca</i> (Perú); <i>fresco de aguacate, guarapo costeño</i> (Colombia).
Habilidades y actitudes interculturales	y §1.1. Configuración de una identidad cultural propia.	Explicación de comportamientos de la cultura española o la de otros países hispanos empleando los referentes propios de la cultura de origen del interlocutor.

Podemos constatar que en ocho de los trece inventarios del *PCIC* se incluyen anotaciones sobre variedades lingüísticas y culturales hispanoamericanas. No se incluyen anotaciones ni en el inventario de «Funciones» ni en el de «Saberes y comportamientos socioculturales», donde se dice explícitamente que las especificaciones se refieren únicamente a España y no a países de Hispanoamérica debido a la dificultad de realizar un análisis de la enorme variedad y a la falta de estudios rigurosos que permitan reflejar de forma completa el fenómeno de la variación. Tampoco aparecen anotaciones

hispanoamericanas en los inventarios «Géneros discursivos y productos textuales» ni en el de «Procedimientos de aprendizaje» puesto que no resulta pertinente.

El inventario restante, que corresponde a las «Nociones específicas» del español, es el centro de esta tesis. Nuestra propuesta es incorporar en el repertorio léxico del *PCIC* una serie de anotaciones dialectales que permitan integrar la variedad diatópica del español de manera rigurosa y con un nivel de exhaustividad similar al resto de inventarios plurilectales:

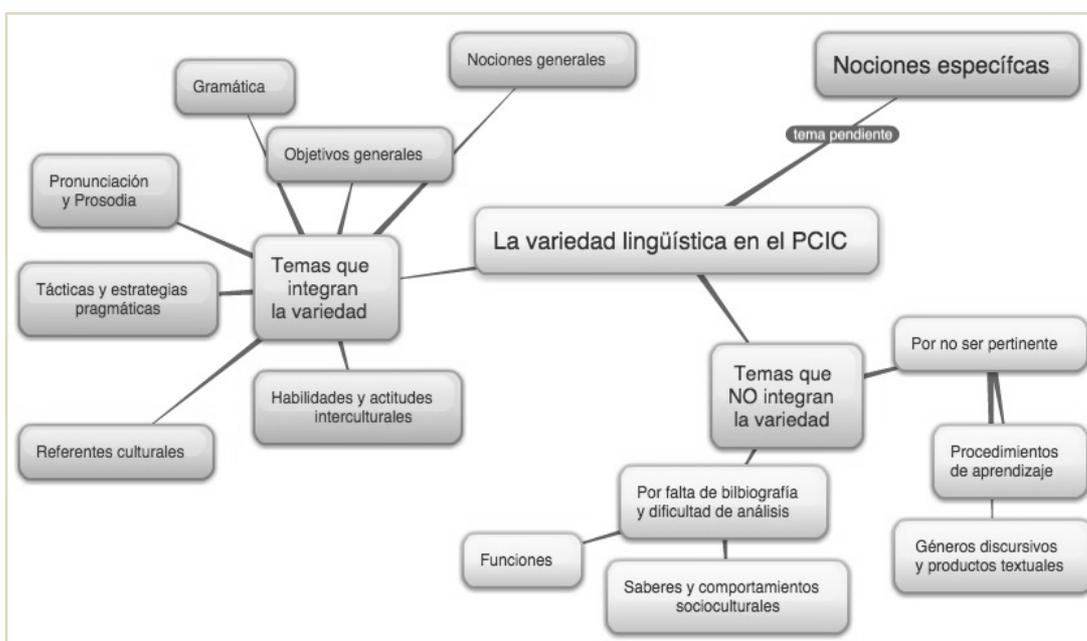


Figura 2: La variedad lingüística en el PCIC

2.4. Las «Nociones específicas» del *PCIC* y la variación léxica

Como ya hemos advertido, mientras que en otros inventarios se incluyen estructuras relacionadas con las variedades diatópicas del español, en el inventario de «Nociones específicas» no aparecen marcas geolectales que aporten una visión integradora de la variedad.

Reflejar el fenómeno de la variación léxica y secuenciarlo en niveles de aprendizaje a través de un inventario de carácter taxonómico, constituye un trabajo

verdaderamente complejo para el que sería necesario desarrollar una investigación profunda sobre rentabilidad, disponibilidad y frecuencia nociofuncional, lo que excede los objetivos de un plan curricular. Sin embargo, obviar las diferencias léxicas geolectales en las «Nociones específicas» tampoco se corresponde con la visión integradora de la variación que sostiene el *PCIC* en el resto de inventarios.

De hecho, en la introducción de las «Nociones específicas» se explica:

[A]unque se han incluido, como en otros inventarios, especificaciones sobre las variedades del español de América en relación con determinados elementos, ha de tenerse presente que las series de exponentes responden a un criterio de selección basado en la variedad centro–norte peninsular española y en las características idiosincrásicas de España (Instituto Cervantes, 2006).

Resulta llamativo que se afirme que se han incluido especificaciones sobre las variedades del español porque, aunque quizá ese fuera el propósito inicial del proyecto, no encontramos anotaciones relacionadas con la variación léxica en ninguna de las secciones del inventario. De hecho, como veremos a lo largo de esta tesis, algunas de las nociones del *PCIC* no cuentan con gran extensión en el mundo hispánico y se utilizan casi exclusivamente en el territorio peninsular. En estos casos es importante incorporar, si existen, nociones equivalentes que sean representativas de la mayoría de hispanohablantes.

La inexistencia de especificaciones sobre las variedades diatópicas en este inventario puede deberse, en primer lugar, a la gran variabilidad del componente léxico. También puede explicarse porque en el momento de redacción del *PCIC* aún no se había publicado el *Diccionario de americanismos* (2010). El carácter variable del léxico así como la falta de exhaustividad en las investigaciones lexicográficas de gran dominio dificultan la inclusión de variantes léxicas dialectales fiables y representativas, como veremos en el capítulo 2 de esta tesis.

Más adelante, el *PCIC* invita al profesor de ELE a adaptar el repertorio léxico a otras situaciones de enseñanza:

[Q]uienes utilicen el inventario en situaciones de enseñanza y aprendizaje en las que se practique una variedad diferente habrán de hacer las adaptaciones oportunas, tanto

en la selección léxica propiamente dicha como, en su caso, en la distribución por niveles (Instituto Cervantes, 2006).

Cabe preguntarnos a qué materiales podríamos acudir para encontrar información sobre la extensión, la rentabilidad y la frecuencia de las voces lingüísticas. Al no existir una guía didáctica sobre el tratamiento de la variación geolectal ni repertorios léxicos panhispánicos para la enseñanza de ELE tampoco parece realista pedir a un profesor de español que se dedique a «buscar en libros, revistas especializadas, tesis, trabajos de investigación, actas de congresos... la información pertinente que le proveyera de conocimientos teóricos precisos» para cada contexto (Andión Herrero, 2008: 11).

En este sentido, la oportunidad de utilizar el *PCIC* como documento base para la enseñanza del léxico y enriquecerlo con las variantes léxicas de otras variedades diatópicas, de la misma manera que se hace con la fonología, la gramática, la ortografía o la pragmática, constituye un proyecto de investigación de excepcional interés. Sorprende el nivel de exhaustividad que incluyen algunos inventarios del *PCIC*, como hemos mostrado en la Tabla 1, y la ausencia de variantes léxicas relacionadas con el español de América.

Nuestra tesis constituye una propuesta holística del léxico del español fundamentada en que el español es una lengua plural y diversa, cuya diversidad «no impide el mutuo entendimiento de sus hablantes y puede enseñarse y aprenderse para conseguir una comunicación satisfactoria con cualquiera de ellos» (Instituto Cervantes, 2006).

Dado que el *PCIC* defiende la inclusión de variedades del español en América en la mayoría de los inventarios, es imprescindible plantear una revisión exhaustiva de las «Nociones específicas». Asumimos, primeramente, que la variación dialectal no responde a reglas categóricas sino a diferencias de frecuencia o a valores socioculturales (Bustos Gisbert, 1995). Asimismo sabemos que no es posible —ni pertinente para un plan curricular— reflejar todas las variedades lingüísticas de cada noción léxica en los 22 países de habla española. Partiendo de

ambas consideraciones planteamos una ampliación del inventario léxico del *PCIC* en tres direcciones:



Figura 3: Propuesta de revisión y ampliación del *PCIC*

(1) *Variedades léxicas panhispánicas: panhispanismos*

Más allá de las diferencias de vocabulario, las semejanzas léxicas son las que garantizan una comunicación eficaz. En el caso del español, los panhispanismos son las voces que garantizan la comunicación con hablantes de cualquier procedencia. Aunque el estudio de las semejanzas puede dar falsas áreas dialectales, por no ser de naturaleza transitiva, estamos de acuerdo con Goebel al considerar que:

[E]l eterno problema de las isoglosas tomaba la cuestión al revés. No se trata de registrar sin más las diferencias, porque son las semejanzas las que aseguran la comunicación de los hablantes. Por eso propone el estudio de las semejanzas interpuntuales, de los interpuntos en función comunicativa (García Mouton, 1991).

Aunque las semejanzas no nos sirven para delimitar áreas dialectales, en esta tesis defendemos que las variedades comunes del español nos permiten adoptar una perspectiva ecolingüística ante la diversidad, que tenga en consideración las implicaciones de escoger determinantes variedades en la enseñanza. Resulta fundamental, para garantizar la comunicación entre los hablantes, ir más allá de la fragmentación lingüística puesto que «una teoría científica ecológica aspira a un holismo epistémico que sea críticamente consciente de los déficit de las especializaciones eventuales» (Finke, 2002: 46). En este sentido, las variantes panhispánicas aportan esta mirada holística que supera la fragmentación de las variedades regionales.

La incorporación de términos panhispánicos que, aún sin ser necesariamente los más utilizados, son los que mejor se comprenden en todo el territorio hispanohablante constituye un acercamiento a una *koiné* panhispánica, es decir, una lengua estándar en la que se desdibujan los rasgos particulares. Esta homogeneización léxica, lejos de eliminar las variantes regionales, supone «un enriquecimiento del caudal expresivo de los hablantes, que tendrán a su disposición un lexicón más amplio, tanto activo como pasivo» (Demonte, 2001). Si las palabras panhispánicas comunes son tan imprescindibles como inevitables en aras de la comunicación entre hispanohablantes, cuánto más lo serán para los estudiantes de español. Como consecuencia de la globalización, la estandarización del léxico convierte al español en una lengua mundial que supone una ventaja por cuestiones demográficas, ecolingüísticas, interculturales, sociales, de aprendizaje y comerciales (Beaven y Garrido, 2000).

(2) Variedades léxicas americanas: americanismos

Respecto a la inclusión de voces americanas, en la mayoría de los inventarios del *PCIC* ya se incluyen anotaciones dialectales de Hispanoamérica, por lo que la justificación de nuestra aportación se imbrica en la política lingüística de la propia institución. Además, contribuye al acercamiento bidireccional de los hablantes españoles al español atlántico, en línea con la mirada panhispánica de la última década a la que nos hemos referido en el §2.2:

Con un espíritu de tolerancia mutua, con la voluntad de conocer y de comprender la variante lingüística del otro, será mucho más fácil mantener la unidad de la lengua. En vez de aislarse o de querer imponer la propia variante de español a los otros, hay que integrar la enorme riqueza de una lengua que en un espacio tan dilatado presenta naturalmente una gama de variedades, a las que tampoco hay que subestimar con un triunfalismo unitarista (Haensch, 2002: 58–59).

En términos de política lingüística, la variedad de las lenguas es consustancial a la identidad personal y social, por lo que su estudio y descripción es fundamental. Es legítimo y primordial reivindicar las especificaciones locales. Desde la

ecolingüística (Stork, 2008)⁵ aboga por la integración de variedades dialectales como un elemento cultural valioso que merece ser protegido y no como una desviación del ideal de lengua (Mühlhäusler, 2000).

(3) Marca [Esp] de españolismos

En relación con la marca [Esp], aunque la existencia de españolismos se justifique por la elección de una variedad lingüística española en el *PCIC*, el hecho de señalar las palabras exclusivas de España permite que el profesor busque alternativas panhispánicas o americanas cuando el contexto de enseñanza lo requiera. Por supuesto, la decisión de enseñar la palabra peninsular o la mayoritaria en el mundo hispánico estará condicionada tanto por la variedad nativa del profesor como por el contexto de enseñanza, pero con esta marca mostramos que ciertas palabras atribuidas al estándar en las obras de ELE son en realidad nociones restringidas al español de España.

Los tres puntos mencionados constituyen las líneas de ampliación fundamentales que seguiremos en nuestra tesis. Sabiendo que el actual director del Instituto Cervantes, García de la Concha, declaró en el 2012 que «el eje del Cervantes será América», el estudio holístico de las «Nociones específicas» adquiere excepcional interés.

2.5. Didáctica de la variedad léxica

Una vez expuesta la importancia de incorporar la variedad léxica al currículo de la enseñanza de español, la pregunta inmediata se refiere a los mecanismos didácticos que permiten aprender y utilizar nuevas formas léxicas a los estudiantes. Aunque podríamos extendernos mucho más en esta cuestión relacionada con la adquisición léxica, presentamos una síntesis del funcionamiento del léxico en nuestra mente, aportando algunas claves didácticas que favorecen la adquisición de palabras.

⁵ Recomendamos la lectura de este excelente artículo en el que se aplican los postulados ecolingüísticos a la situación de la lexicografía hispanoamericana.

La idea central es que la variedad léxica no se puede enseñar como algo desestructurado o anecdótico sino que debe integrarse en el currículo a modo de eje transversal que va incorporando nociones en los distintos niveles, al tiempo que se retoman y amplían significados y usos de palabras ya aprendidas.

Flórez Márquez (2000) defiende que «la forma más efectiva de darles este *input* comunicativo» de las variaciones es a través de actividades orientadas a la adquisición, en donde las variantes aparezcan en contextos comunicativos reales. Esta presentación de las variaciones debería ser gradual no solo en cuanto al número sino también respecto al tipo de variantes.

El aprendizaje de una palabra es un proceso complejo, que es posible gracias a la dotación genética para desarrollar el lenguaje. Cada vez que aprendemos una palabra, nuestra mente realiza una serie de operaciones para incorporarla de manera eficaz a un almacén de palabras que denominamos el *lexicón mental*. Se trata de un conjunto de piezas léxicas que constituye un sistema articulado de rasgos, donde se especifican las propiedades fonéticas, semánticas y sintácticas de las palabras. El lexicón permite tanto la comprensión (vocabulario *pasivo* o *receptivo*) como la producción del habla (vocabulario *activo* o *productivo*). Además de estos dos tipos de vocabulario, existe el vocabulario *en potencia*, que representa aquellas palabras que, aunque no forman parte del lexicón, se pueden deducir mediante estrategias.

La incorporación de una unidad léxica al lexicón mental supone, por tanto, un procesamiento de distintas informaciones. No solo almacenamos el significado sino que, en nuestra mente, adquirir una palabra supone una labor compleja que codifica:

- a. la forma (fonética, fonológica, ortográfica, ortoépica, morfológica)
- b. la función sintáctica (categoría y función)
- c. el significado (semántica)
- d. el conocimiento experiencial
- e. el conocimiento enciclopédico
- f. su variación de uso (sociolingüística)
- g. su valor intencional y comunicativo (pragmática).

De este modo, el lexicon no es un simple listado de palabras, como un diccionario, sino que almacena una gran cantidad de información interrelacionada de cada unidad léxica. Sería algo más parecido a la red semántica del programa *Visual Thesaurus*, donde las unidades aparecen en medio de una red de relaciones (semánticas, morfológicas, pragmáticas, etc.) sin orden jerárquico:

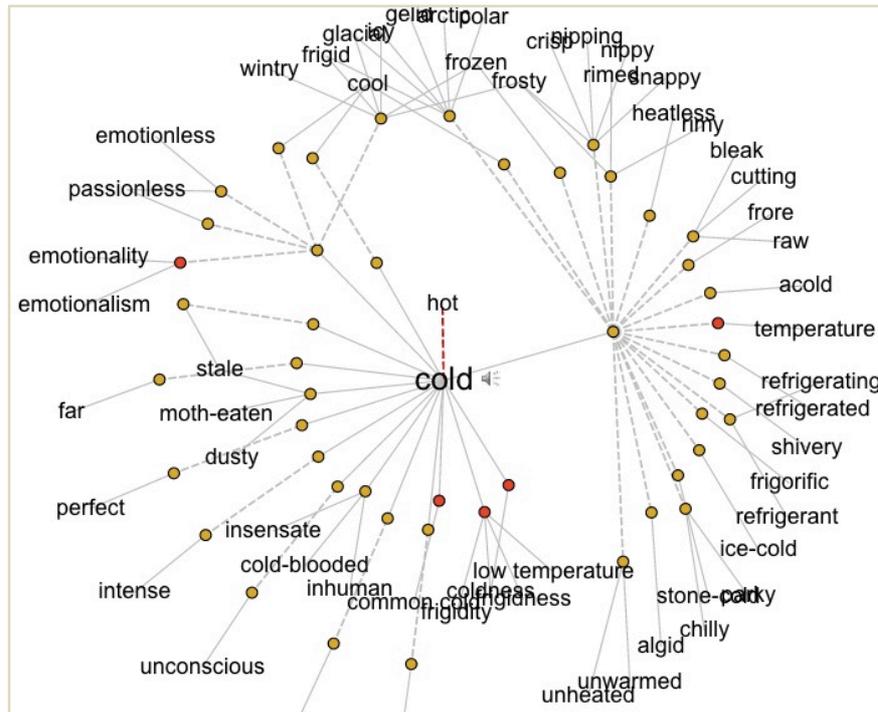


Figura 4: El concepto [COLD] según el Visual Thesaurus

En este sentido, si conocer una palabra implica tal cantidad de relaciones (fónicas, morfológicas, sintácticas, léxicas, semánticas, discursivas, socioculturales, intralingüísticas e interlingüísticas), para que los estudiantes incorporen las variantes léxicas de manera significativa debemos proponer actividades de aprendizaje similares a las que realiza nuestra mente en este complejo procesamiento. A la hora de planificar un manual de español, no tiene mucho sentido incluir un anexo de voces americanas en cada tema de manera aislada y fragmentada del proceso de aprendizaje. La adquisición de una palabra siempre estará relacionada con el esfuerzo cognitivo; para que sea significativo se requiere tiempo de exposición y la

activación de mecanismos de procesamiento que garanticen la fijación del significado y uso de cada unidad léxica.

El profesor de ELE debe proponer actividades que favorezcan asociaciones:

- a. morfológicas: ejercicios de derivación y composición
- b. semánticas: hiponimia / hiperonimia; meronimia / holonimia; sinonimia / antonimia; campos léxicos y asociativos; polisemia; metáfora, ironía
- c. léxicas: estudiar cognados, expresiones idiomáticas
- d. discursivas: trabajar con conectores textuales
- e. pragmáticas: tratamiento de la selección léxica según situación
- f. sintácticas: práctica con las colocaciones

Por otra parte, las actividades de trabajo léxico aisladas de cuestiones gramaticales tampoco tienen sentido. Cuando planificamos la incorporación de variedades léxicas en el currículo es importante recordar que el léxico se aprende siempre en contexto lingüístico y extralingüístico. Así, para trabajar con las unidades léxicas debemos ofrecer siempre una situación comunicativa (*contexto*) y otras unidades de referencia con las que se combine (*cotexto*). De hecho, el aprendizaje del estudiante variará en función de la calidad de los rasgos léxico-semánticos de la situación comunicativa que presentemos en las actividades. A partir de la inferencia que pueda extraer del contexto y cotexto de la unidad léxica, será capaz de especificar los rasgos semánticos de cada entrada y relacionarlos con los de otras palabras.

En este sentido, son interesantes los ejercicios rellena-huecos en un texto (*open cloze*), los puzzles de palabras, los *mind gaps* (redes de palabras ordenadas por temas), las transformaciones de frases, los ejercicios morfológicos que fomenten la distinción entre raíces y afijos o el estudio contrastivo de pares de palabras cercanas que podrían crear interferencias. En un contexto de aprendizaje de la variedad, las actividades que relacionen el léxico con la pronunciación, la ortografía y las

imágenes visuales también son imprescindibles, aunque sean poco frecuentes en los materiales de ELE.

Para la adquisición del léxico también es fundamental la práctica: cuanto más lento sea el procesamiento, mejor permanecerá la unidad léxica en nuestro lexicón. Se trata de un proceso cíclico por lo que el estudiante necesitará varias presentaciones de la unidad léxica, en actividades de distintos tipos, para consolidar la fijación de la misma en la memoria a largo plazo. Por ello, si introducimos las variantes léxicas de manera adyacente sin integrarlas en la organización del currículo, los estudiantes no incorporarán estas unidades de manera significativa.

La adquisición también debe estar orientada a mejorar las estrategias de aprendizaje e incorporar técnicas de repaso o reactivación. Los estudiantes deben aprender a incorporar las unidades léxicas en la memoria a largo plazo. Para ello, las variantes léxicas deben trabajarse con actividades de retención y memorización, aprendiendo reglas mnemotécnicas, ejercicios de descubrimiento de nuevas palabras o mediante la individualización del vocabulario (proceso que consiste en crear un glosario personalizado, con el léxico que considere relevante cada estudiante según sus propias necesidades de aprendizaje, añadiendo imágenes, definiciones, frases, etc.) Además de trabajar la adquisición del léxico activo y pasivo, también puede resultar muy útil trabajar el léxico en potencia, aprendiendo técnicas de decodificación y deducción, a partir del contexto y cotexto, o de su análisis morfológico, por ejemplo. Que los estudiantes aprendan estrategias para enfrentarse a la variación geolectal del español es fundamental ya que se trata de una característica básica de nuestra lengua.

Por último, el aprendizaje del léxico como proceso cíclico, exige una fase de recuperación así como un sistema de evaluación. El aprendizaje de vocabulario necesita mecanismos que garanticen la reutilización de las palabras aprendidas, tomando conciencia del vocabulario aprendido y favoreciendo su uso y mejora. Así, es importante tener en cuenta que retomar y reutilizar las variantes

dialectales aprendidas es una labor fundamental para fijar estas unidades en la mente de los estudiantes.

3. Conclusión

En este capítulo hemos presentado una descripción general del *PCIC* con sus características y antecedentes, examinando su estructura interna y poniendo de manifiesto su gran utilidad e importancia. Asimismo, hemos explicado las líneas de ampliación que desarrollaremos en este estudio: primero, vamos a incorporar términos panhispánicos que son comunes a una mayoría de regiones del mundo hispánico; a continuación, valoraremos la representatividad de voces americanas representativas equivalentes a las del *PCIC*, cuya incorporación es incuestionable si pretendemos dotar de una mirada plurinormativa a este material; por último, marcaremos aquellas palabras que son exclusivas del español de España. Hemos reflexionado también sobre el tratamiento de las variantes léxicas en la clase de español, teniendo en cuenta los mecanismos de incorporación de palabras al lexicón mental y las actividades didácticas que favorecen este proceso.

Esta tesis defiende la importancia de considerar la variación léxica en clase de ELE. Por ello, la ampliación de las «Nociones específicas» del *PCIC* con palabras representativas del español es un asunto fundamental para la investigación lingüística. Nuestra propuesta multiplica las aplicaciones y usos del *PCIC*, a través de un diálogo interdialectal de la lengua común que enriquece el caudal expresivo y cultural de nuestro idioma, a la vez que contribuye a fomentar una política panhispánica integradora de la variedad.

CAPÍTULO 2

GEOGRAFÍA LINGÜÍSTICA Y LÉXICO HISPÁNICO

En este capítulo vamos a delimitar el marco teórico de nuestra tesis. Para ello, revisaremos los estudios e investigaciones sobre variación lingüística, desde las primeras reflexiones sobre el cambio lingüístico hasta la implantación del modelo laboviano en el ámbito de la sociolingüística y la lingüística cuantitativa.

En primer lugar, analizamos la evolución histórica de la dialectología, desde sus orígenes hasta la actualidad. Después, presentamos los primeros atlas lingüísticos, que dan origen a la geografía lingüística, disciplina en la que se enmarca nuestro trabajo. Como punto de partida definimos los cuatro elementos geolingüísticos esenciales (cuestionario, red de encuesta, encuestadores e informantes), revisando a continuación las críticas y limitaciones de este campo de estudio.

Por último, nos centraremos en los estudios de variación léxica, puesto que nuestra investigación se centra en estas variantes. Tras enumerar los problemas y las limitaciones que caracterizan los estudios léxicos, presentamos las obras y fuentes de consulta a las que se puede acudir para conocer datos fiables de variación. Incluimos, para terminar, una clasificación ejemplificada de la distribución léxica que encontramos al comparar las voces españolas con otras alternativas americanas.

1. Dialectología y geografía lingüística

En los estudios lingüísticos encontramos dos aproximaciones teóricas diferenciadas: la concepción de la lengua como sistema uniforme, con sus aspectos sistemáticos y sincrónicos, y el estudio de la lengua consustancial a la variedad. Esta última es la que vertebra los estudios dialectales sobre variación lingüística, que sostienen que la fuerte diferenciación interna que caracteriza a las lenguas puede originar en última instancia los cambios lingüísticos. La diversificación interna de las lenguas les confiere una inherente complejidad así como una serie de rasgos de dialectalismo o variación. La lengua presenta, por tanto, variación dentro de su unidad; es lo que denominamos un *diasistema* o conjunto de dialectos, niveles y estilos de lenguaje (Coseriu, 1981). Existen diferentes rasgos lingüísticos que pueden estar sujetos a variación, ya sean fenómenos fonéticos, morfosintácticos, pragmáticos o léxicos. Estos últimos son la base de nuestra investigación.

Los estudios sobre variación se centran en rasgos lingüísticos específicos que se denominan *variables lingüísticas* —unidades que pueden manifestarse de formas distintas sin que exista un cambio de significado—, con sus diversas *variantes lingüísticas* —manifestaciones o realizaciones posibles de un fenómeno lingüístico.

La variación lingüística generalmente está ligada a factores extralingüísticos. Cuando la variación se produce por el paso del tiempo, nos referimos a *variaciones históricas*; cuando el contexto comunicativo actúa como factor determinante, nos referimos a *variaciones estilísticas* y cuando la variación se produce en función de un espacio geográfico determinado hablamos de *variación geolingüística* (Moreno Fernández, 1998). Esta tesis se centra en este último tipo de variación.

1.2. Dialectología: perspectiva histórica

El estudio de la variación lingüística nace ligado a la dialectología a finales del siglo XVIII, época en la que los lingüistas prestan atención a la historia y

evolución de las lenguas, que dio origen a los primeros estudios monográficos dialectales (Abad, 1993: 75). Aunque en siglos anteriores se encuentran referencias a cuestiones de índole dialectal¹, es hacia finales del siglo XVIII cuando los lingüistas comenzaron a estudiar los dialectos de modo constante (Pop, 1950). Así, el nacimiento de los estudios dialectales trae consigo la incipiente sistematización de un conjunto de variedades lingüísticas en el marco unitario de un sistema de lengua.

Sobre el estudio de la variación en el siglo XIX, podemos destacar las aportaciones de Schmidt, Schuchardt, Wenker o Rousselot, entre otros. En 1841, Biondelli publicó un *Atlas lingüístico de Europa*, que constituye el inicio de la cartografía lingüística, con el objetivo de buscar relaciones entre lenguas y familias de lenguas. Los mapas, desde entonces, se asociaron al estudio de la variación lingüística por su capacidad para evocar una impresión espacial de las lenguas.

Schmidt, en 1872, propuso su «Teoría de las ondas» para explicar que los cambios lingüísticos se originan no exclusivamente por genealogía sino desde varios centros, como las ondas de un lago, produciéndose en la formación de las lenguas diversas innovaciones que se superponen.

Por su parte, Schuchardt subrayó la importancia de estudiar el habla concreta de los hablantes puesto que toda lengua es una suma de variaciones, en función de la edad, el sexo o el nivel sociocultural de los hablantes, entre otros factores. En su obra *Vocalismus des Vulgärlateins* (1866–1868) resalta la dificultad de fijar los límites de los dialectos. Desde entonces toma fuerza la idea de que una lengua es un *continuum* en el que se entremezclan múltiples variaciones que son inherentes a todo lenguaje, de la que nos hablará, un siglo más tarde, Manuel Alvar (1969).

En 1873 Ascoli fundó el *Archivio glottologico italiano*, que contribuyó a impulsar la dialectología italiana. Por su estudio sistemático de los dialectos italianos se le ha considerado el padre de la dialectología como disciplina científica.

¹ Para ampliar más sobre las nociones en torno al cambio lingüístico anteriores al siglo XIX, véase: Navarro (1993).

Wenker, en 1876, lanzó una serie de encuestas por correspondencia para estudiar los dialectos alemanes, a partir de las que publicó más de 1500 mapas dibujados a mano. En su siguiente trabajo de 1878 diseñó los primeros mapas sintéticos fonéticos. Así, Wenker elaboró el primer atlas lingüístico, con una colección de mapas fonéticos en los que trazaba una serie de áreas lingüísticas, que podrían considerarse las primeras *isoglosas*. Podríamos considerarle, por tanto, el precursor de la geografía lingüística (Coseriu, 1955: 39).

En la última década del XIX, Rousselot en su obra *Modifications phonétiques du langage étudiées dans le patois d'une famille de Cellefrouin (Charente)* demostró que la unidad lingüística absoluta no existe ni siquiera en el seno de una misma familia. En este sentido, Rousselot puso de manifiesto que toda comunidad de hablantes es diversa y compleja y está sujeta a la diferenciación y a la interferencia de rasgos lingüísticos variados.

En la Península aparecen también algunos trabajos lingüísticos que muestran interés dialectal, como el estudio de Jovellanos *Apuntamiento sobre el dialecto de Asturias* (1858–1859). En Hispanoamérica, Esteban Pichardo es el promotor de la dialectología con su *Diccionario provincial casi razonado de voces y frases cubanas*, que constituye el primer trabajo lexicográfico dedicado exclusivamente al español americano. También es fundamental la figura de José Cuervo, que se centró en la variación lingüística en función del contexto comunicativo, tal y como apreciamos en sus *Apuntaciones críticas sobre el lenguaje bogotano* (1867–1872).

Pero no es hasta el siglo XX cuando encontramos la primera investigación directa y sistemática sobre variación lingüística. Edmond Edmont, que no era lingüista ni dialectólogo, accedió a colaborar con Gilliéron para elaborar el proyecto del *Atlas Linguistique de la France* en todo el territorio francés. A partir de entonces nació un nuevo método de estudio de la variación dialectal, la geografía lingüística, que estudiaremos en detalle en la siguiente sección.

En España, no nace verdaderamente la dialectología hasta 1903, fecha en que Menéndez Pidal expone en una carta dirigida a Unamuno su interés por realizar dos

monografías exhaustivas del leonés y el aragonés, con el objetivo de trazar posteriormente una historia de la lengua española. Desde entonces, a lo largo de todo el siglo XX, se desarrollaron multitud de estudios sobre variación lingüística del español y de otras lenguas peninsulares, elaborándose gran cantidad de atlas lingüísticos que comentaremos más adelante.

A mediados del siglo XX, la dialectología sufrió cierto estancamiento ante los nuevos planteamientos de la lingüística teórica: los lingüistas cuestionan a los dialectólogos por divagar sobre fenómenos lingüísticos concretos, sin contemplar el sistema global de la lengua, al mismo tiempo que los dialectólogos desprecian a los lingüistas por ser demasiados abstraccionistas y olvidar los datos reales de la lengua. Desde el estructuralismo asistimos a una exclusión de la dialectología como disciplina científica así como a grandes problemas teóricos, derivados del generativismo, para establecer su marco de estudio. Por otra parte, el hecho de que la recogida de datos dialectales se llevara a cabo a lo largo de muchos años, se interpretó como un trabajo con datos desactualizados o poco fiables.

Sin embargo, fue hacia los años 60 cuando se produjo un cambio de perspectiva en los estudios de variación. Además de los factores lingüísticos, los factores sociales comienzan a tenerse en cuenta en los análisis de la lengua (Rotaetxe, 1988). Nació entonces un nuevo enfoque en el estudio de la variación que dio lugar a la sociolingüística, disciplina preocupada por explicar la variación lingüística según factores sociales y situacionales.

Desde los años 60, el estudio de la variación ha estado ligado a la denominada *lingüística cuantitativa*. Nazar (2008) propone referencias muy anteriores sobre los conatos de esta disciplina, atribuyendo los inicios de la lingüística cuantitativa a la publicación de un trabajo de Markov en 1913, donde se analizan las probabilidades de aparición de las distintas secuencias de letras en el poema *Eugene Onegin* de Pushkin, dando a conocer las *cadena de Markov*, que todavía se utilizan en distintas disciplinas y que fueron la base de la posterior «Teoría de la información» de Shannon (2001). Nazar (2008) cita otros estudios pioneros

sobre aspectos cuantitativos del lenguaje como los de Mosteller y Wallace en 1963 sobre la autoría de los *Federalist Papers* o el análisis estadístico textual de Muller sobre la obra de Corneille. En relación con la lingüística histórica, también se aplicó un modelo léxico–estadístico con el objetivo de determinar el grado de relación entre el vasco y otras lenguas caucásicas (Tovar, 1968). En otro plano, Chomsky y Harris fueron quienes apostaron definitivamente por una aplicación directa de las matemáticas a la lingüística, lo que dio lugar a posterior gramática generativa y transformativa.

A partir de los años 70 —y sobre todo, a partir de los años 90— la estadística ocupó un lugar fundamental en multitud de estudios, para la extracción de léxico bilingüe, el procesamiento de lenguaje natural, las probabilidades sintácticas o léxicas, la generación automática de tesauros o la traducción automática, entre otras aplicaciones².

En los trabajos sobre variación lingüística, la estadística se ha empleado de manera sistemática para el tratamiento y análisis de los datos. De hecho, el estudio de la variación se consolidó como modelo teórico explicativo con el planteamiento probabilístico laboviano. Los análisis de probabilidad de Labov en los años 60 se centraron en analizar la variación relacionada con factores sociales, lo que dio lugar a la sociolingüística. Este tipo de análisis variacionista proporciona datos rigurosos basados en el cálculo de la probabilidad de que aparezca un rasgo lingüístico en unas circunstancias lingüísticas, sociológicas o contextuales específicas (Moreno Fernández, 1994). Así, la sociolingüística se enmarca en la llamada lingüística cuantitativa, con el objetivo de descubrir regularidades estructurales en la lengua que puedan ser objeto de una formalización matemática (Nazar, 2008: 1).

Lógicamente el modelo matemático aplicado a la lingüística tiene importantes ventajas, como la posibilidad de realizar generalizaciones, predicciones y de formalizar las reglas. En este sentido, hay un punto de inflexión en el estudio de la

² Nazar (2008) nos remite a una excelente explicación sobre la aplicación de los métodos estadísticos en el Procesamiento del Lenguaje Natural en Manning y Schütze (1999).

variación a partir del tratamiento informático de los datos, con la lingüística de corpus y la generalización de Internet. Es fundamental, por ejemplo, la importancia del programa estadístico informatizado VARBRUL que ha facilitado el análisis estadístico a los lingüistas³.

Todos estos avances no hacen que los estudios actuales sobre variación lingüística sean radicalmente distintos a los estudios tradicionales. Los estudios sobre variación se nutren, todavía hoy, de investigaciones dialectológicas tradicionales. Así, «la sociolingüística es hija de su tiempo y continuadora de unas experiencias acumuladas, principalmente, durante los últimos cincuenta años» (Moreno Fernández, 1994: 95). Pero la lingüística cuantitativa actual permite explorar fenómenos lingüísticos desde otro ángulo gracias a sus particularidades, entre las que destacamos: el registro de los datos de un gran número de hablantes, su posterior comprobación en términos de representatividad a través de corpus informatizados, el estudio del habla en centros urbanos, el análisis de contextos de aparición (no solo en términos de presencia/ausencia sino de coexistencia), el procesamiento de datos con programas estadísticos que permitan indagar sobre motivos lingüísticos y/o sociales que propician la variación o la posibilidad de hacer predicciones sobre la distribución de fenómenos lingüísticos.

En todo caso, cabe preguntarse si el modelo cuantitativo es suficiente para interpretar la variación lingüística. Algunos teóricos cuestionan que las variables sociales sean suficientes para explicar la variación, señalando que los cambios se producen sobre la base de estrategias pragmáticas. En todo caso, es evidente que el estudio de la variación lingüística no puede limitarse a exponer resultados cuantitativos:

Si el objetivo es intentar explicar el porqué de la variación y determinar los esquemas de covariación de un determinado fenómeno lingüístico con factores internos y/o externos, la necesidad de realizar un análisis complejo se hace evidente (Bentivoglio, 2001).

³ Tal y como nos recuerda Bustos Gisbert (1995), no podemos obviar que también existe entre los lingüistas cierto escepticismo ante el método probabilístico en relación con el tipo de datos que permite estudiar y con el alcance de sus explicaciones.

Aún así, hoy en día es indudable que el método cuantitativo aplicado al estudio de la variación lingüística (Labov, 1983; López Morales, 1983) reporta unos beneficiosos inconmensurables para los estudios lingüísticos⁴.

1.3. Geografía lingüística

Aunque todo estudio sobre variación lingüística se relaciona con el campo de la dialectología, el análisis de la variedad y la variación de la lengua se puede formular desde diversos métodos y modelos de investigación. Según Gimeno Menéndez (1990), las diferentes posibilidades de aproximarse al análisis han dado lugar a diversas disciplinas como la dialectología tradicional, la dialectología histórica, la dialectología estructural, la dialectología social, la dialectología transformativa, la sociolingüística, la sociolingüística histórica o la geografía lingüística, en la que se enmarca nuestro trabajo.

El problema fundamental al que nos enfrentamos es el aspecto multidimensional de la lengua, que hace difícil analizar un acto comunicativo desde la perspectiva de una única disciplina. Así, los mismos procesos de variación lingüística pueden ser objeto de análisis dentro de diferentes ámbitos, como parte de la geografía lingüística, de la lingüística histórica, de la sociolingüística o de la pragmática, y cada una de estas disciplinas establecerá un determinado modelo de análisis de acuerdo con sus fundamentos teóricos y exigencias empíricas. Moreno Fernández (1990) explica que quizá sería conveniente definir una *lingüística del habla* que permita dar cobertura teórica y metodológica a la existencia de múltiples modelos descriptivos válidos según la posición teórica que adoptemos: la geográfica, la social, la histórica o la comunicativa.

Si bien la dialectología no dispone de un método específico para recoger y presentar sus materiales, la geografía lingüística surgió del empleo de una

⁴ Otro ejemplo fundamental de estas aplicaciones lo encontramos precisamente en el ámbito de la geografía lingüística: los mapas lingüísticos combinados con una serie de formulaciones complejas matemáticas han dado origen a la dialectometría, disciplina que mide las distancias dialectales ayudándose con fórmulas taxonómicas aplicadas informáticamente.

metodología estricta para el estudio de los fenómenos de variación, mediante encuestas directas y unitarias a lo largo de una red de puntos geográficos específicos, previamente seleccionados, con una presentación cartografiada de los resultados. Frente a la exhaustividad y precisión de los estudios tradicionales de los dialectos, la geografía lingüística pierde inevitablemente en profundidad pero arroja una panorámica clara, rápida y precisa de la distribución de los cambios lingüísticos que tiene gran interés para la ciencia lingüística. Evidentemente, los exhaustivos glosarios y textos que nacen del estudio de los dialectos son un complemento necesario a los mapas de la geografía lingüística, cuyo interés es delimitar la extensión y la distribución geográfica de los cambios lingüísticos. Así, ambas disciplinas son tan diferentes como complementarias.

Podemos definir la geografía lingüística como el «conjunto de métodos para recopilar de un modo sistemático los testimonios de las diferencias dialectales» (Chambers y Trudgill, 1994: 37). En otras palabras, se trata de un método de investigación que reúne la variación lingüística a través de una metodología específica previamente diseñada y representa el estado sincrónico de uno o más elementos lingüísticos. García Mouton destaca que esta disciplina no termina su labor en el cartografiado de mapas sino que debe dar un paso más:

Se llama Geografía Lingüística a todo el proceso que va desde la recogida de materiales por medio de encuestas hasta su cartografiado de mapas; aunque hay que insistir en que la verdadera Geografía Lingüística cubre una etapa más, la del estudio de interpretar la información que contienen esos mapas (García Mouton, 2007: 342).

1.3.1. Geografía lingüística en Europa

La geografía lingüística —que algunos autores han denominado *lingüística geográfica* o *geolingüística*— se consolidó como disciplina científica con la elaboración de los primeros atlas, que incluyen una serie de mapas sintéticos o analíticos que ofrecen la visión espacial simultánea de un mismo fenómeno lingüístico.

A principios del siglo XX, Gilliéron publicó el *Atlas Linguistique de la France* (ALF), que ya hemos citado, en el que estudiaba 1900 preguntas sobre fenómenos lingüísticos distribuidos a lo largo de más de 600 localidades de toda Francia. A través de las encuestas realizadas por Edmont a un único hablante representativo de cada localidad que no tuviera estudios y representara el habla dialectal, se procede a la elaboración de mapas diferentes para cada fenómeno lingüístico y a su posterior estudio e interpretación. Gilliéron no se centra exclusivamente en registrar fenómenos fonéticos sino que también registra la variación morfológica y léxica.

Nace así la geografía lingüística, disciplina en la que se enmarca nuestra tesis doctoral. El estudio y la interpretación de los mapas permitieron observar la migración de las palabras, los centros de prestigio o de irradiación, las regiones aisladas, las estratificaciones lingüísticas de diferentes épocas, las innovaciones, etc. En este sentido, es fundamental la obra posterior de Jud (1914–1915), que estudió la superposición de estratos de palabras mediante la llamada *estratigrafía lingüística*, demostrando la complementariedad y la conciliación de los estudios geográfico–lingüísticos e históricos.

Tras el éxito del ALF germinaron nuevos proyectos de atlas lingüísticos. Así, Jaberg y Jud publicaron el atlas italo-suizo *Sprach– und Sachatlas Italiens und der Südschweiz* (AIS), que resulta fundamental puesto que en él se plantean novedades metodológicas, como la utilización de cuestionarios diferentes según la localidad, la incorporación de las ciudades como núcleos de estudio de la variedad o las encuestas a personas de diferentes edades o clases sociales. Además, este atlas rompe con la idea de trabajar con un investigador no lingüista e incorpora el modelo etnográfico para relacionar las palabras y las cosas designadas (en el marco del movimiento *Wörter und Sachen*), incluyendo en sus encuestas casi dos mil dibujos y fotografías.

La orientación del AIS la sigue también el *Atlante linguistico–etnografico italiano della Corsica*, proyecto realizado por Bottiglioni. Este trabajo incluía notas y

observaciones así como la traducción de las frases de encuesta al alemán, al francés y al inglés.

Por otra parte, se desarrolló el proyecto del *Atlas lingüístico rumano* (ALR), de gran perfección metodológica, aunque se vio interrumpido por la guerra. En Italia, el *Atlas lingüístico italiano* (ALI), de Bartoli, Bertoni y Vidossi, presenta un formidable cuestionario y una espesa red de puntos, aunque el proyecto también sufrió una fuerte parada con el fallecimiento de su investigador principal, Pellis, hasta que se consolidó el nuevo equipo de investigación.

Todos estas obras pueden considerarse *atlas de gran dominio* (Jaberg, 1955), puesto que su propósito es aportar una imagen lingüística visual de territorios muy extensos. Las únicas lenguas romances que no han publicado sus atlas lingüísticos son el español y el portugués. Paralelamente a los atlas de gran dominio comenzaron a realizarse *atlas de pequeño dominio*, centrados en el estudio dialectal exhaustivo de las regiones⁵.

En España, el primer proyecto de atlas fue precisamente un atlas de gran dominio, el *Atlas Lingüístico de la Península Ibérica* (ALPI), cuyo mentor fue Menéndez Pidal en 1910 y que durante décadas dirigió su discípulo Navarro Tomás. Se trata del proyecto más importante de geografía lingüística española y movilizó a filólogos de diferentes dominios lingüísticos peninsulares. La trágica historia de este proyecto se debe a su interrupción por la Guerra Civil cuando el Centro de Estudios Históricos se disolvió, provocando el exilio, la dispersión y la separación de los investigadores. Para proteger los cuadernos del ALPI, Navarro Tomás los trasladó a América y, aunque el lingüista nunca volvió a territorio español, en 1951 los envió de vuelta para asegurar la continuación del proyecto.

⁵ Con esta nueva metodología, Dauzat coordinó una revisión del *ALF*, el *Nouvel Atlas linguistique de la France* (NALF), que permitió conocer con mayor rigor la transición entre estructuras lingüísticas y la distribución de la variación en las áreas de estudio.

A través de 1320 preguntas, el objetivo del ALPI era estudiar el componente fonético y morfosintáctico por un lado, y el componente léxico–etnográfico por otro. Así, durante décadas el proyecto se centró en el estudio del habla de 528 localidades rurales. Lamentablemente hasta el momento solo se ha publicado un tomo de este ingente trabajo pero esperamos que las iniciativas del filólogo canadiense David Heap por digitalizar los mapas y ponerlos al servicio de la comunidad investigadora, hagan posible el acceso y la divulgación de estos valiosísimos materiales, tarea para la cual también colabora el CSIC.

A pesar de la historia truncada del ALPI, en España se fueron elaborando otros atlas regionales y etnográficos de pequeño dominio, dirigidos todos ellos por Manuel Alvar, lo que les otorga cierta homogeneidad. En los años 60, se publicó el *Atlas Lingüístico y Etnográfico de Andalucía* (ALEA), que constituyó la referencia para los atlas regionales posteriores. Del ALEA derivó el *Atlas Lingüístico y Etnográfico de las Islas Canarias* (ALEICan), planteado entre 1975 y 1978, que permitió observar la vinculación con Andalucía y la norma sevillana y con el español de América. Entre 1978 y 1983 se publicó el *Atlas Lingüístico y Etnográfico de Aragón, Navarra y La Rioja* (ALEANR), que constató el desvanecimiento del aragonés. En 1995 se publicó el *Atlas Lingüístico y Etnográfico de Cantabria* (ALECant), que estudia un territorio de sumo interés por ser colindante con el vascuence y el asturleonés.

Otros atlas en proceso de elaboración son el *Atlas Lingüístico (y etnográfico) de Castilla La Mancha* (ALECMan), dirigido por García Mouton y Moreno Fernández y del que ya existen bastantes materiales publicados en una zona que no estaba registrada entre los proyectos de Alvar, o el *Atlas Lingüístico de Castilla y León* (ALCyL), dirigido por Alvar pero con una metodología distinta al resto de atlas lingüísticos.

Además de los atlas regionales, también son interesantes los proyectos de *macroatlas* internacionales como el *Atlas linguarum Europae* (ALE), atlas de cuarta generación que reelabora materiales ya recogidos. El proyecto fue lanzado en 1970

por la UNESCO y cubre seis familias de lenguas en Europa en 2631 localidades. También destacan el *Atlas lingüístico de España y Portugal*, diseñado por Alvar en 1974, y el *Atlas linguistique Romane* (ALiR), en el que participan más 30 universidades y centros de investigación⁶.

1.3.2. Geografía lingüística en Hispanoamérica

El estudio del español en América destaca por sus importantes aportaciones en relación a las influencias que ejercen las lenguas en contacto. Las primeras investigaciones sobre lenguas indígenas desde principios del siglo XX vinieron acompañadas del interés por la hipótesis del sustrato lingüístico. En la actualidad, sabemos que la influencia de las lenguas amerindias se advierte considerablemente en los indígenas que aprenden español, más que en sentido contrario, pero a lo largo de las décadas han sido muchos los estudios en los que se ha afirmado la influencia sustratística de las lenguas indígenas sobre el español hablado en América. Lope Blanch (1995) insiste en que ya en 1947 Malmberg expuso una serie de factores que determinan la interferencia lingüística, según los cuales podemos afirmar que, en la mayoría de los casos, no se cumple la hipótesis del sustrato entre las lenguas indígenas y el español. De hecho, se ha estudiado detalladamente lo contrario: cómo el español ha influido en las lenguas indígenas ocasionando su modificación o, incluso, su desaparición.

En relación con el contacto de lenguas, abundan las investigaciones sobre el español en contacto con el inglés (Ramírez, 1992; López Morales, 1992) o con el portugués, como se refleja en el *Atlas Lingüístico Diatópico y Diastrático del Uruguay*, que comentaremos más adelante. Bustos Gisbert (1995) nos recuerda que también tienen gran interés las situaciones lingüísticas de *continuum* español–inglés que manifiestan cambios o mezclas de códigos, como las descritas por Valdés (1979).

⁶ Para más información sobre el ALiR, véase García Mouton (2003).

Pero la geografía lingüística como tal llegó a Hispanoamérica en 1948 de la mano del exiliado español Navarro Tomás, que publicó más de 70 mapas sobre el habla de Puerto Rico. Destaca, además, su *Cuestionario lingüístico hispanoamericano* (1943), donde se plantea la cuestión de cómo coordinar los diferentes estudios sobre variación lingüística en toda Hispanoamérica.

A partir de los años 70 se publicaron nuevos atlas, entre los que destacan: el *Atlas Lingüístico–Etnográfico del Sur de Chile* (ALESuCH) o el *Atlas Lingüístico–Etnográfico de Colombia* (ALC). Pero a partir de los años 90 se produce la explosión de los estudios geolingüísticos hispanoamericanos. Entre ellos destacamos dos: el *Atlas Lingüístico de México*, dirigido por Lope Blanch, y el *Atlas Lingüístico Diatópico y Diastrático del Uruguay*, dirigido por H. Thun y A. Elizaincín.

El *Atlas Lingüístico de México* tiene unas dimensiones monumentales y abarca datos de 193 localidades, con más de 1300 informantes. Incluye 952 mapas y las respuestas indican el índice de frecuencia. Para algunos dialectólogos, plantea ciertos problemas metodológicos:

Es tanta la información que se proporciona, que los mapas llegan a perder su fuerza visual porque exigen un trabajo enorme de desentrañamiento. No sé si algo así se realizará algún día, pero toda esa información presentaría su más alto grado de utilidad si pudiera informatizarse y consultarse a través de sistemas de bases de datos (Moreno Fernández, 2005: 94).

El *Atlas Lingüístico Diatópico y Diastrático del Uruguay* es especialmente interesante por el alto nivel de análisis e interpretación de los materiales publicados en los mapas. Además de ofrecer información sociolingüística, presenta el contraste con el portugués de Brasil, lo que permite estudiar las influencias entre las dos lenguas colindantes. El aspecto cartográfico es muy diferente al de los atlas anteriores, por la pluralidad de informantes y las diferencias recogidas, lo que produce una acumulación de datos en el cartografiado que obliga a establecer distintos tipos de mapa: mapas fenotípicos, mapas monofásicos monosexuales, mapas plurifásicos, etc. Según García Mouton, son mapas difíciles de leer, por lo que «en algunas ocasiones surge la pregunta de si tiene verdaderamente sentido

cartografiar, porque el texto y los diagramas resultan más ilustrativos» (García Mouton, 2006: 121).

Como señala Moreno Fernández y Otero (2008), junto a estos dos grandes atlas, existen varios proyectos que también son destacables, como el *Atlas Lingüístico de Nuevo México*, dirigido por Neddy A. Vigil y Garland D. Bills; el *Atlas Lingüístico de Cuba*, dirigido por García Riverón y Choy; el *Atlas Lingüístico–Etnográfico de Costa Rica*, dirigido por Miguel Ángel Quesada; el *Atlas Lingüístico de Antioquia*, de Osorio, Betancour, García y Zuluaga; el *Atlas Lingüístico de Cuyo*, dirigido por César Quiroga, y el *Atlas Lingüístico Guaraní–Románico*, dirigido desde Alemania por Wolf Dietrich, Harald Thun y Haralambos Symeonidis:

Todos estos atlas constituyen bases de datos privilegiadas, con datos recibidos de los hablantes, que conforman una parte fundamental del patrimonio lingüístico del español de América, y cuyas posibilidades de uso en el futuro resultan todavía insospechadas. Quizá no esté tan lejano el día en el que los dialectólogos americanos se reúnan para hacer, a partir de ellos, una elaboración conjunta que represente el español de América en sus variedades (García Mouton, 2006: 122).

En América también surgió un proyecto de macroatlas del español, el *Atlas Lingüístico de Hispanoamérica*, dirigido por M. Alvar y A. Quilis. En 1984 publicaron el cuestionario y la filosofía del proyecto, sin pretensiones de exhaustividad y con colaboradores americanos: «Lo que se pretende no es profundizar en cuanto se sabe, sino aprender lo que se ignora y vincular las mil partes de ese gran todo» (Alvar y Quilis, 1984: 15). En el año 2000, Quilis publicó la última relación actualizada del estado de las encuestas en un volumen monográfico de la revista *Géolinguistique*, donde expuso los avances de la geolingüística en América Latina y el estado del atlas lingüístico de Hispanoamérica. Esperamos que la culminación de un proyecto tan notable permita ofrecer una imagen global de la distribución geográfica del español, que será esencial para los estudios lingüísticos.

1.3.3. La metodología de la geografía lingüística

Como ya hemos apuntado, la geografía lingüística hace referencia a un método «dialectológico y comparativo (...) que presupone el registro en mapas

especiales de un número relativamente elevado de formas lingüísticas (fónicas, morfológicas, sintácticas o léxicas), mediante encuesta directa y unitaria en una red de puntos de un territorio determinado» (Gimeno Méndez, 1990: 117). Se trata, por tanto, de estudiar la difusión y extensión de los rasgos lingüísticos en un área geográfica concreta.

La consideración de la geografía lingüística como método, más que como disciplina autónoma, no impide que se haya desarrollado tan ampliamente que hoy puede considerarse un campo de estudio particular dentro del conjunto de las ciencias lingüísticas. Desde los estudios de Gilliéron, que ya hemos comentado, se han desarrollado nuevos planteamientos teóricos respecto a los límites dialectales y la difusión de los fenómenos lingüísticos. Afirmamos, por tanto, que la geografía lingüística se caracteriza, sobre todo, por presentar una metodología estricta y definida de antemano.

En todo estudio geolingüístico existe primero una etapa de preparación en la que se delimitan los siguientes aspectos: la *red de encuesta*, cuyos puntos de encuesta tendrán diferente densidad según el tipo de estudio o el número de habitantes, entre otros factores; el *tipo de encuesta*, oral, mediante transcripciones, grabaciones o escrita; el *cuestionario*, con diferentes modelos preguntas: interrogaciones directas, interrogaciones indirectas, definiciones, equivalencias o geosinónimos, completar espacios en blanco, dibujos, gráficos o, incluso, imágenes; y los *informantes*, aunque tradicionalmente se prefería a un hombre rural mayor y sedentario, poco a poco se introdujeron otro tipo de informantes, incluyendo a las mujeres y el entorno urbano, lo que permitió observar aspectos de variación relacionados con el sexo, la edad, la clase social, la dedicación profesional, etc.

Cuando se ha preparado el estudio, la siguiente fase se centra en la recolección de los datos. Suele tratarse de un proceso dilatado en el tiempo en el que tradicionalmente colaboraba un único entrevistador o encuestador pero poco a poco se prefirió incluir a varios colaboradores formados metodológicamente para realizar

con rigurosidad el trabajo, lo que permitía ahorrar tiempo y emprender proyectos de mayor envergadura.

Una vez que se recogen todos los datos, se procede a la publicación de los resultados cartografiados en mapas de distintos tipos. Existen mapas *lingüísticos*, que recogen información fonética, gramatical o léxica, y mapas *etnográficos*. También pueden confeccionarse mapas *interpretativos*, que representan la distribución de los fenómenos lingüísticos, permitiendo delimitar las fronteras de diversas áreas dialectales a través de isoglosas; y los mapas *sintéticos*, que reflejen la frecuencia de aparición del fenómeno lingüístico en cuestión. En una última fase se estudian, se interpretan o se reanalizan los datos, según sea el propósito de la investigación.

1.3.4. Críticas y limitaciones de la geografía lingüística

Es indudable que la geografía lingüística aportó al ámbito de la dialectología una serie de contribuciones fundamentales: el hecho de acudir directamente a los informantes de una zona, seleccionarlos, transcribir sus respuestas, trabajar con un mismo cuestionario en distintos puntos de encuesta, la elaboración de dicho cuestionario a partir de unos objetivos específicos de investigación o la idea de lengua como *continuum* en las diferentes áreas, por ejemplo. No obstante, las críticas a esta disciplina, sobre todo en su etapa incipiente, también fueron notables.

En primer lugar, los atlas lingüísticos fueron criticados por su aparente superficialidad, en contraste con los glosarios y monografías que elaboraban los dialectólogos. La geografía lingüística pierde en profundidad y exhaustividad frente a los estudios dialectales, pero ofrece otras claves fundamentales que permiten llegar a conclusiones lingüísticas de gran interés, señalando áreas conservadoras e innovadoras, relacionando distintos mapas o centrándose en la distribución geográfica de un fenómeno específico. Como ya hemos apuntado, hoy en día nadie duda de que los objetivos de ambas disciplinas son tan diferentes como complementarios.

Otras críticas que se han hecho a la geografía lingüística son, entre otras, la limitación de los mapas a un momento temporal concreto; la necesaria selección de unos puntos determinados para el estudio que no representan la totalidad del territorio; el contexto formal de realización de las encuestas, que puede condicionar las respuestas de los informantes; la imposibilidad de reflejar todo el léxico, toda la fonética o toda la sintaxis de un territorio; las entrevistas a informantes rurales, que son representativos únicamente de un estrato social; la metodología pregunta–respuesta, que suele obtener una única respuesta perdiéndose el contraste de diferentes datos, etc.

De entre todas las críticas que pueden hacerse a esta disciplina, podemos afirmar que cualquier investigación lingüística tiene muchas de estas limitaciones. García Mouton (1996: 76) se pregunta: «¿Qué investigación se ocupa de *todos* estos fenómenos?». Las críticas habría que encuadrarlas en el trabajo de investigación de cada estudioso, según sea su propósito de conocimiento. Además, algunas de esas limitaciones se pueden salvar incorporando hablantes de distintos niveles sociales, elaborando encuestas múltiples para obtener varias respuestas, favoreciendo un contexto de naturalidad para que el informante olvide que se encuentra en una situación formal, etc. Lo que está claro es que la geografía lingüística no pretende en modo alguno hacer una monografía en cada punto de encuesta. No se trata de una compilación de estudios dialectales ni de elaborar una suma exhaustiva de monografías. La geografía lingüística pretende realizar un estudio lingüístico en un momento del tiempo, a través de unos puntos geográficos determinados, y presentarlos a través de mapas con el objetivo de llegar a lecturas, interpretaciones o reelaboraciones que resulten sustanciales.

2. El estudio de la variación léxica

El componente *fonético–fonológico* es el que mejor se ha estudiado e interpretado en los estudios de variación. La facilidad para demostrar la alternancia de los elementos del sistema, su alta frecuencia en el habla, el hecho de tratarse de un

grupo cerrado de elementos y su correlación con factores sociales, hacen que las unidades fonológicas sean las más adecuadas para un análisis de variación.

La variación *gramatical* o *morfosintáctica* también viene determinada por factores lingüísticos y sociales, aunque su análisis resulta más complejo que el de la variación fonológica en cuanto a establecer el grado de equivalencia entre variantes (Silva–Corvalán, 2001). Hasta la década de los años 70 no se produce un verdadero interés por los estudios variacionistas sobre sintaxis. Lavandera (1978) destaca que estas primeras investigaciones generaron cierto escepticismo teórico sobre la existencia de variación en este plano lingüístico, lo que no impidió que se reforzaran en los años 90, incorporando las herramientas del análisis del discurso y de la pragmática:

[T]he difficulties involved in dealing with meaningful units have constituted a challenge and have placed sociolinguistics in a closer relationship with studies of the semantics and pragmatics of discourse (Silva Corvalán, 1991: 122).

El plano lingüístico sobre el que se han realizado menos análisis de variación es el *léxico*. A la hora de estudiar las alternancias entre palabras, uno de los principales problemas del análisis léxico radica en la existencia o no de voces equivalentes o sinónimas. Con razón, Baldinger (1970) afirmó que la sinonimia era «el plato fuerte de la semántica».

Son varios los lingüistas que niegan la existencia de la sinonimia absoluta, puesto que en muy pocos casos dos unidades léxicas son conmutables en todos los contextos. Otros teóricos, en cambio, defienden la sinonimia con vehemencia. Estas discrepancias surgen dependiendo de la concepción teórica en que nos basemos —sinonimia total o sinonimia parcial— y dependen de si tenemos en cuenta o no los valores connotativos y del contexto.

El primer teórico que niega la sinonimia es Bréal, en 1884, con su *ley de repartición del significado*, según la cual los términos que se consideran sinónimos tienden a diferenciarse semánticamente, repartiendo algunas de sus acepciones o connotaciones. Así, el fenómeno de la sinonimia se complica puesto que muchos términos tienen un carácter polisémico. Lo que sucede es que, habitualmente, una

palabra se puede intercambiar por otra *en alguna de sus acepciones* comunes (Martín Camacho, 2008).

Encontrar casos de sinonimia absoluta es infrecuente; generalmente, alguno de los términos es más general que el otro, o más intenso, más emotivo, más profesional, más coloquial, más literario, más infantil, implica una percepción moral o está determinado por el área dialectal (Salvador, 1985). La sinonimia, por tanto, en sentido riguroso, para muchos autores no existe, puesto que siempre observamos algún valor diferencial —social, emotivo, estilístico— que distingue, aunque sea ligeramente, un par de palabras (Berruto, 1979). Desde esta perspectiva, la variación «siempre es funcional, pues la producción lingüística es intención, estrategia y reflexión; está condicionada y condiciona la interpretación del locutor» (Serrano Montesinos, 1999: 12).

Otros lingüistas, en cambio, consideran que sí podemos hallar voces que presentan una sinonimia completa, es decir, significados exactamente iguales como observamos en *malaria–paludismo* o *marido–esposo* (Gutiérrez Ordóñez, 1996). En todo caso, este tipo de sinonimia absoluta sería más frecuente en el léxico científico.

En el estudio de la variación léxica, otro problema complejo es la dificultad de localizar las variantes léxicas en el discurso natural. Por este motivo, a pesar de los inconvenientes que presenta la utilización de cuestionarios fijos —que condicionan la fiabilidad de los resultados según la *paradoja del observador* de Labov—, en este tipo de estudios sobre variación léxica resulta esencial recurrir a encuestas.

Por último, otra de las preocupaciones de este tipo de estudios está relacionada con la frecuente alternancia entre palabras simples y unidades fraseológicas, o entre diversas unidades fraseológicas entre las que existe una clara relación de sinonimia (Penadés Martínez, 1999) como, por ejemplo, estas expresiones con el significado de ‘dormir mucho’: *dormir como un lirón–dormir como un bendito–dormir a pierna suelta...*

2.1. Fuentes de consulta y documentación

Una vez expuestas las limitaciones que podemos encontrar en el estudio de la variación léxica la primera pregunta que nos hacemos es a qué fuentes de consulta y documentación podemos acudir para conocer la variedad léxica del español. No encontramos una respuesta fácil. Colón (2002) advierte que no existe ninguna historia completa del léxico hispánico.

Las diferencias léxicas son las más notables cuando comparamos el español de América y el español peninsular, o cuando contrastamos el español hablado en las distintas regiones hispanoamericanas. De hecho, el estudio del léxico ha sido constante en la dialectología, con el objetivo de elaborar diccionarios, monografías, vocabularios o repertorios dialectales que den cuenta de la variedad. No obstante, el léxico representa un porcentaje de diferenciación muy inferior a la fonética o a la morfosintaxis, en el sentido de que sus variantes no cambian la estructura del sistema. Las palabras migran con facilidad entre diferentes regiones sin afectar al sistema lingüístico, adoptando valores de uso —positivos o negativos— que condicionan inevitablemente su utilización.

Es cierto que existen multitud de obras y materiales que recogen la variación léxica pero muchas de ellas no obedecen a investigaciones puramente lingüísticas ni a criterios sistemáticos sino que presentan repertorios folclóricos o anecdóticos. De hecho, algunas de estas obras parten de una intención práctica, de tipo enciclopédico, más que de una preocupación metalingüística (Fajardo Aguirre, 2010). Nos referimos, por ejemplo, a manuales de especialidad o a estudios etnográficos (Haensch, 1986).

En términos generales, los especialistas podemos consultar las siguientes fuentes sobre variedad léxica: atlas lingüísticos, monografías, artículos científicos, estudios de especialización, diccionarios, repertorios bibliográficos, corpus y bases de datos.

2.1.1. Atlas lingüísticos

Los atlas lingüísticos son fundamentales para mostrar la variación léxica puesto que dan cuenta, a través de mapas, de cómo un mismo concepto recibe diferentes denominaciones en un mismo territorio. A pesar de su volumen y del difícil manejo práctico, los atlas lingüísticos regionales y peninsulares poseen un enorme valor para: registrar voces antes desconocidas, registrar nuevas acepciones de palabras, localizar geográficamente las voces léxicas, fijar o comprobar etimologías teniendo en cuenta la localización geográfica de los términos y observar cuestiones de adstrato (Alvar, 1980: 56).

Es cierto que los atlas lingüísticos son «grandes, pesados, difíciles de examinar. Requieren agilidad para su manejo y han de consultarse de pie, abiertos sobre largas mesas despejadas. Pero eso no exime de su utilización» (Salvador, 1985: 56). A pesar de sus innegables ventajas, el rendimiento que se ha hecho de los atlas lingüísticos es claramente escaso, en comparación con la riqueza y cantidad de materiales que reúnen (Sousa Fernández, 2010).

El hecho de que sean proyectos tan largamente prolongados en el tiempo dificulta inevitablemente su uso, puesto que los datos de un atlas no pueden estar permanente actualizados sino que se refieren a un estado de la lengua concreto y determinado. Por otra parte, las dificultades de acceso a los atlas impiden en muchos casos su uso por parte de los investigadores. Es innegable que deberíamos encaminarnos hacia un nuevo soporte digital para los atlas lingüísticos, en versión electrónica y con facilidades de acceso —como está realizando con los mapas del ALPI el investigador David Heap o en la línea del moderno *Atlante Lessicale Toscano*— a fin de que toda la información que presentan no se pierda en el olvido y contribuya verdaderamente a los avances lingüísticos, a la elaboración de diccionarios, a los estudios sobre variación, a la confección de manuales rigurosos de enseñanza, etc.

Una aplicación directa de la información que ofrecen los atlas es la elaboración de diccionarios. Este trabajo es muy complejo y exige gran

minuciosidad⁷. Primero, se transforman los mapas en listas de palabras ordenadas con datos de frecuencia y distribución. Es preciso explicitar la definición o el significado de los datos distinguiendo, cuando las haya, diferentes acepciones; los cuestionarios en los que se ha basado el atlas facilitan esta parte del proceso. Por último, se debe dar cuenta de cómo un mismo concepto se denomina de diferentes formas, las regiones donde se emplea cada una de ellas y su frecuencia de uso⁸.

2.1.2. Diccionarios

Para una descripción más amplia del léxico hispánico —regional o global— podemos consultar un gran número de diccionarios, a pesar de que muchas de estas obras resultan insuficientes, fragmentarias, poco actualizadas o carentes de rigor desde un punto de vista metodológico (Haensch, 2002: 37). Esto se debe a que muchos de estos diccionarios se basan en obras primarias sobre voces americanas que, como ya hemos comentado, presentan un léxico poco actualizado y se inspiran en criterios folclóricos sin establecer unas bases lingüísticas. En otras ocasiones, los diccionarios no incluyen marcas de frecuencia, registro o diferentes acepciones de las palabras, y presentan cierta arbitrariedad a la hora de registrar las marcas diatópicas, señalando como generales de América palabras que son exclusivas de una zona o viceversa.

El primer diccionario sobre léxico hispanoamericano, que constituyó el punto de partida de tantos otros trabajos lexicográficos, fue el *Diccionario geográfico–histórico de las Indias Occidentales* de Antonio de Alcedo (1786–1789), máximo representante de la incipiente lexicografía colonial.

En el siglo XIX, se publicó el primer diccionario diferencial de Hispanoamérica, el *Diccionario provincial de voces cubanas* (1836), de Esteban

⁷ Sobre la aplicación de los datos de un atlas al diccionario es recomendable acudir a Castillo Peña (1991).

⁸ Un buen ejemplo de todo este proceso es el proyecto gallego de elaboración de un diccionario dialectal integral a partir de datos del ALPI, del que nos habla Sousa Fernández (2010).

Pichardo, que incluye aquellos términos específicos de la variedad cubana que tienen un significado diferente al peninsular o que no se utilizan en España.

Fue en el siglo XX cuando se despertó un verdadero interés por la lexicografía, que generó la elaboración de multitud de diccionarios. En una primera fase, los diccionarios provinciales, o de americanismos comunes a varias regiones, constituyeron una reafirmación nacional contra la supremacía de la norma lingüística española. Las referencias sobre el *uso lingüístico* —correcto o incorrecto— de las voces eran constantes. Según avanzó el siglo XX, aumentó el interés por el estudio del léxico en sí mismo así como el rigor lexicográfico, por lo que muchos diccionarios suprimieron las indicaciones prescriptivas e incluyeron voces propias regionales, tabúes, coloquialismos o voces incluso jergales. Ya en el siglo XXI se ha elaborado por primera vez un diccionario panamericano, el *Diccionario de americanismos* (2010), coordinado por las instituciones representantes de los países hispanoamericanos, que describiremos con mayor detalle en el capítulo 4 de esta tesis.

Casi todos ellos se han concebido como una suma de voces léxicas particulares de América, imperando el criterio diferencial y contrastivo. No encontramos casi ninguno que abarque el idioma español en su totalidad. Así, aunque su pretensión fuera valorar una variedad lingüística propia, estos diccionarios no consiguen desligarse de la referencia constante a la norma castellana (Zimmermann, 2003).

En Hispanoamérica existen dos grandes diccionarios integrales de lengua española. En primer lugar, el *Diccionario del español de México*, bajo la dirección de Luis Fernando Lara. Este diccionario se basa en el *Corpus del español mexicano contemporáneo (1921–1974)* y recoge todas las voces generales de la variedad mexicana del español, sin importar el origen ni las regiones que las utilicen⁹.

⁹ Subrayamos que en México no se han elaborado apenas diccionarios diferenciales o contrastivos como en los otros países hispanoamericanos (Fajardo Aguirre, 2010: 339).

También es destacable el *Diccionario integral del español de la Argentina* (DIEA), dirigido por Plager en 2008. Se fundamenta en el *Corpus inicial del español argentino*, que reúne muestras lingüísticas a partir de los años 80 y se elaboró expresamente para la creación de este diccionario. En él se presenta el habla usual de Argentina, incluyendo voces americanas o españolas indistintamente.

En todo caso, la lexicografía integral tiene grandes dificultades pues exige una gran inversión de tiempo, dinero y trabajo. Esto explica que la mayoría de diccionarios americanos adopten una visión diferencial, lo que tiene como consecuencia que, incluso hoy en día, el diccionario más consultado de lengua española sea el de la Real Academia Española. Pero la consulta de este diccionario no está exenta de problemas: por una parte, no incluye gran cantidad de voces americanas, y por otra, presenta multitud de voces que son exclusivas de la Península sin marca diferencial alguna.

Las primeras voces americanas que aparecieron en un diccionario de español se incluyeron en el *Diccionario de Autoridades* (1726–1739). A partir de ese momento, se fueron registrando cada vez más americanismos en los diccionarios de español de autoría europea, como en el *Nuevo diccionario de la lengua castellana* (1846), de Vicente Salvá, donde se incorporó por primera vez un gran número de voces de América.

A lo largo del siglo XX se concede mayor representatividad a los términos americanos, aumentando exponencialmente su presencia en los diccionarios:

[O]tra gran oleada de americanismos fue llegando con la conmemoración del V Centenario del Descubrimiento a las dos últimas ediciones del *DRAE* (1992 y 2001), que compensaron la carencia de un diccionario general de americanismos académico con una nueva incorporación masiva de americanismos, aportados por las Academias correspondientes de la Española, de manera que el número de marcas americanas en los artículos del diccionario creció de los 12.494 de la edición de 1992 a las 28.171 de 2001 (Fajardo Aguirre, 2010: 350).

No obstante, y a pesar del trabajo de las diferentes Academias americanas en la comisión del *DRAE*, todavía hoy los americanismos están infrarrepresentados.

Además, supone un grave problema lexicográfico el hecho de no marcar las palabras estrictamente peninsulares:

En la edición de 1992 del diccionario de la Academia, aparece por primera vez la marca de uso *españolismo*, aunque sólo en doce casos (cfr. Werner, 2002: 81–83). Entre ellos figuran tanto vocablos que se refieren a fenómenos no específicamente españoles, tales como *cañón*, como voces para fenómenos específicamente españoles, tales como *bonoloto*. En unos 500 casos, si bien no es posible hallar la marca de uso *españolismo*, sí encontramos, en cambio, indicaciones de que una determinada palabra o bien un determinado uso son privativos de España (Werner, 2002: 82). (Stork, 2006: 595).

A pesar de los esfuerzos por mantener una mirada ecuánime ante las variedades, lo cierto es que tanto la Real Academia Española como muchas otras academias americanas contribuyen a una preeminencia del español de España en los diccionarios generales frente al español de América. En este sentido, en palabras de F. J. Pérez (2000):

[C]obra sentido estudiar en los diccionarios los procedimientos descriptivos no sólo para privilegiar la pericia del redactor sino para observarlos como acercamientos epistemológicos: visiones del mundo, reflejo de las mentalidades, propuestas ideológicas, estructuras sociológicas o filosóficas, captación de la historia, palpitaciones espirituales, rencores y afectividades y, en suma, el tono de la vida que el diccionario, sin saberlo, [ha] impreso en sus observaciones pormenorizadas de las cosas (Pérez, 2000: 34).

Hay un diccionario que merece la pena destacar por su tratamiento del español americano: el *Diccionario de uso del español de América y España* (2002), de Lucena Cayuela. Esta obra, descrita en detalle por Stork (2006), incluye: acepciones distintas según su uso peninsular o americano, palabras que no se emplean la Península Ibérica y definiciones redactadas en un español panhispánico.

Respecto a los diccionarios que se ocupan exclusivamente del léxico hispanoamericano, encontramos dos tipos: aquellos que aspiran a recoger el léxico global de Hispanoamérica y los que incluyen voces exclusivas de un país concreto.

Los primeros diccionarios *globales* hispanoamericanos que no recogen el léxico exclusivo de un país sino que pretenden reunir unidades léxicas de toda América nacen en el siglo XX. Con frecuencia este tipo de diccionarios incurren en

inexactitudes y generalizaciones. El más importante de ellos fue el *Diccionario de americanismos* de Malaret en 1925.

En relación con los diccionarios *nacionales* o *regionales* del español de América, la mayor crítica que se ha hecho es su concepción lexicográfica diferencial, además de su falta de rigor metodológico al incluir un gran número de españolismos junto a voces específicamente americanas. En algunos casos, además, se aprecia una falta de rigor en las marcas diatópicas que se incluyen.

Respecto al tipo de información léxica que nos ofrecen estos diccionarios, Werner (2001) explica que sería especialmente interesante que estas obras presentaran los datos en dos direcciones: no solo interesa saber que *coche* apenas se emplea en Hispanoamérica, sino que es interesante conocer cuál es el término equivalente en cada uno de los países americanos, cómo se usa, si tiene precisiones de significado, etc. Además, sería necesario que los diccionarios incluyeran rasgos de variación respecto a la frecuencia, al uso, al estilo o al registro de las variedades.

Dentro del grupo de diccionarios regionales o nacionales destacan los diccionarios del proyecto Augsburgo, coordinados por Haensch y Werner desde 1993, que representan una búsqueda de rigurosidad metodológica y un deseo de contribuir a la lexicografía americana aportando materiales no registrados hasta el momento. De ellos destacamos la búsqueda de informantes que garanticen la *actualidad mínima* de las palabras usuales en diferentes países y el empleo de un español panhispánico en la redacción de sus entradas. Según Stork (2006: 597), los diccionarios de este proyecto son «los más perfectos de entre todos los diccionarios contrastivos, y con toda razón han sido muy elogiados». Aún así, estos diccionarios siguen siendo consultados, sobre todo, por especialistas, más que por un público general. En todo caso constituyen obras lexicográficas de gran valor para el estudio de la variación léxica.

Para que las múltiples informaciones que recogen todos estos diccionarios americanos fueran verdaderamente útiles para los estudios lingüísticos, sería necesaria una informatización de estas obras, aportando facilidades de acceso y uso

de los datos a los investigadores y haciendo un esfuerzo por mantener actualizados los estudios léxicos que se vienen realizando sobre el español de América. Solo así constituirían una fuente de información valiosísima para la descripción del léxico hispánico y evitaríamos que cayeran en el olvido los esfuerzos de tantos investigadores.

2.1.3. Estudios especializados y bases de datos

Además de los atlas lingüísticos y los diccionarios, para obtener información acerca de la distribución del léxico hispánico podemos consultar multitud de monografías, tesis doctorales, artículos de investigación, bibliografías, tesoros lexicográficos y glosarios dialectales exhaustivos que informan sobre el léxico de una región o de un campo semántico.

Aunque no existe un tratado completo sobre el léxico español, es relativamente sencillo encontrar estudios especializados sobre la historia viva de algunas palabras, sobre bases etimológicas, las evoluciones semánticas de los términos o cuestiones de interferencias lingüísticas.

A pesar de tantas publicaciones, nuestro conocimiento sobre el léxico hispánico es demasiado fragmentario. El investigador debe esmerarse en analizar exhaustivamente monografías, lecturas especializadas, repertorios bibliográficos o diccionarios para hacerse una idea *global* de la historia y distribución del léxico. Así, muchos de estos estudios presentan con exhaustividad la distribución de voces particulares o incluso del habla de una región o de un país, pero no son capaces de ofrecer una imagen completa de la variación léxica que nos permita comparar la distribución entre diversos territorios ni conocer el grado de frecuencia de uso de las variedades. En relación con nuestro objeto de estudio, resulta francamente difícil

encontrar inventarios léxicos fiables y sistemáticos sobre las equivalencias de voces del español peninsular y del español hablado en América¹⁰.

Entre los diversos estudios recientes sobre variación léxica del español nos gustaría destacar los siguientes: los estudios de disponibilidad léxica de López Morales (1995); los estudios de frecuencias léxicas del español como los de Terrádez Gurra (2001) o Cantos, Sánchez, Sarmiento y Almela (2005); los estudios generales y monográficos sobre el léxico hispánico de Colón Doménech (2002); la historia del léxico español de Lüdtke y Schmitt (2004); el trabajo de Ávila (2001) sobre el léxico hispánico en los medios de comunicación; el trabajo sobre léxico americano en manuales de ELE de Andión Herrero (2003); las investigaciones léxico-estadísticas de territorios hispánicos privilegiados por la profundidad y alcance de sus estudios léxicos, como es el caso del español de Puerto Rico; y los artículos sobre dialectometría y variación léxica hispánica de Ueda (1995, 2003, 2007a, 2013).

Por otra parte, en las últimas décadas, además de asistir a la informatización de un gran número de diccionarios, observamos una importante tendencia a trabajar con bancos de datos electrónicos y recursos de Internet:

Cabe preguntarse si presentaciones del alcance diatópico del léxico en forma de diccionarios de *mexicanismos*, *argentinismos*, *cubanismo*s etc. constituyen una solución adecuada para representar estos hechos. Por lo menos desde el punto de vista técnico parecen más prometedores banco de datos electrónicos (Zimmermann, 2006).

Así, podemos consultar en línea, por ejemplo, el *Portal de Léxico Hispánico* (<http://portaldelexico.es/>), que es un centro de información que almacena datos bibliográficos, lingüísticos y documentales sobre el vocabulario de las lenguas romances de la Península Ibérica y sus respectivas variedades. El proyecto fue elaborado desde el departamento de Filología Española de la Universidad Autónoma de Barcelona y se ha desarrollado en colaboración con la Real Academia Española.

¹⁰ Vid. Sobre esta escasez de investigaciones lingüísticas orientadas al estudio del léxico hispánico nos hablan diversos autores, entre los que destacamos la contribución de Haensch, Ettinger, Wolf y Werner (1982) y la de Colón (2002).

En las últimas décadas, los estudios e investigaciones sobre variedades léxicas se han centrado en otro aspecto que complementa a los índices de frecuencia: la llamada *disponibilidad léxica*. El interés nace a mediados del siglo pasado cuando los lingüistas dedicados al francés como lengua extranjera percibieron que las palabras más frecuentes de un idioma no son las que necesita aprender un estudiante. Dicho de otro modo, si se enseñan las palabras más frecuentes de un idioma —las que predominan en el discurso según los métodos estadísticos— se pierden palabras esenciales que no aparecen en corpus lingüísticos. De esta idea nace la necesidad de estudiar las palabras *disponibles*, es decir, las palabras que vienen a la mente de un hablante nativo ante un tema de conversación determinado. Esta metodología de estudio del léxico ha tenido grandes adeptos en los estudios léxicos hispánicos. Es preciso destacar las investigaciones del *Proyecto Panhispánico de Disponibilidad Léxica* (<http://www.dispoplex.com/>), común a España e Hispanoamérica, para la creación de un diccionario con índices de disponibilidad de las palabras del español. También existe el programa *LexiDisp*, que aplica, sobre un conjunto de datos, una fórmula matemática para el cálculo de la disponibilidad léxica (Moreno Fernández, 1995).

Otros estudios relacionados con el léxico del español se centran en la *comprensibilidad léxica* entre los distintos países hispánicos. Los resultados siempre dependerán de la metodología empleada en las investigaciones. Así, por ejemplo, López Chávez obtiene en sus cálculos una compatibilidad léxica muy baja: «Los 100 primeros vocablos son compatibles por abajo del 70 % y ninguna comparación de los lexicones como totalidad alcanza un 56 % de compatibilidad» (López Chávez, 1993: 1006). Sin embargo, en 1990, Raúl Ávila (1990) había obtenido conclusiones estadísticas muy diferentes, situando la compatibilidad léxica del español en torno al 92 %.

Por su parte, Lope Blanch (2000) afirma que los datos de López Chávez se pueden contradecir adoptando otra visión metodológica, para lo que realiza un estudio en el que pretende determinar la compatibilidad léxica entre hablantes

madrileños y mexicanos, tomando como base unas encuestas de Alvar y Quilis. Sus resultados muestran tres opciones posibles: (a) la palabra utilizada en Madrid se conoce en México con otro significado; (b) la palabra utilizada en Madrid se conoce en México pero los mexicanos prefieren emplear otra; (c) la palabra utilizada en Madrid no se conoce ni utiliza en México. El autor corrobora que la mayor parte de *madrileñismos* inusitados en México pertenecen al ámbito familiar, popular, figurado o a la jerga estudiantil. Por lo general, observa que el 99 % de las palabras utilizadas en el discurso de los madrileños son conocidas o utilizadas por mexicanos. Estas diferencias de porcentaje vienen determinadas, sobre todo, por incluir tanto términos gramaticales como denotativos en el estudio. Como vemos, según el tipo de estudio, encontramos diversas cifras respecto a la compatibilidad léxica.

Queremos destacar también las investigaciones que se están desarrollando respecto al denominado *español neutro*, que constituye «un modelo estandarizado que sin perder los filtros restrictivos de lo común, toma muestras de normas asociadas a núcleos o comunidades de prestigio más concretos dentro del mundo hispánico y las expande por todos sus usuarios» (Andión Herrero, 2008: 15). Algunos trabajos interesantes que reflejan la importancia de este español panhispánico son los siguientes: Ávila (2001), Demonte (2001), López Morales (2001), López González (2002), Gómez Font (2003) o García Izquierdo (2009).

Sobre la variación entre España y América existe tan solo un proyecto internacional que se coordina desde la Universidad de Tokio. El estudio de Varilex: *Variación léxica del español en el mundo* (<http://lecture.ecc.u-tokyo.ac.jp/~cueda/varilex/>), que surgió con el objetivo de estudiar la variación de las palabras en las distintas zonas geográficas del español. El proyecto trata de formar una red internacional de investigación sobre el léxico español moderno-urbano, reuniendo una gran cantidad de datos con el objetivo de analizarlos y cartografiarlos en múltiples mapas de grandes ciudades hispanohablantes. Este proyecto recuerda, por tanto, a los atlas lingüísticos y se vertebra sobre bases propias de la geografía lingüística. La versión digital de los datos, informatizada y actualizada a tiempo real, facilita enormemente cualquier análisis de la variación. No

nos detenemos ahora a explicar el proyecto puesto que en el siguiente capítulo presentamos un análisis exhaustivo del mismo.

2.1.4. Lingüística de corpus

Antes de terminar esta sección no podemos olvidar la importancia que han cobrado los corpus en los estudios lingüísticos sobre variación. Un corpus lingüístico es «un conjunto, normalmente muy amplio, de ejemplos reales de uso de una lengua» (Giráldez Ceballos-Escalera, 2009: 98). Se trata de documentos organizados e informatizados con un propósito de investigación concreto. Crystal explica que un corpus es una colección de datos que pueden ser utilizados como punto de partida para descripciones lingüísticas o como un medio de verificación de hipótesis acerca de una lengua (Crystal, 1997).

La lingüística de corpus es la disciplina que, asociada a la lingüística computacional, estudia los rasgos lingüísticos a través de una variedad significativa de muestras de lengua. Constituye una herramienta de apoyo con enorme potencial para estudiar de manera exhaustiva la lengua y la comunicación: «Using a corpus will for some years to come be a voyage of discovery at every level of education — the student, the teacher, the class, the institution, the educational authority, the curriculum planners, the publishers» (Sinclair, 2004: 294).

Todo corpus constituye un conjunto de datos lingüísticos sistematizados según diversos criterios. Por tanto, cualquier conjunto amplio de textos no puede considerarse un corpus. La muestra de un corpus debe ser amplia (en principio, cuantos más textos mejor) pero también significativa, de manera que responda a los objetivos de una investigación. Así los corpus deben ser *representativos* de un tipo de discurso, de manera que permitan generalizar los resultados que de ellos se obtengan.

Si analizamos la historia de la lingüística de corpus observamos su rasgo incipiente. No obstante, llama la atención la rapidez con que ha calado el estudio de corpus entre los lingüistas. En palabras de Sánchez *et al.* (1995):

Quienes han trabajado con los corpus lingüísticos se sienten pronto inmersos en «un mundo en expansión», con posibilidades sin fin en cuanto al aprovechamiento de los datos que se ofrecen al investigador. Probablemente los lingüistas ya no pueden permitirse el lujo de prescindir de estas ayudas (Sánchez *et al.*, 1995: 13).

Así, desde los años 90 se ha producido un resurgimiento de los métodos empíricos y estadísticos en el análisis de la lengua. En la actualidad se desarrollan muchas investigaciones basadas en corpus sobre adquisición de conocimiento léxico, estudios socioculturales o estilística. El estudio de corpus también se ha aplicado a otras disciplinas como la traducción automática, la recuperación de información o la interacción en el discurso (Pérez Hernández, 2002).

Como se puede observar, para la lingüística aplicada el trabajo con corpus es fundamental porque permite extraer conclusiones fiables a partir del análisis de datos reales. El estudio con corpus pone de manifiesto que, en ocasiones, el uso *real* que los hablantes hacen del idioma no se atiene a lo esperado por los gramáticos puesto que frecuentemente observamos una disociación evidente entre el lenguaje *real* y el *oficial* (Sánchez *et al.*, 1995). Frente a las reflexiones introspectivas del lingüista experto, el corpus aporta la evidencia empírica, objetiva, de los datos; esto es, la lengua viva en sí misma. Así podemos considerar que los corpus no solo son útiles, sino que son esenciales para el estudio lingüístico.

Existen muy diversos tipos de corpus, según diferentes criterios. Los corpus que suponen una representación de las lenguas, que se utilizan para elaborar diccionarios, gramáticas u otras obras similares, se denominan corpus *de referencia*. Aquellos corpus que exponen un tipo de lengua, con algún propósito específico, se denominan corpus *especializados*. Si los corpus recogen el registro oral, se denominan corpus *orales*, mientras que los que se compilan en varias lenguas son *multilingües*. Por último, los corpus de tamaño constante a los que se van añadiendo datos automáticamente se denominan corpus *monitor*.

La mayor parte de los corpus actuales presentan textos etiquetados. Esto significa que las palabras que los componen llevan anotaciones, a modo de etiquetas, con información específica. Esta característica es fundamental para la investigación

puesto que nos permite acceder a los rasgos morfológicos, fonéticos, sintácticos o semánticos de las palabras que analizamos. Sánchez (1995) explica que los corpus etiquetados tienen enormes ventajas para realizar consultas lingüísticas, a través de programas diseñados específicamente para gestionar este tipo de datos.

En lo que se refiere al léxico, las consultas de vocabulario son el primer objetivo perceptible de un corpus o, al menos, el más perceptible (Sánchez *et al.*, 1995: 16). Una de las aplicaciones más conocidas del trabajo con corpus es precisamente la extracción de listados de frecuencias, que permiten obtener conclusiones fidedignas respecto al *uso* del léxico por parte de los hablantes¹¹.

En este sentido, los corpus son fundamentales para la elaboración de diccionarios, como por ejemplo, el *Gran diccionario de uso del español actual*, basado en el corpus CUMBRE. Estos diccionarios pueden ofrecer información sobre la frecuencia de palabras o incluir, en ellos, ejemplos reales de uso. Por otra parte, los corpus también son útiles en el campo de la enseñanza de segundas lenguas, pues permiten distinguir las nociones léxicas elementales (*léxico básico*) y aumentar gradualmente la expresividad (*léxico marginal*).

En un trabajo como el nuestro, es muy interesante utilizar un corpus para reflexionar sobre el vocabulario empleado en España y América. Los corpus presentan el *uso vivo* del idioma, lo que nos permite acceder a las palabras de mayor rentabilidad léxica. Para ello, sería fundamental acceder a un corpus que reúna textos de España e Hispanoamérica, con el objetivo de arrojar datos empíricos sobre la utilización real de las voces léxicas que pretendemos analizar. También es importante trabajar con textos *etiquetados* que marquen la región geográfica en la que se haya registrado cada palabra. Los corpus de libre acceso que permiten estudiar el léxico de España y América son: el CORPES, el CREA, el CORDE y el PRESSEA. Respecto a las limitaciones e impedimentos que hemos encontrado en esta investigación para el trabajo con corpus hablaremos en el capítulo 4.

¹¹ Véase, por ejemplo, el excelente trabajo sobre frecuencias léxicas del español coloquial de Terrádez Gurrea (2001).

2.2. Descripción general de la variación léxica

El fenómeno de la extensión geográfica de las variedades léxicas es muy complejo de analizar. El léxico varía por condiciones lingüísticas (semánticas, pragmáticas, sintácticas, fonéticas) y extralingüísticas (personales, sociales, históricas, geográficas, psíquicas, etc.). Según el sexo, la edad, la educación, el nivel económico, la experiencia de viajes, el contacto con hablantes de otras procedencias, la inteligencia, la religión, etc., cada individuo emplea activamente unas palabras u otras.

Dentro del léxico hispánico podemos distinguir: *panhispanismos*, utilizados tanto en España como en América; *americanismos*, utilizados en uno, varios o todos los países de América; y *españolismos*, voces exclusivas de España. Para profundizar sobre el léxico y su variación, Chacón García (2014) analiza la variedad léxica del español atendiendo a diferentes ámbitos: vocabulario de civilización (*arepa, carpincho, piraña...*), falsos arcaísmos (*bravo, durazno, pollera...*), usos excluyentes (*abarrotes, saco, rompehuelgas...*), desplazamientos de significado (*¿le provoca un tinto? > ¿le apetece un café?*), equivalencias de significado (*autobús vs. chiva*), diferencias estilísticas, morfológicas, pragmáticas (*estrecho–angosto; florería–floristería*) así como las palabras tabú, los disfemismos y los eufemismos (*concha, bicho*). Por su parte, Haensch (2001; 2002) presenta una clasificación exhaustiva y rigurosa de las diferencias léxicas entre el español americano y el peninsular, fundamentada en sus investigaciones lexicográficas.

Partiendo de las publicaciones mencionadas, presentamos a continuación una tabla que sintetiza la tipología de la variación léxica del español que encontramos al comparar el vocabulario de España e Hispanoamérica. En ella podemos observar la complejidad inherente al tratamiento de este fenómeno. No pretendemos ofrecer un repertorio léxico exhaustivo sino una síntesis de las diferencias léxicas esenciales entre el español europeo y americano.

Tabla 2
Diferencias léxicas entre España e Hispanoamérica

	Hispanoamérica	España
Diferente grafía	<i>soya</i>	<i>soja</i>
Diferente acento	<i>video</i> [bidéo]	<i>vídeo</i> [bídeo]
Diferente género	<i>el radio</i> (aparato) [Mx, Centroamérica, las Antillas, Ec, Co y Ve]	<i>la radio</i> (aparato)
Verbos reflexivos y verbos intransitivos	<i>regresarse</i>	<i>regresar</i>
Diferente uso de preposición o de régimen	<i>en la mañana</i>	<i>por la mañana</i>
Adjetivos en función adverbial	<i>ella canta muy lindo</i>	<i>ella canta muy bien</i>
Adverbios típicamente americanos	<i>acá/allá</i>	<i>aquí/allí</i>
Cambios de sufijo	<i>friolento</i>	<i>friolero</i>
Falsos arcaísmos	<i>pollera</i>	<i>falda</i>
Geosinónimos	<i>camión</i> [Mx]	<i>autobús</i>
Desplazamientos de significados: homonimia, polisemia	<i>tinto</i> (<i>café solo</i> , [Co])	<i>tinto</i> (<i>vino tinto</i>)
Misma unidad léxica formal, diferencias connotativas	<i>enojarse</i>	<i>enfadarse</i>
Misma unidad léxica formal, diferente frecuencia	<i>demorar</i>	<i>tardar</i>
Distinto uso contextual y pragmatismo lingüístico	<i>¡Atención, faenas!</i> [Ch]	<i>¡Precaución, obras!</i>
Tabuización, destabuización y eufemismos	<i>coger</i>	<i>realizar el coito</i>

3. Conclusión

Aunque el estudio de la variación lingüística ha sido tradicionalmente objeto de la dialectología, desde principios del siglo XX asistimos al auge de la geografía lingüística. En esta disciplina, que parte de una metodología estricta para el estudio de la variación, se enmarca la presente tesis doctoral.

Tanto en España como en Hispanoamérica se han desarrollado diversos estudios que registran la variación de los fenómenos lingüísticos a través de mapas geográficos que delimitan su extensión y distribución de manera sistemática. Sin embargo, la distribución diatópica de las variantes léxicas del español carece de estudios exhaustivos para la totalidad del territorio hispanohablante. A pesar del enorme número de diccionarios y de las fuentes de consulta sobre variación léxica ya comentadas, las limitaciones que existen para el estudio riguroso del tratamiento del vocabulario son evidentes. Por ello, el componente léxico ha sido el plano que, por el momento, menos se ha interpretado y sistematizado a nivel geolectal.

Ante las dificultades que encuentra un lingüista para consultar datos rigurosos sobre la distribución geográfica de la variedad léxica, nuestro estudio pretende mostrar la representatividad de 182 conceptos del *PCIC*, en función del número de hablantes y de países en que se recogen las voces. Para calcular el grado de representatividad es necesario identificar el léxico característico de las diferentes regiones. Para ello, en este proyecto trabajamos con diccionarios, corpus, estudios especializados y tomamos como base los datos del proyecto Varilex, del que hablaremos con mayor detalle en el siguiente capítulo.

Por otra parte, es necesario explicar que nuestra investigación, casi por necesidad epistemológica, es defensora de la existencia de la sinonimia, al menos en el plano discursivo: «las unidades léxicas, como ocurre con determinadas construcciones sintácticas, pueden verse neutralizadas semánticamente en el discurso, en el uso comunicativo de la lengua» (Moreno Fernández, 2009: 32).

Esto no quiere decir que exista una neutralización absoluta entre las dos palabras, pero admitir cierta neutralización permite afrontar el estudio de las alternancias léxicas en unas condiciones lingüísticas y extralingüísticas determinadas.

Antes de concluir, remitimos al Anexo I donde presentamos una selección de diccionarios y fuentes de documentación relevantes para aproximarse al fenómeno de la variación léxica del español. No hay que olvidar que a pesar del elevado número de obras sobre variación lingüística publicadas hasta la fecha, «la situación actual es particularmente insatisfactoria en cuanto a diccionarios, vocabularios, glosarios, etc. del español de América, la mayoría de los cuales dejan mucho que desear» (Haensch, 2001: 65). Se trata, por tanto, de complementar la información de este capítulo con una revisión crítica personal. No se han incluido monografías regionales puesto que no son objeto de nuestro trabajo. El objetivo es presentar un panorama general con las obras fundamentales sobre la variación léxica en el mundo hispánico.

CAPÍTULO 3

PROYECTO INTERNACIONAL DE VARIACIÓN LÉXICA DEL ESPAÑOL EN EL MUNDO: VARILEX

1. Introducción

El objetivo fundamental de esta tesis es estudiar la variación léxica de 182 nociones del *PCIC*, incorporando voces representativas del español junto a las nociones propuestas por este documento de referencia para la enseñanza de ELE. En el capítulo anterior hemos presentado un recorrido histórico sobre los estudios de variación lingüística, centrándonos en las investigaciones sobre la variación diatópica del componente léxico. En este capítulo vamos a describir uno de ellos: el *Proyecto de Variación Léxica del español en el mundo* (Varilex), que nos ha dado acceso a su base de datos, gracias a la cual ha sido posible plantear nuestra investigación.

Siguiendo las huellas de los atlas lingüísticos de gran dominio, el proyecto Varilex propone una metodología rigurosa y sistemática para la recolección de datos de variación léxica, por lo que resulta idóneo para nuestro estudio. Por otra parte, la ausencia de inventarios globales de léxico hispánico justifica igualmente nuestra elección. Aunque los diccionarios y otros estudios específicos sobre el vocabulario del español son útiles para mostrar una aproximación general del léxico de cada región, la oportunidad de contar con los datos de un proyecto de estudio léxico

internacional a gran escala constituye una ocasión única para estudiar la distribución de las voces del español moderno urbano.

Nos centraremos, primero, en describir las características de Varilex, analizando las fases de desarrollo e implementación del proyecto, el método de obtención y análisis de datos, la relevancia y el impacto que ha tenido en estudios posteriores así como otros aspectos destacados. También vamos a describir su metodología, analizando la red de puntos, las encuestas, los informantes y el trabajo de los colaboradores internacionales. Describimos, además, la base de datos digitalizada de Varilex, que es la fuente de datos primaria de nuestra investigación, especificando las características técnicas y las posibilidades que ofrece. Para terminar, exponemos una síntesis con los puntos fuertes y limitaciones de este proyecto, que son muy importantes a la hora de abordar nuestra investigación.

2. Origen de Varilex

El proyecto Varilex es fruto de la colaboración de investigadores y dialectólogos en una red de investigación internacional, algo que no es novedoso en los estudios de léxico¹. En vez de utilizar la colaboración de investigadores únicos

¹ La lingüística hispánica ha desarrollado otros proyectos coordinados por lingüistas a modo de red internacional, algunos de los cuales ya han sido comentados en el capítulo 1. Desde que Lope Blanch lanzó el proyecto del PILEI o «Proyecto de estudio del habla culta de las principales ciudades de Hispanoamérica», se han ido creando otros proyectos internacionales de gran alcance. Destaca, sobre todo, el ya mencionado «Atlas Lingüístico de Hispanoamérica» de Manuel Alvar y Antonio Quilis, que reunió múltiples materiales aunque hasta el momento solo algunos de ellos han sido parcialmente publicados.

Otros proyectos interesantes son, entre otros, el proyecto «Difusión Internacional del Español» (DIES) dirigido por Raúl Ávila, el proyecto «Corpus Integral del Español Actual», propuesto por Fernando Lara; los «Estudios del Discurso de Cortesía de Español» (EDICE), coordinados por Diana Bravo; y el «Proyecto para el estudio sociolingüístico del español de España y América» (PRESEEA), que aporta un corpus con materiales reales de lengua representativos del mundo hispánico.

Como hemos visto, en la actualidad también las academias de la lengua española han comenzado a trabajar conjuntamente para publicar diccionarios, entre los que destaca el *Diccionario de Americanismos* (2010), además de la redacción de la *Nueva gramática de la lengua española* (2008) y la *Ortografía de la lengua española* (2010).

como en la dialectología tradicional, es precisamente la colaboración internacional la que hace posible la viabilidad de un proyecto de esta envergadura, reuniendo datos de 20 países hispanohablantes a partir de unas mismas encuestas y de una metodología común.

Tal y como ya hemos constatado en el capítulo 2, en los estudios de variación se habla recurrentemente de la imposibilidad de sistematizar el léxico de toda la geografía hispánica. Casi todos los investigadores coinciden en que el problema se explica por la falta de datos. Si no disponemos de información suficiente, el investigador tiene dos opciones: abandonar la misión de estudiar la variación léxica o aventurarse a elaborar repertorios de vocabulario modestos que cubran algunas de las lagunas existentes.

El proyecto Varilex nació para elaborar una gran compilación de datos. Este grupo de investigación japonés, con centro en la universidad de Tokio, consideró la pertinencia de desarrollar un proyecto que permitiera estudiar el léxico hispánico a través de un trabajo sistemático. Así, tras una fase de estudio preliminar, en 1993 nació Varilex, con la intención de retratar el léxico hispánico urbano a través de una metodología rigurosa, mediante la colaboración de investigadores ubicados en diferentes regiones del mundo.

El objetivo fundamental de Varilex es, por tanto, recolectar datos de forma sistemática y unificada. Se trata, en primer lugar, de suplir este vacío de información que permitirán después estudiar en profundidad fenómenos de lenguas en contacto, interferencias, focos de irradiación, etc. El trabajo de Varilex es fundamental puesto que si no contamos con datos, nunca llegarán los estudios y aplicaciones posteriores.

El primer problema con el que se encuentra una investigación de este tipo es el dilema sobre la exhaustividad o la generalidad. Una lengua hablada y extendida por territorios tan alejados geográficamente y con costumbres, tradiciones, modelos económicos y culturales, tan dispares, presentará grandes diferencias de vocabulario. Un atlas geográfico, en este caso léxico, de gran dominio se ve obligado a renunciar a las particularidades regionales con el objetivo de presentar las peculiaridades de

una lengua común, lo que no entra en conflicto con la elaboración de otro tipo de estudios que retraten características específicas de las diferentes zonas.

Varilex apuesta por afrontar el estudio del léxico desde un enfoque unitario que favorezca la visión simultánea de todo el dominio hispánico y la coherencia de los materiales de estudio. Recoger una lengua común permite descubrir si existe una *koiné* del mundo hispánico, lo que tiene implicaciones evidentes en la comunicación y aplicaciones inmediatas. Este proyecto de investigación se centra en el estudio de una lengua común más allá de las grandes diferencias, al tiempo que pone de manifiesto la diversidad geolectal del continente americano y de España.

3. Descripción general del proyecto

El proyecto Varilex surge para estudiar la variación léxica del español en el mundo, teniendo en cuenta los diferentes factores que explican el cambio lingüístico en la sociedad actual. En primer lugar, la era de la información hace que los habitantes de las ciudades estén en contacto con una abundancia informativa sin precedentes que condiciona inevitablemente el préstamo léxico y las interferencias lingüísticas. Así, la democratización de la información a través de Internet así como la difusión internacional de programas de radio, televisión o de películas de habla española, favorecen el intercambio léxico y el viaje de las palabras más allá de las fronteras tradicionales.

Por otra parte, el aumento de las relaciones directas entre los habitantes de las distintas zonas, por factores de inmigración o de turismo, explican el intercambio de bienes, personas y, por supuesto, de palabras, que tienen como consecuencia un efecto de nivelación u homogeneidad cultural y, en último término, lingüística. Así, se implantan términos que proceden de otros países y culturas, en función de la moda o de las influencias comerciales; palabras que pueden ser transitorias y pasajeras o, por el contrario, de fuerte arraigo según los casos. Además, se crean nuevos términos para conceptos o realidades nuevas, lo

que favorece la renovación constante del vocabulario por influencia de la sociedad de consumo y de la publicidad.

Ante este mundo cambiante que responde a una velocidad vertiginosa de influencias e innovaciones léxicas, pero que además cuenta con un proceso de nivelación determinado por el mayor contacto entre los hispanohablantes, no podemos estudiar el léxico urbano utilizando las mismas técnicas que la dialectología tradicional. En un proyecto de investigación sobre léxico urbano interesan términos relacionados con la vida actual, más que un vocabulario histórico. Además, los informantes responderán a un perfil culto o de un nivel sociocultural medio, frente al informante rural que se buscaba en la dialectología. accediendo a núcleos urbanos, en vez de a zonas rurales. Por último, el cartografiado de mapas se realizará por puntos que representan grandes núcleos población, sin centrarse en las tradicionales zonas dialectales ni el dibujo de isoglosas lingüísticas.

3.1. Fase preliminar e inicio del proyecto

Antes de comenzar el proyecto Varilex se realizaron varias investigaciones lingüísticas en España, México y Chile, con el objetivo de observar las semejanzas y diferencias léxicas en estos tres puntos. A través de un estudio de conversaciones semidirigidas y de listas de palabras, se realizaron una serie de entrevistas con el objetivo de investigar acerca del léxico activo de las tres regiones y determinar las características del futuro cuestionario del proyecto.

Durante 1992 y 1993 ampliaron las encuestas y se obtuvieron variantes en Panamá, Colombia, Ecuador, Venezuela, Perú, Uruguay, Argentina, Cuba y Puerto Rico, además de en México, España y Chile. Se realizaron una serie de análisis de correlación que pusieron de manifiesto la existencia de cinco zonas diferenciadas en el análisis léxico (norte, andina, cono sur, mar Caribe y España).

Para complementar el estudio extensivo del léxico se realizó un estudio intensivo del vocabulario de la Ciudad de México, para incluir materiales que

permitieran estudiar el léxico diferencial del español. Además de consultar informes e investigaciones publicados hasta la fecha, se planteó una revisión y contra-revisión de estos datos, acompañada de observaciones directas para analizar la realidad lingüística y documentar parte del léxico diferencial. Todas las palabras recogidas eran candidatas para realizar un estudio comparativo del léxico moderno urbano del español.

Así se gestó el proyecto de *Variación Léxica del español del mundo*, que se presentó en 1993 en México, en el X Congreso de ALFAL. Está coordinado por Hiroto Ueda de la Universidad de Tokio, Toshihiro Takagaki, de la Universidad de Estudios Extranjeros de Tokio y Antonio Ruiz Tinoco de la Universidad Sophia de Tokio.

El objetivo principal era muy ambicioso: una red internacional de investigación que permita estudiar el léxico urbano actual de los distintos países hispanohablantes, para después analizarlos con métodos multivariados y elaborar mapas lingüísticos que den cuenta de la distribución de la variedad, tal y como lo haría un atlas lingüístico. No existía –ni existe hasta la fecha– ningún otro estudio exhaustivo e informatizado del léxico urbano, por lo que muchos colaboradores entendieron este proyecto como una oportunidad única para el estudio del léxico hispánico.

Los datos que se obtienen de este proyecto son de libre acceso, por lo que pueden ser consultados y utilizados por los investigadores así como por los participantes del proyecto. Se pretende crear un banco de datos léxicos que esté siempre actualizado y que permita realizar estudios dialectológicos, lexicográficos, taxonómicos, históricos, etc.

Así, en Varilex se combinan técnicas de investigación de la geografía lingüística con técnicas estadísticas e informáticas para procesar grandes cantidades de datos. La característica de trabajar con un grupo abierto internacional, junto a investigadores de todo el mundo, es fundamental para una lengua con la extensión geográfica del español que cambia con gran dinamismo en el plano léxico. Las

enormes distancias del dominio hispánico hacen necesaria una investigación exhaustiva, a modo de red internacional de cooperación para el estudio de la variación léxica. Así, se sumaron a Varilex importantísimos investigadores como Manuel Alvar, Bernard Pottier, Hunter Haensch, Humberto López Morales, Raúl Ávila, Brian Steel, Gerd Wotjack, David Heap, Carmen Silva–Corvalán, Francisco Moreno, Pilar García Mouton, Violeta Demonte o José G. Moreno de Alba, entre otros muchos. Sin ellos habría sido imposible crear un banco de datos léxicos actualizado del español.

3.2. Relevancia y aplicaciones del proyecto

El hecho de crear un proyecto internacional con un archivo libre, abierto y constantemente actualizado supone una renovación en la investigación dialectológica. Debido a las limitaciones que tiene el estudio del léxico hispánico, relacionadas con la dimensión temporal —los cambios lingüísticos rápidos que traspasan las tradicionales áreas dialectales— y espacial —la enorme magnitud del territorio de habla española—, Varilex constituye una aportación novedosa y de gran valor para el estudio del léxico moderno urbano de nuestra lengua.

La trascendencia e importancia de este proyecto ha fundamentado nuestra elección y puede resumirse en los siguientes puntos:

- 1) No existe hasta el momento ninguna otra investigación exhaustiva del léxico urbano actual de las grandes ciudades hispánicas.
- 2) Varilex es un proyecto abierto, con un banco de datos disponible para todos los usuarios. Utilizar bases de datos en línea no solamente ofrece la posibilidad de analizar los datos con *software* sino que permite trabajar en grupo con facilidad.
- 3) El proyecto está en constante dinamismo y actualización, por lo que se ofrecen los datos más actualizados en cada momento. Es una innovación en la metodología lingüística.

- 4) Del proyecto se generan múltiples y diversas publicaciones y estudios. De hecho, con la colaboración de los hispanohablantes, los datos recogidos permitirían elaborar un atlas lingüístico mundial de la lengua española.
- 5) Está complementado y colabora actualmente con otro proyecto internacional sobre los textos de prensa del mundo hispánico: *Difusión Internacional del Español por Radio, Televisión y Prensa* (DIES–RTP).

Según Ueda (2009), si pensamos en las aplicaciones del proyecto Varilex, los estudios relacionados pueden agruparse en seis ámbitos diversos:

- 1) *Metodología*: Sobre el método de investigación, la encuesta y los colaboradores destacan las publicaciones de Takagaki (1993) y Ueda (1994, 1996a, 1999, 2003). Respecto a la cartografía automática y la base de datos remitimos a Aoto (2004) y Ruiz Tinoco (1999, 2000, 2001). En relación con las técnicas estadísticas para el análisis de los datos los coordinadores del proyecto se han centrado en investigar la matriz de coeficientes de correlación y el análisis de cluster y en la elaboración de un método propio de patronización.
- 2) *Datos de distribución geográfica*: desde 1993 el Equipo de investigación Varilex (1993–2006) publica los resultados tanto de forma impresa como en la página web: <http://gamp.c.u-tokyo.ac.jp/~ueda/varilex>
- 3) *Comentarios*: Varilex publica informes sobre las variedades locales específicas estudiadas, analizando de manera crítica los resultados obtenidos. Algunos ejemplos pueden consultarse en Fernández (1994–2004), Nakajima (1999–2000), entre otros.
- 4) *Temas lingüísticos*: son importantes las aportaciones de Ueda (1995a) sobre fonética y fonología, Koike (2001) sobre fraseología, Nishimura (2002) sobre diminutivos, Aoto (2003) sobre las

- perífrasis verbales de gerundio, Ueda (1996b, 2000, 2005) sobre el léxico y Ruiz Tinoco (2002) sobre el parámetro sexo.
- 5) *Temas dialectales*: destaca la aportación de Ueda (1995b) sobre la zonificación del español en el mundo, Aoto (2005b) dedicado a la variedad de Paraguay y Ueda (2007) sobre la variedad del léxico de Guinea Ecuatorial con los datos reunidos tanto *in situ* como en París con la ayuda de los residentes ecuatoguineanos.
 - 6) *Aplicaciones de métodos para la enseñanza de español*: Ueda (2003) presenta algunas reflexiones sobre el tema de la aplicación del estudio dialectológico a la enseñanza del español. Uno de los resultados concretos puede ofrecerse en forma de diccionario como los dos diccionarios de Rubio y Hueda (1992, 1994) o el diccionario de bolsillo de Takagaki (2003), elaborados con datos recogidos por Varilex.
 - 7) *Otras aplicaciones*: otra aplicación posible es ofrecer los datos léxicos y propuestas específicas a los medios de comunicación, que se enfrentan diariamente al problema de la variabilidad léxica en un mundo globalizado (Ueda 1996c), a empresas de doblaje o, incluso, a multinacionales que trabajan en distintos puntos de la geografía hispánica.

Por todo ello, consideramos que Varilex es el proyecto adecuado para esta investigación, aportando datos fiables y sistemáticos sobre la variación que nos permiten determinar la representatividad de las distintas variantes léxicas de nuestro estudio.

3.3. Metodología

3.3.1. Obtención de los datos

Para el estudio de la variación léxica pueden recurrirse a tres métodos de obtención de datos: observación directa, encuesta directa o encuesta por entrega. Presentamos a continuación las características de cada uno.

La observación directa permite al investigador tomar contacto general con el área que va a investigar, para obtener un conocimiento básico de los aspectos fundamentales necesarios para preparar un cuestionario o un estudio de investigación (Sierra Bravo, 1992). Al entrar en contacto con el fenómeno que se pretende estudiar y analizarlo, el investigador se familiariza con la realidad y puede formular las preguntas de investigación. La observación directa permite reflexionar sobre el fenómeno de estudio de manera natural y con exhaustividad.

La encuesta directa, en cambio, se caracteriza por obtener datos representativos de una realidad más numerosa, lo que le concede un mayor alcance. Aporta seguridad y precisión de los datos recogidos puesto que el investigador *in situ* plantea cuestiones directamente a los informantes.

La encuesta por entrega tiene gran tradición en la geografía lingüística, como ya vimos en el capítulo 1 con los pioneros trabajos de Wenker en Alemania. Las ventajas de este tipo de encuesta son, sobre todo, prácticas, respecto al tiempo de recogida de los datos y el menor número de recursos financieros, de trabajo y logísticos que requieren. Además, estas encuestas tienen carácter sincrónico y pueden realizarse simultáneamente en muchos lugares. Es cierto que las encuestas por entrega tienen el problema de que a veces no consiguen altas tasas de respuestas pero las ventajas del método son innegables (Wood, 1989): la posibilidad de procesar las respuestas informáticamente incorporándolas directamente a una base de datos, de donde se pueden extraer conclusiones y comparar resultados con rapidez, rigor y eficacia; y la ordenación de los datos de acuerdo a las necesidades de cada investigación.

Varilex ha optado por combinar diferentes métodos de investigación en sus estudios de variación. En una primera etapa, se recurre al estudio de materiales publicados y al método de observación directa para realizar lo que denominan las *calas léxicas*. A través de encuestas *in situ* y de un exhaustivo análisis de fuentes secundarias (documentos y otras investigaciones) se registra un estudio general de la variación léxica hispánica.

En una segunda fase, se compone un cuestionario exploratorio para reunir el mayor número de variantes léxicas. Por último, se redacta un cuestionario definitivo en el que se incluyen las preguntas y posibles respuestas con el fin de que los informantes elijan la más adecuada, pudiendo además aportar otras propias.

3.3.2. Encuesta tradicional / Encuesta web

Las encuestas de Varilex hacen referencia a distintos ámbitos conceptuales que se relacionan con la vida urbana actual, como pueden ser las prendas de vestir, los instrumentos usados en la cocina, los muebles, las partes del automóvil, el material de escritorio, etc.

Para responder a los cuestionarios, el sistema de encuesta tradicional en dialectología consiste en buscar a un informante idóneo al que entrevistar, conocedor de la vida local que generalmente pertenece al ámbito rural. El encuestador realiza una entrevista directa a través de la que obtiene los datos. En este tipo de encuesta es importante la relación encuestador–encuestado. De hecho, algunas investigaciones se basaban en el criterio de utilizar un *investigador único* para garantizar uniformidad en lo que se refiere a esta relación. Este tipo de encuesta permite recoger los registros de un momento determinado «como si se tratase de una fotografía instantánea, sin perseguir más fenómenos ni corregir los datos registrados ni obtener más informaciones relevantes» (Ueda, 2011: 2–3).

La metodología de encuesta tradicional es válida para determinados estudios, según sean los objetivos, las condiciones y el ámbito de investigación que tenga cada uno. No obstante, si se pretende realizar encuestas dialectales tradicionales para analizar el léxico de todo el dominio hispánico nos encontramos inevitablemente con

dos problemas: en primer lugar, la elección de un informante idóneo debe basarse en criterios diferentes, puesto que el objeto de investigación está relacionado con el ámbito urbano y no con las tradiciones rurales. Por otra parte, debido a las exigencias temporales, financieras y metodológicas de las encuestas *in situ*, lo más probable sería que el proyecto nunca llegara a concluir o, en caso de que prosperara con éxito, el desfase temporal de los datos entre unas zonas y otras sería un obstáculo para comparar e investigar los resultados.

Por ello, Varilex decidió formar una red de investigadores en cada país, que repartirían cuestionarios a los hablantes que consideraran más adecuados en cada localidad, generalmente capitales de país y otras ciudades importantes. Así, Varilex logra llegar a informantes locales de la ciudad seleccionada, pertenecientes a un estrato socioeconómico medio o alto y con un nivel cultural equivalente. No existe límite de tiempo para cumplimentar la encuesta. Después, los datos son enviados a Tokio para ser procesados y analizados. Las encuestas utilizadas estudian el léxico variable de diferentes campos semánticos que son susceptibles de variación: prendas de vestir, instrumentos de cocina, muebles, partes del automóvil, material de escritorio, etc.

Una de las características innovadoras de Varilex, desde el punto de vista metodológico, es el empleo de un cuestionario *abierto*. En dialectología con frecuencia se afirma que el mejor cuestionario es el que se elabora una vez concluida la investigación. En este sentido, Varilex ofrece un cuestionario que puede reabrirse y modificarse según avanza la investigación. Como el proyecto realiza encuestas en series sucesivas, siempre se plantea la posibilidad de incluir nuevas variantes registradas, de manera que cuando aparecen respuestas interesantes en los cuestionarios estas se incluyen en la lista de respuestas posibles, con el objetivo de averiguar si otros informantes utilizan o no estas voces en otras localidades.

En 1999, el equipo de investigación japonés consideró que para ordenar la información, almacenarla y consultarla, era preciso utilizar Internet y una base de datos digital de libre acceso, con las ventajas que ello conlleva. A partir de ese

momento, las encuestas también podrían hacerse por Internet a mayor número de informantes. Ueda (2011) pone de manifiesto estos puntos positivos:

- 1) La preparación de cuestionarios es más sencilla.
- 2) El envío de resultados de encuestas es instantáneo por la red, lo que ahorra gastos, tiempo y seguridad en el registro de los datos.
- 3) El análisis de las respuestas es muy rápido debido a su inclusión directa en la base de datos, lo que ahorra tiempo de codificación y evita errores de manipulación.
- 4) La corrección de los errores puede hacerse de manera casi inmediata, esté relacionada con el modo de preguntar de los encuestadores o con el modo de contestar de los encuestados.
- 5) Los datos dialectales recopilados se pueden descargar y consultar, lo que permite que estos registros se utilicen de manera activa en investigaciones relacionadas con diversos ámbitos de estudio.
- 6) La colaboración a modo de red internacional es una realidad que facilita el intercambio, las críticas y las observaciones de otros expertos.
- 7) La posibilidad de utilizar técnicas de programación permite la generación dinámica de mapas a través de programas accesibles y gratuitos. Frente a los tradicionales atlas lingüísticos, Varilex propone un sistema mucho más rápido y menos laborioso, que permite elaborar distintos tipos de mapas con perspectivas y objetivos de investigación diferentes, generando, editando e imprimiendo únicamente aquellos datos que necesitemos para nuestro estudio².

Como aspecto negativo cabe destacar que la encuesta web carece inevitablemente de la relación humana encuestador–encuestado. En cambio, tiene

² Para más información sobre cartografía automática, consultar el artículo de Ruiz Tinoco (2001).

como ventaja que el contacto entre ambos continúa en el tiempo; la relación no se pierde cuando se realiza la encuesta. De hecho, se pide a los encuestados que formen parte del proyecto de diferentes maneras: un mismo colaborador puede ser informador, investigador, revisor de datos, etc., al tiempo que puede incluir notas y comentarios sobre el uso particular que él hace de las palabras estudiadas.

Otra dificultad de este tipo de encuestas surge del hecho de que cualquier hablante pueda responder al cuestionario, sin buscar informantes idóneos que respondan a un perfil específico. Esto permite contar con un mayor número de informantes potenciales, aunque exige un filtro riguroso por parte del investigador. Desde la página web, cualquier informante puede completar el cuestionario, lo que exige un control de quién ha respondido a la encuesta, de qué manera, etc. Por ello, se pide a los informantes que aporten sus datos personales y se parte de la premisa de que muy pocas personas van a tener interés por utilizar la encuesta web de manera inadecuada, puesto que se trata de una que tiene una duración aproximada de una hora. En todo caso, el control de respuestas que hacen los investigadores permite eliminar de la base de datos los registros erróneos con gran rapidez.

La encuesta web comienza con una breve introducción en la que se explican los objetivos del proyecto Varilex para analizar el uso de palabras y expresiones de la vida cotidiana de todos los hispanohablantes. Además, se invita al colaborador a facilitar su correo electrónico para enviarle estudios y publicaciones relacionadas con esta investigación en forma de agradecimiento. Se requieren, por supuesto, datos de la ciudad y origen del informante, aparte de otros datos básicos para posteriores estudios sociolingüísticos como la edad, el sexo, la profesión, etc. Los niveles de colaboración que considera Varilex en su investigación son los siguientes:

- 1) *Coordinadores del proyecto*: especialistas que tienen acceso total al sistema, los programas, el diseño de la base de datos y pueden añadir, suprimir, alterar datos, permitir o denegar acceso, etc.

- 2) *Colaboradores locales*: profesionales que de forma voluntaria y bajo la supervisión de los coordinadores pueden acceder, rectificar y añadir datos.
- 3) *Informantes en general*: añaden datos pero no pueden alterarlos ni borrarlos. No son especialistas, tan solo usuarios de lengua.
- 4) *Usuarios generales de la base de datos*: pueden consultar los datos.

Los niveles de colaboración de Varilex se representan en la siguiente figura:



Figura 5: Niveles de colaboración de Varilex

El cuestionario está dividido en diferentes partes que tratan determinados temas. Para cada pregunta de encuesta, Varilex ofrece una lista de respuestas posibles, algo que no es habitual en la dialectología tradicional para evitar el efecto de inducción en el encuestado. La opción de ofrecer las respuestas se justifica por tres consideraciones (Ueda, 1999):

- 1) Para realizar un estudio dialectal del léxico urbano, los datos deben ser cuantificables, además de ser comparables. Por ello, es necesario que los datos estén bien especificados respecto a su uso.
- 2) Para cuantificar y sistematizar el uso del léxico urbano es muy interesante conocer todas las formas que utilizan. Aportando todas las formas de variación registradas, accedemos tanto a datos

positivos como negativos, lo que nos permite estar más seguros de que la respuesta del informante incluye todas las variedades que él utiliza o conoce. Además, el informante puede añadir otras formas distintas de las que aparecen en la lista.

- 3) La elección de incluir una lista con las respuestas posibles se corresponde con un método extensivo que presenta más respuestas de las que *existen* en la realidad. La utilización de un método intensivo podría caer en el peligro de ser menos representativo de lo que es la realidad. En todo caso, el informante no va a marcar una palabra que no utiliza o que no conoce.

Antes de proceder a rellenar la encuesta, se deben completar los datos personales que relacionen las variantes lingüísticas con las características individuales de cada informante, para lo que se pide aportar: nombre y apellido, ciudad, país, sexo, edad, afiliación (universidad, colegio, jubilado, estudiante, labores de casa...), correo electrónico y un mensaje a los coordinadores.

A continuación mostramos dos ejemplos de preguntas de encuesta de Varilex. Podemos observar que además de las definiciones también suelen incorporarse imágenes, tanto en los cuestionarios tradicionales como en las encuestas web:



[A013] RIBBON: Tira de tela que se emplea para adorno de la cabeza.

(1)balerina, (2)banda, (3)bufanda, (4)cinta, (5)cinta de pelo, (6)cintillo, (7)diadema, (8)felpa, (9)vincha.

&Otro(s) _____, #No se me ocurre.

\$Comentario:

Figura 6: Pregunta [A013] de Varilex en el formato de encuesta tradicional

[A011] PYJAMAS (US:PAJAMAS): Traje de dormir, formado por pantalón y chaqueta, ambos de tela ligera.



ELIJA LA(S) FORMA(S) QUE USTED UTILIZA.

bata; el payama; el pijama; el piyama; la payama; la pijama; la piyama;

OTRA(S) FORMA(S): Si hay más de una forma, divídalas por una coma (,).

ESCRIBA SU COMENTARIO SOBRE EL USO DE ESTA(S) FORMA(S):

ENVIAR

Figura 7: Pregunta [A011] de Varilex en formato de encuesta web

Si nos fijamos en la Figura 3 observamos cómo en la parte superior se aporta la definición ‘Traje de dormir, formado por pantalón y chaqueta, ambos de tela ligera’, acompañada del dibujo del objeto y la indicación en letras mayúsculas de aquello que se le pide hacer al encuestado. Además, existe la posibilidad de incluir alguna otra forma léxica o algún comentario sobre el uso de la misma.

El apoyo en imágenes para los cuestionarios fue una innovación metodológica de Pellis en el ALI, que también se incluyó en el ALPI o en el ALGA. Las ventajas de este tipo de soporte visual en las encuestas son innegables, como

expuso Stanley M. Sapon (1962) en el *Pictorial Linguistic Interview Manual* (PLIM): reducen el habla del investigador o la pregunta, se economiza tiempo y puede ser aplicado en contextos y lugares muy diversos. Varilex, opta por combinar varios procedimientos en sus cuestionarios, porque no hay ninguno que no presente problemas o limitaciones. Se trata de que la forma de responder sea fácil e intuitiva para los informantes.

Otro aspecto fundamental de la encuesta web de Varilex es el hecho de que el informante accede a los datos que se han recogido hasta el momento, identificando en una tabla las formas que han contestado otros informantes de su país, con su correspondiente frecuencia. Este cuadro de distribución puede inducir la respuesta del informante, pero se aporta para facilitar la realización de la encuesta. En todo caso, los datos que le aparecen al colaborador que realiza la encuesta, tanto en la lista de las opciones posibles como en la tabla de distribución inferior, están siempre actualizados y organizados según la última persona que contestó la encuesta. Esto quiere decir que el sistema incluye automáticamente las nuevas respuestas de otros informantes, con el fin de corroborar si efectivamente se utilizan esas voces en una región determinada.

3.3.3. Red de encuesta

Para referirnos a la red de encuesta de Varilex es preciso, en primer lugar, destacar que el objetivo del proyecto es estudiar el léxico hispánico del mundo urbano, es decir, de las ciudades. Este es otro de los motivos que hacen que Varilex sea un proyecto de variación adecuado para nuestro estudio puesto que el léxico que se enseña en las clases de ELE es el léxico urbano contemporáneo.

Frente al interés de la dialectología tradicional por el estudio de áreas rurales, más conservadoras a nivel lingüístico (lo que permitía estudiar con mayor rigor la historia y la evolución de la lengua), Varilex estudia el léxico de las grandes áreas urbanas, que caracterizan nuestra realidad social. En todo caso, tal y como expone Thun (1991), carece de sentido establecer una frontera entre la dialectología urbana y

la dialectología rural pero es necesario tener claro que los objetivos y las condiciones de estudio son necesariamente distintos.

Por el momento, tan solo el *Proyecto de estudio coordinado de la norma lingüística culta de las principales ciudades de Iberoamérica y de la Península Ibérica*, ha emprendido el estudio lingüístico de las ciudades, aunque este se ha centrado fundamentalmente en la pronunciación de los hablantes cultos de las ciudades más importantes de España e Hispanoamérica, registrando entrevistas grabadas con un formato preestablecido que se llevan a cabo por especialistas locales en cada ciudad. Gracias a este proyecto se han publicado monografías sobre los usos lingüísticos de Buenos Aires, Bogotá, La Habana, México o Sevilla, entre otros:

El excesivo apego de la Dialectología a la investigación de áreas rurales ha tenido como consecuencia que desconozcamos casi por completo lo que sucede en las grandes áreas urbanas que son, precisamente, las que caracterizan nuestra realidad social. Con no poca razón, muchos investigadores abogan porque la Dialectología sea hoy básicamente urbana, máxime si queremos comparar lo que sucede en espacios geográficos distintos, y, para ello, será requisito indispensable el estudio de los procesos de variación sociolingüística (Bustos Gisbert, 1995: 51).

Una vez que sabemos que el interés de la investigación se centra en el habla de las ciudades es preciso delimitar la red de puntos de encuesta. Partiendo de la base de que los atlas de gran dominio no pretenden plasmar la riqueza local del territorio sino que pretenden ofrecer una imagen de continuidad, no se puede trabajar con una red demasiado espesa, puesto que dificultaría la visión global del léxico.

Para establecer una red de puntos generalmente se combinan diferentes criterios (geográficos, históricos, socioculturales...). En Varilex se estudia el léxico de las capitales, identificando los puntos con mayor densidad de población dentro de cada país. Así, además de las capitales, Varilex intenta recoger el léxico de otras ciudades representativas. Esos puntos están fijados previamente pero se pueden ampliar, como de hecho ha ido haciendo Varilex cada año. Hasta el momento, han estudiado las siguientes ciudades:

Tabla 3
Ciudades y países donde Varilex ha estudiado la variedad léxica

Código	Ciudad	País
ES-COR	La Coruña	España
ES-SCO	Santiago de Compostela	España
ES-OVI	Oviedo	España
ES-STD	Santander	España
ES-SLM	Salamanca	España
ES-ZAR	Zaragoza	España
ES-BAR	Barcelona	España
ES-GDL	Guadalajara	España
ES-MAD	Madrid	España
ES-VAL	Valencia	España
ES-SEV	Sevilla	España
ES-GRA	Granada	España
ES-MLG	Málaga	España
ES-ALM	Almería	España
ES-MUR	Murcia	España
ES-TEN	Santa Cruz de Tenerife	España
ES-PAL	Las Palmas de Gran Canaria	España
MA-TET	Tetuán	Marruecos
GE-MAL	Malabo	Guinea Ecuatorial
CU-HAB	La Habana	Cuba
CU-SCU	Santiago de Cuba	Cuba
RD-STI	Santiago	Rep. Dominicana
RD-SDO	Santo Domingo	Rep. Dominicana
PR-SJU	San Juan	Puerto Rico
PR-DOR	Dorado	Puerto Rico
PR-MAY	Mayagüez	Puerto Rico
EU-NYK	Nueva York	Estados Unidos
EU-NOR	Nuevo Orleans	Estados Unidos
MX-MON	Monterrey	México
MX-AGS	Aguas Calientes	México
MX-MEX	Ciudad de México	México
MX-MRD	Mérida	México
GU-GUA	Guatemala	Guatemala
EL-SSV	San Salvador	El salvador
HO-TEG	Tegucigalpa	Honduras
HO-NAC	Nacaome	Honduras

NI-LEO	León	Nicaragua
NI-MAN	Managua	Nicaragua
CR-LMN	Puerto Limón	Costa rica
CR-SJO	San José	Costa rica
PN-PAN	Panamá	Panamá
CO-MED	Medellín	Colombia
CO-BOG	Santafé de Bogotá	Colombia
VE-CAR	Caracas	Venezuela
VE-MER	Mérida	Venezuela
VE-VNC	Valencia	Venezuela
VE-TAC	Tachira	Venezuela
EC-QUI	Quito	Ecuador
PE-LIM	Lima	Perú
PE-ARE	Arequipa	Perú
BO-PAZ	La Paz	Bolivia
CH-ARI	Arica	Chile
CH-SCH	Santiago	Chile
CH-CON	Concepción	Chile
CH-TEM	Temuco	Chile
PA-ASU	Asunción	Paraguay
UR-MTV	Montevideo	Uruguay
AR-SAL	Salta	Argentina
AR-SJN	San Juan	Argentina
AR-NEU	Neuquen	Argentina
AR-BUE	Buenos Aires	Argentina
FI-MNL	Manila	Filipinas
FI-ZBO	Zamboanga	Filipinas

3.3.4. Encuestadores

A diferencia de un cuestionario de entrega directa a los informantes, el proyecto Varilex opta por crear una red de colaboradores locales que son los encargados de distribuir los cuestionarios entre los hablantes de cada ciudad. La ventaja que supone contar con esta red de apoyo se traduce, sobre todo, en un elevado registro de respuestas, que garantiza la continuidad del proyecto.

Los colaboradores locales son los encargados de elegir a los informantes, tarea de excepcional importancia para obtener conclusiones fiables y representativas

sobre el uso lingüístico. Además, cada medio año los colaboradores deben responder a unas doscientas preguntas de contestación libre y a otras de respuesta selectiva. Pueden incluir sus comentarios sobre el uso de las palabras. También pueden aportar nuevas palabras diferenciales o ideas metodológicas diversas que pudieran resultar útiles para el estudio.

3.3.5. Informantes

Como ya hemos apuntado, uno de los cambios más importantes de la dialectología moderna frente a la dialectología tradicional es el perfil de los informantes seleccionados. Si bien en los estudios tradicionales se buscaba hablantes mayores, preferiblemente varones, de ámbito rural y con un nivel cultural básico –el denominado perfil NORM de Chambers y Trudgill (1994): *Non–mobile, Older, Rural, Males*– en Varilex el objetivo es radicalmente distinto. Los informantes que arrojarán datos sobre el léxico urbano hispánico actual serán hombres y mujeres, residentes de una ciudad, con un nivel sociocultural medio alto.

Los colaboradores locales seleccionan un mínimo de cuatro informantes –dos mujeres y dos hombres– de entre 18 y 39 años y de entre 40 y 80. Los informantes no tienen un tiempo limitado para contestar; cuando terminan la encuesta, la devuelven al colaborador local, que es el encargado de enviar los datos a la universidad de Tokio.

Si atendemos a las respuestas de los informantes a las encuestas, observamos que la mayoría escoge una respuesta de la lista y, ocasionalmente, aporta una voz léxica nueva. Los informantes aportan una media de 1,2 respuestas por pregunta. Existe, no obstante, una minoría que considera detenidamente las respuestas y que llega a ofrecer tres, cuatro y hasta cinco respuestas extra de media a lo largo de la encuesta.

Respecto a las diferencias observadas según el tipo de informantes, Ruiz Tinoco (2008) concluye que hombres y mujeres de ámbito urbano no presentan grandes diferencias. En los datos analizados, la edad tampoco resulta ser un factor determinante del uso lingüístico, aunque sí se aprecian diferencias de sexo y de edad

si se compara la distribución del léxico en campos semánticos específicos. En el estudio de Ruiz Tinoco (2008) hay algunos vocablos entre los veinte más frecuentes que no aparecen en la lista del sexo opuesto. En lo que respecta a la edad, puede haber palabras menos frecuentes en personas mayores que en jóvenes, como lo que sucede en el mundo de la tecnología, aunque las diferencias no son significativas.

Todos estos factores son determinantes para reflexionar sobre la metodología de las encuestas y considerar los factores que pudieran afectar a los resultados de la investigación.

4. La base de datos Varilex como recurso de investigación

4.1. Características

Desde los inicios de las investigaciones de Varilex, el proyecto ha ido necesitando acumular cada vez más datos, con las dificultades de procesamiento que ello conlleva. En un primer momento, los datos se procesaban en PC con programas específicos para el proyecto diseñados por Hiroto Ueda.

Desde 1999, Varilex empezó a funcionar con una base de datos digitalizada a través de Internet, debido a la ingente cantidad de registros. Esta fecha supuso un cambio de enfoque en el proyecto, introduciendo Internet como componente fundamental para la recopilación y consulta de datos. A partir de entonces, los datos registrados se vuelcan en Internet, quedando a disposición de cualquier interesado y facilitando el acceso libre y gratuito a los registros. Así, cualquier investigador puede realizar consultas a través de comandos para realizar búsquedas complejas.

Para ello se instaló un servidor en la página <http://varilex.call.sophia.ac.jp/> con el sistema operativo LINUX para PC (kernel 2.2.), un programa de administración de base de datos con lenguaje SQL, MySQL ver. 3.22.15, junto con PHP3 ver. 3.0.6. Además el proyecto Varilex utiliza otro servidor experimentalmente, con el sistema operativo Linux, el servidor Apache, el motor de base de datos MySQL y el lenguaje de programación PHP.

Las decisiones técnicas sobre la base de datos del proyecto se justifican por criterios como la estabilidad, la facilidad de uso, la posibilidad de incluir caracteres de la lengua española o su disponibilidad sin coste.

4.2. Estructura

Para desarrollar nuestro análisis de la variación lingüística hemos tenido que manejar la base de datos de Varilex, que está formada por cuatro tablas principales interrelacionadas, que se refieren a informantes, ciudades, respuestas y conceptos. En ellas existen algunas redundancias que no afectan en ningún caso a las consultas pero que facilitan la sistematización de los datos.



Figura 8: Tabla interrelacionadas de la Base de Datos de Varilex

La tabla de *informantes* recoge la siguiente información:

- *código*: referencia del informante, que coincide con el campo *informante* de tabla de *respuestas*.
- *codciu*: código de ciudad.
- *país*: código del país.
- *ciudad*: código de ciudad.

- *sexo*: 1 hombre, 2 mujer.
- *edad*: edad del informante.
- *tipo*: tipo de ocupación del informante.
- *ocupación*: descripción de la ocupación del informante.
- *id*: número de orden dentro de la tabla.

La tabla de **ciudades** incluye:

- *id*: relaciona esta tabla con la de informantes.
- *ciudad*: también incluida en informantes.
- *x, y*: posición de la ciudad dentro del mapa blanco que se usa de base para la cartografía automática.

La tabla de **respuestas** incluye:

- *id*: número de identificación del dato.
- *encuesta*: código, que se corresponde con *código* de la tabla concepto.
- *informante*: código del informante, que relaciona esta tabla con la de informantes.
- *forma*: respuesta dada por el informante.

La tabla de **conceptos** se refiere a:

- *código*: que permite ordenar los conceptos utilizados; tiene una letra que indica el orden de la encuesta y tres números que ordenan el concepto utilizado.
- *inglés*: nombre del concepto en inglés a modo de referencia neutra para representar el ámbito conceptual.
- *concepto*: descripción verbal del ámbito conceptual en cuestión para el que buscamos variaciones léxicas; Varilex no colecciona sinónimos en el sentido tradicional del término sino variaciones léxicas que responden que se derivan de una definición o imagen en las encuestas.
- *id*: número de orden para las entradas.

4.3. Métodos de búsqueda

Los datos de las encuestas se procesan en una base de datos en un servidor experimental de entorno LAMP (LAMP: Linux como sistema operativo; Apache como servidor web; MySQL para administrar la base de datos y PHP como lenguaje de programación).

La información correspondiente a los detalles de las encuestas se almacena en estas tablas. Se puede acceder a esta información a través de diferentes interfaces virtuales, que son creadas con lenguaje PHP y el comando SELECT en MySQL. Cualquier investigador interesado en los datos no necesita ser experto en estos lenguajes informáticos puesto que se pueden obtener las informaciones seleccionando iconos de manera intuitiva. Para usuarios de MySQL, existen otras interfaces más complejas que permiten realizar búsquedas más sofisticadas. La estructura actual se encuentra en un proceso de mejora y futura ampliación.

Indudablemente la utilización del lenguaje SQL de bases de datos relacionales resulta muy útil para este tipo de estudios aunque «echamos en falta la existencia de *software* especializado para el tratamiento de los datos de variación léxica o de dialectología en general» (Ruiz Tinoco, 2008: 88).

4.4. Cartografía automática

En la página web se puede obtener el mapa de distribución de uso de la forma en cuestión. Es muy importante destacar que los mapas que se obtienen en las búsquedas no están almacenados sino que se generan dinámicamente cuando realizamos la consulta. En este sentido, el sistema supone una gran ventaja puesto que los datos se pueden consultar a lo largo del desarrollo del proyecto, observando cambios, variaciones, incluso añadiendo correcciones y revisiones en caso de que los datos presenten alguna contradicción. Todo esto era imposible en los atlas lingüísticos en papel, que se consultaban una vez estaban terminados, sin editarse, corregirse, ni añadir más apreciaciones una vez impresa la versión definitiva.

Varilex utiliza un tipo de cartografiado simple, aunque se podría emplear la tecnología SIG para incluir datos de concentraciones urbanas, clima, altura, ingresos, zonas bilingües o multilingües, etc.

5. Puntos fuertes y limitaciones de Varilex

Después de más de veinte años desde la preparación de las bases del proyecto, podemos confirmar que Varilex ha continuado creciendo y ampliándose, incluso ha generado nuevos proyectos como el *Corpus de variación del español* (Covares), que pretende reunir un corpus digitalizado de textos de todas las áreas hispanohablantes.

Desde los años 90, Varilex ha trabajado tanto la dimensión lingüística del proyecto como la dimensión geográfica. Desde la primera encuesta, en la que se analizaron 219 unidades léxicas en 25 ciudades, Varilex ha llegado a estudiar más de 1350 unidades léxicas (de las que derivan aproximadamente 11 200 variantes) en 63 localidades. El aumento de los conceptos estudiados y de las localidades se ha ido ampliando poco a poco, incorporando nuevos puntos a la red de encuesta.

A lo largo del proyecto se han realizado 11 encuestas diferentes a un total de más de 1700 informantes, que se centran en los siguientes campos semánticos:

Tabla 4
Encuestas de Varilex con los correspondientes campos semánticos estudiados.

Encuesta	Campos semánticos
A	Prendas de vestir, casa, comida, transporte.
B	Comida, casa, prendas de vestir, escritorio, ocio–deporte, transporte, ciudad.
C	Ocio–deporte, naturaleza, comida, español general, ciudad, escuela, transporte.
D	Español general, escuela, comida.
E	Español general, ciudad, ocio–deporte.
F	Español general.
G	Gramática.
H	Gramática.
I	General, naturaleza, ocio–deporte, comida.
J	Gramática.
K	Colores.

Existen ciertas limitaciones que observamos en Varilex y que debemos tener en cuenta a la hora de realizar nuestra investigación. Un problema importante está relacionado con las equivalencias biunívocas entre dialectos. Hay que tener en cuenta los valores connotativos y denotativos de cada voz. Así, el hecho de utilizar palabras distintas dentro de comunidades lingüísticas diferentes no solo responde a un contraste diatópico; también puede ocurrir que dos voces sean diferentes porque se refieran a realidades diversas o porque tengan un valor específico según las regiones. Toda variación se puede explicar por diversos factores, más allá del componente diatópico. No obstante, en la mayoría de los casos, existen voces léxicas variadas en las diferentes ciudades encuestadas, sin otras implicaciones que las geográficas. Es decir, que en la mayoría de los casos, las voces se oponen connotativamente desde un

punto de vista diatópico, existiendo una correspondencia biunívoca entre los valores de las formas léxicas.

No obstante, como señala Ávila (1997), hay otros casos en los que no se produce esta variación diatópica simple, por lo que resulta necesario atender al componente *connotativo* de las formas léxicas –teniendo en cuenta las variaciones sociolingüísticas que puedan existir entre las voces– así como al componente *denotativo* –observando las relaciones lingüísticas que se establecen dentro de cada dialecto. Todo ello requiere un análisis más detallado de la variedad. Existirán, por tanto, casos en los que voces distintas se refieren a un mismo concepto, presentando variaciones relacionadas con el grupo de edad, el sexo, la frecuencia, las marcas de uso (anticuado, rural, juvenil...), el estrato social o el registro. Al mismo tiempo, podemos encontrar voces idénticas que se refieren a conceptos distintos.

En el proyecto Varilex se pretende solventar este problema ofreciendo a los hablantes una definición y un dibujo que explica el significado de las voces que se pretenden registrar. Las variaciones tienen una interpretación. Además de presentar una lista de voces más o menos frecuentes en los diferentes países hispanos, es necesario describir cómo se articulan estas variaciones dentro del sistema lingüístico de cada dialecto, de manera que sea posible realizar comparaciones interdialectales. Este tipo de análisis exhaustivo acerca del uso específico de los términos lingüísticos exigiría realizar un estudio pormenorizado de las nociones estudiadas por Varilex, objetivo difícil de asumir debido a la larga lista de voces analizadas.

No obstante, es necesario mantener esta mirada escéptica para no dar un valor absoluto a nuestros resultados. Creemos que las encuestas de Varilex procuran evitar este tipo de cuestiones, incluyendo definiciones precisas exhaustivas acompañadas, en muchos casos, de ilustraciones que permiten disipar las dudas del colaborador a la hora de responder. De hecho, Varilex tiene el mismo enfoque que otros proyectos de estudio léxico como *El estudio coordinado de la norma lingüística culta*. No obstante, en algunos casos concretos encontraremos, ligada a la variación diatópica,

elementos connotativos o denotativos que explica la existencia de determinadas variaciones.

En todo caso, tal y como señala el propio Ávila (1997) en su investigación sobre la connotación y la denotación en los datos de Varilex, tengamos presente que en la mayoría de los casos analizados, a un determinado significado le corresponde una forma distinta en cada ciudad, tratándose exclusivamente de variantes o alternancias geográficas.

Respecto al *efecto de inducción* de los colaboradores, que ya hemos comentado, no encontramos una justificación sólida para destacar en otro color las respuestas que ya han sido contestadas por otros hablantes de la misma región geográfica, aunque Varilex lo justifica para facilitar a los informantes la realización de las encuestas.

Por otra parte, el programa de Internet que se utiliza hoy día está obsoleto. Se podría actualizar con una base de datos que permitiera *utilizar la página de manera interactiva* por parte de los investigadores: incorporando datos, apreciaciones o correcciones. Esta plataforma también debería permitir ir publicando las investigaciones relacionadas con este proyecto de manera clara y precisa, para que el investigador conozca en qué estado se encuentra el proyecto y las investigaciones que de él se derivan.

Sería interesante procesar una *mayor cantidad de datos* a través de nuevas páginas web que permitan las consultas lingüísticas a tiempo real, con el objetivo de ofrecer información actualizada a cualquier especialista interesado en el estudio del léxico.

Además, los datos que surgen de Varilex deberían *compararse* con los datos extraídos de *otros proyectos* de investigación. En el capítulo 6 de nuestra tesis, realizaremos una comparación de los datos de Varilex y los del *Diccionario de americanismos* para proponer una ampliación del léxico del *PCIC*, lo que permitirá contrastar ambos materiales y comprender las diferencias entre ambos. Así, el hecho de que algún dato no aparezca en las encuestas de Varilex pero sí se recoja en otras

fuentes de consulta fiables sobre la variación léxica, debería analizarse. Podría ocurrir que, al tratarse de una encuesta de tipo selección, alguna respuesta que no se ofrezca de antemano no llegue a registrarse, a pesar de proporcionar una casilla en la que el colaborador puede incluir otras palabras.

Para lograr una mayor *representatividad* de los datos, se podría trabajar en estas dos líneas:

- 1) Establecer más contactos en algunas ciudades importantes de Hispanoamérica puesto que hay países en los que solo se ha tenido en cuenta una localidad en la investigación. Es el caso de Guatemala, El Salvador, Costa Rica, Panamá, Ecuador, Bolivia, Paraguay, Uruguay y Guinea Ecuatorial.
- 2) Los conceptos estudiados podrían ampliarse y llegar, al menos, a 2000.
- 3) El hecho de que los cuestionarios se expandan cada vez con más conceptos que se añaden a la base de datos, hace necesario que la participación de los usuarios sea mayor. Es preciso valorar cómo podemos invitar a los informantes a colaborar.

A pesar de estas limitaciones, los logros del equipo de Varilex son innegables. El proyecto, que se inició con el objetivo de conocer los cambios léxicos diacrónicos y diatópicos, ha logrado contribuir al desarrollo de la metodología de la investigación lingüística geolectal. Asimismo, ha supuesto una fuerte innovación para la geografía lingüística, aportando una base de datos abierta, libre, gratuita, relacional, flexible y constantemente actualizada.

El tratamiento estadístico de los datos que hace el equipo Varilex es, sin duda, convincente. Además, se adopta una posición ecuaníme ante las variedades de español: el término que encabeza las preguntas es el más frecuente.

6. Conclusión

Para desarrollar esta investigación hemos utilizado los datos de distribución léxica recopilados por el proyecto Varilex. El estudio sistemático del léxico que ha desarrollado este grupo de investigación tiene como objetivo mostrar la diversidad léxica del español actual urbano. Un proyecto de tales dimensiones se basa en una red de cooperación internacional que ha permitido reunir datos de variación en los veinte países hispanohablantes. Así, el proyecto Varilex propone un trabajo sistemático de la variación, que no se limita a presentar generalidades de vocabulario sino que estudia, con una metodología rigurosa, la distribución de los elementos léxicos de forma unificada. Como proyecto lingüístico y geográfico, Varilex renuncia a retratar las peculiaridades regionales para presentar las características de una lengua común en el mundo hispánico.

En este capítulo hemos explicado la metodología de este proyecto, diferente de la utilizada en la dialectología tradicional. Los objetivos del proyecto condicionan también la elección de los informantes, escogiendo tanto a mujeres como a hombres de cualquier edad, que residan en ciudades y cuenten con un nivel socioeconómico medio alto. También nos hemos detenido en las diferentes etapas y fases del proyecto.

Desde 1993, Varilex ha recogido datos léxicos mediante encuestas tradicionales y encuestas en línea. Hemos visto cómo estas encuestas tienen un carácter flexible y abierto, por lo que se pueden modificar a lo largo del tiempo. La posibilidad de realizar encuestas en línea ha permitido acceder a mayor número de usuarios, aunque exige un control más estricto de los informantes. Por otra parte, la digitalización del proyecto ha permitido: preparar cuestionarios de forma más sencilla, enviar los resultados con rapidez, incluirlos directamente en la base de datos ahorrando tiempo de codificación y fallos de manipulación, corregir errores, habilitar la descarga por parte de cualquier usuario y generar mapas con los resultados que nos interesen.

Así, toda la información de Varilex se almacena en un banco de datos léxicos de libre acceso que hemos analizado en este capítulo. La creación de un proyecto internacional con un archivo libre, abierto, relacional y constantemente actualizado supone un cambio de paradigma en la investigación dialectológica. Esta base de datos hace posible desarrollar estudios dialectológicos, lexicográficos, taxonómicos, históricos o cartográficos. De hecho, los cálculos demolingüísticos que hemos propuesto en nuestra investigación no habrían sido posibles sin el acceso a este material. Por otra parte, es imprescindible tener en cuenta algunas limitaciones de los registros, como la falta de correspondencia biunívoca entre algunos datos analizados.

Para mejorar el alcance del proyecto, se deberían incrementar los conceptos estudiados así como la red de puntos de encuesta, sobre todo en los países que solo han registrado el léxico de una única ciudad. También sería interesante que en la página web de Varilex se pudieran comparar sus resultados con los de otros proyectos de investigación léxica. En todo caso, los logros de Varilex son innegables y han supuesto una enorme contribución al desarrollo de la metodología lingüística. El tratamiento estadístico de los datos resulta convincente y presenta la realidad del léxico del español actual en todo el dominio hispánico.

Por todo ello, consideramos que los datos de Varilex son concluyentes para nuestra investigación. El tratamiento de los mismos nos permite estudiar la representatividad de las voces, mediante un cálculo ponderado que tiene en cuenta los resultados de las encuestas, además de los datos demográficos de cada país. Como veremos en el siguiente capítulo, el cálculo de la representatividad de las variantes léxicas es fundamental para seleccionar el léxico de la clase de ELE, que debe cumplir los requisitos de rentabilidad y extensión. Gracias a los datos de Varilex, en esta tesis proponemos un cálculo demolingüístico para medir la variable latente de la representatividad, lo que nos permitirá proponer una ampliación léxica fundamentada de las *Nociones específicas* del PCIC.

CAPÍTULO 4

PROPUESTA DE ANÁLISIS DEMOLINGÜÍSTICO DE LA VARIACIÓN LÉXICA

1. Introducción

Como ya se ha expuesto en los capítulos anteriores, la presente investigación tiene el objetivo de comentar y hacer una propuesta de ampliación del léxico que se presenta en las «Nociones específicas» del *PCIC*. Nuestra hipótesis inicial es que, a la luz de los recientes estudios geolectales, pueden añadirse marcas dialectales en este importante material de referencia del español como lengua extranjera o segunda lengua. Para ello vamos a desarrollar un estudio demolingüístico que permita considerar la extensión de los términos de referencia, integrando variantes léxicas verdaderamente representativas, extensas y rentables del español actual.

Consideramos que el establecimiento de grados cuantitativos de representatividad en las voces permite identificar con rigor el alcance y la extensión del vocabulario del *PCIC*. La publicación de un repertorio léxico representativo de la distribución del léxico del español es crucial para orientar el tratamiento de la variedad léxica en la enseñanza de ELE.

En este capítulo explicaremos nuestra propuesta demolingüística para medir la extensión y la frecuencia de las voces, en función de factores demográficos y lingüísticos. Para medir un concepto multidimensional como la representatividad, que constituye la variable latente de esta investigación, necesitamos construir indicadores que puedan expresarse numéricamente. A partir de las informaciones recogidas por el grupo de investigación de Varilex, realizamos un procesamiento original de los datos que permite operacionalizar la representatividad de las voces en función de las respuestas de los informantes en cada país hispanohablante. Asignamos así un número específico que determina la representatividad, de manera que podamos medir y ordenar el léxico en función de su extensión y uso. Nuestro cálculo demolingüístico permite identificar las voces panhispánicas o americanas más representativas del español y, a partir de ahí, hacer una propuesta para integrarlas en el *PCIC*.

En los siguientes apartados enmarcaremos, en primer lugar, nuestra investigación, definiendo el origen, los objetivos y las hipótesis de trabajo. A continuación presentaremos nuestra propuesta demolingüística para el análisis de la representatividad. Por último, explicaremos los aspectos metodológicos específicos, describiendo el tipo de estudio, los materiales de trabajo, así como los datos analizados y las técnicas empleadas para su recolección y procesamiento.

2. Origen de la investigación

Esta investigación nace de la conciencia de la dificultad que experimentan los profesores de español para determinar qué vocabulario es, a la vez, representativo, general y rentable, para un aprendiz de ELE. La ausencia de investigaciones sobre la distribución geográfica de la variación léxica hace que el tratamiento riguroso de la misma sea verdaderamente complejo. El docente, diseñador curricular o autor de materiales de español, además de incorporar elementos lingüísticos de diferentes variedades dialectales, también debe plantearse la secuenciación en niveles de aprendizaje.

Si acudimos al *PCIC* como obra de referencia para la enseñanza de español, observamos que el capítulo dedicado al léxico carece de indicaciones de variación geolectal. Reconocemos que otras secciones del *PCIC* sí incluyen rasgos dialectales diversos, pero el repertorio léxico propuesto por el Instituto Cervantes acusa la ausencia de voces pertenecientes a otras variedades diferentes del español peninsular. El carácter abierto de este inventario permite hacer ya, pasada una década de su publicación, una ampliación de léxico representativo que cumpla con criterios de rentabilidad y extensión eficaces para el currículo, los manuales, los exámenes o la preparación de las clases.

Nuestra tesis pretende ofrecer una respuesta razonada y documentada a la ausencia de bibliografía específica, admitiendo la complejidad del tratamiento del léxico a nivel diatópico, pero tratando de desbloquear la situación actual que impide aproximarse con rigor a la variación. Para ofrecer nociones altamente representativas del español vamos a definir matemáticamente el concepto de *representatividad*, asignando a las voces léxicas un Grado de Representatividad (GR) específico, que nos permita realizar comparaciones entre las variantes e identificar aquellas que sean más representativas.

3. Hipótesis y objetivos de investigación

En esta investigación vamos a analizar el léxico incluido en el inventario de «Nociones específicas» del *PCIC*. Las hipótesis que guían nuestro estudio y que permiten profundizar en la interpretación de este documento son las siguientes:

- H1: Las nociones léxicas del *PCIC* son una muestra del español de España que puede complementarse con voces representativas de otras variedades de español.
- H2: En el *PCIC* existen variantes léxicas que son exclusivas del español peninsular (españolismos).

- H3: En el léxico del español podemos encontrar variantes panhispánicas que son comunes a una mayoría de hispanohablantes (panhispanismos) que deben formar parte del léxico de ELE.
- H4: En el léxico del español existen variantes comunes a las regiones de Hispanoamérica que no existen en el español peninsular (americanismos) pero que deben formar parte del léxico de ELE.
- H5: La identificación de las variantes léxicas definidas en H2, H3 y H4 permite proponer una ampliación del *PCIC* representativa del español urbano actual.

Así, el objetivo principal de esta investigación es realizar una propuesta de ampliación de las «Nociones específicas» del *PCIC* que sea representativa de la diversidad geolectal del español. Nuestro propósito es triple: por un lado, vamos a proponer voces representativas de la totalidad de la geografía hispánica; por otro, señalaremos aquellas voces americanas que cuentan con gran rentabilidad y extensión; por último, marcaremos aquellas variantes léxicas que son exclusivas del español de España.

Para lograr este objetivo central, hemos establecido otros objetivos específicos que marcan el planteamiento de nuestro trabajo:

- Proponer una fórmula matemática que permita clasificar las voces en función de su extensión y uso en 20 países hispanohablantes.
- Calcular el Grado de Representatividad (GR) de 182 nociones del *PCIC* y de sus alternativas léxicas correspondientes.
- Diseñar un cartografiado dinámico de los datos que permita la lectura e interpretación de nuestros resultados demolingüísticos.
- Proponer un conjunto de panhispanismos y americanismos con alto GR que sean representativos del léxico del español y que puedan incorporarse al *PCIC*.

- Identificar las nociones del *PCIC* que se utilizan exclusivamente en España (españolismos) para marcarlas en nuestra propuesta de ampliación.

Para desarrollar nuestras hipótesis y lograr los objetivos de investigación, proponemos las siguientes actividades:

- Revisar otras propuestas sobre *representatividad* léxica y establecer una comparación de los resultados demolingüísticos, en función de los diferentes análisis de representatividad.
- Procesar los datos de Varilex en una base de datos específica de esta investigación con formato MySQL que nos permita implementar el concepto matemático de representatividad.
- Incorporar los datos de estimación demográfica más recientes para cada uno de los países analizados a nuestra base de datos MySQL, a partir de los datos publicados por el Instituto Nacional de Estadística (INE) de España.
- Definir las áreas de zonificación dialectal con las que estudiaremos la dispersión de las voces léxicas.
- Valorar diferentes sistemas de representación cartográfica de datos espacio-temporales y seleccionar el que mejor represente nuestros resultados.
- Comparar nuestros datos de representatividad con los datos del *Diccionario de americanismos* y contrastarlos con su índice de frecuencia según corpus reales de la lengua española.

4. Metodología

4.1. Tipo de estudio

En esta investigación estudiamos el léxico del español urbano en 19 países hispanohablantes. Partimos de una aproximación onomasiológica, puesto que

tomamos como punto de partida un concepto —con un significado único— para el que existen varios significantes, según las zonas geográficas a las que nos refiramos. Nuestra investigación se enmarca, por tanto, en lo que Baldinger (1964) llamó el *campo de las designaciones*.

La metodología onomasiológica, inaugurada por Tappolet (1895) y posteriormente desarrollada por Zauner (1903), fue la que siguieron muchos atlas lingüísticos, cuyo propósito coincidía con el nuestro: compilar las variantes léxicas de determinados conceptos en las diferentes áreas geográficas.

Así, en el marco de la geografía lingüística, hemos desarrollado un estudio transversal de la variación léxica para el que utilizamos los datos almacenados por el grupo de Varilex desde el año 1993 hasta la actualidad. Pretendemos mostrar, desde una perspectiva sincrónica, una imagen global del léxico actual, sin detenernos en la evolución específica de cada fenómeno lingüístico.

A continuación, presentamos un estudio descriptivo de 182 conceptos del español a los que les corresponden más de un millar de variantes léxicas. No podemos hablar de un estudio analítico porque no buscamos inferencia causal ni contamos con un número de datos suficientemente representativos de la población de cada país. Partiendo de las investigaciones de Varilex, que ponen de manifiesto la existencia de equivalencias léxicas para determinados conceptos, analizaremos dónde y con qué frecuencia se distribuyen esas voces en la geografía hispánica, de manera que sea posible ampliar las nociones del *PCIC* con otras voces equivalentes que sean altamente representativas del español. En los casos en los que se registran dos voces en un mismo país, nos ocuparemos de las diferencias de frecuencia así como de las diferencias connotativas o denotativas que se pudieran derivar.

Desde el punto de vista teórico, esta tesis aporta un acercamiento demolingüístico diferente de los planteados en otros estudios de variación, pues mide la cantidad de respuestas registradas en cada país en relación al número total de informantes encuestados y a los datos demográficos de la población. En todo caso, nuestra investigación es esencialmente aplicada, ya que permite valorar y ampliar las

voces léxicas de las «Nociones específicas» del *PCIC* en función de su representatividad.

4.2. Cálculo demolingüístico

Uno de los objetivos fundamentales de esta tesis es realizar un análisis demolingüístico que permita determinar el grado de representatividad de las variantes del léxico hispanico. De este modo, obtenemos el porcentaje de uso de cada variante léxica dentro de toda la comunidad lingüística¹ del español.

Para realizar este análisis nos basamos en dos aspectos. Por un lado, tenemos en cuenta la distribución de las formas léxicas, es decir, qué variantes léxicas se han registrado en cada país. Por otro, consideramos los datos demográficos de los países analizados, que informan del número de hablantes potencial de cada voz en el conjunto de la geografía hispanica.

El profesor Ueda (1996c) propone un cálculo demolingüístico que permite hallar un índice de representatividad de las voces léxicas en función de la dispersión geográfica de las mismas. Para ello, propone la siguiente fórmula:

$$gr = \frac{\sum_i n_i P_i}{\sum_i P_i}$$

donde i = lista de países encuestados; n_i = número de países; = 1 (cuando se presenta la voz en el país en cuestión), = 0 (cuando no se presenta la voz); P_i = población de cada país.

Según Ueda (1996c), la Tabla 5 mostraría la distribución geográfica del concepto JACKET, aplicando su cálculo demolingüístico, como vemos a continuación:

¹ Entendemos por *comunidad lingüística* «el conjunto de hablantes que utilizan una misma lengua para comunicarse entre sí en un momento y en un lugar determinado. Por ejemplo, los hispanohablantes de Madrid y La Habana forman parte de una misma comunidad lingüística» (Andión y Casado, 2014).

Tabla 5
Distribución del concepto JACKET según el análisis demolingüístico de Ueda (1996)

Concepto	Voz	N ° países	Suma de habitantes	% Total	Países
JACKET	americana	1	37.746.260	14,03 %	Es
	chaqueta	9	125.391.840	46,61 %	Es, EL, Cu, PR, Ec, Co, Ve, Pa, Ch
	chaleco	1	9.896.971	3,68 %	Cu
	gabán	1	3.196.520	1,19 %	PR
	saco	15	210.275.616	78,17 %	Mx, Gu, EL, CR, Pn, Cu, RD, PR, Co, Ve, Pe, Bo, Pa, Ur, Ar
	vestón	1	11.329.736	4.21 %	Ch

Según estos cálculos, *saco* es la voz más representativa en el territorio hispanico, puesto que se registra en 15 de los 20 países y representa un 78 % del total de hispanohablantes. La voz *chaqueta* ocupa la segunda posición en términos de representatividad puesto que se registra en 9 países, con un total de 125 millones de hablantes, que constituye un 46 % respecto al total de la geografía hispanica. Por el contrario, la voz *americana*, que se registra exclusivamente en España, tiene una representatividad considerablemente inferior, atribuyéndole un 14 % dentro del ámbito hispanohablante, puesto que representa a 38 millones de personas respecto a la totalidad. Según el mismo razonamiento, la voz menos representativa es *gabán*, que se registra tan solo en Puerto Rico, donde existen únicamente 3 millones de hablantes, lo que le otorga un 1 % de representatividad respecto al total de hispanohablantes.

Sin embargo, los cálculos de representatividad planteados no tienen en cuenta el número de personas que contestan una determinada forma en cada país. Dicho de otro modo, este cálculo demolingüístico propone considerar la representatividad de

cada forma léxica, en función de si la voz se ha registrado o no en cada país hispanohablante. En nuestra investigación hemos tenido acceso a cada una de las respuestas registradas por Varilex en cada país, así como al número total de informantes encuestados. Estos datos nos han permitido ajustar el porcentaje de representatividad.

En este sentido, nuestra propuesta consiste en realizar un cálculo demolingüístico ponderado que tenga en cuenta el número de personas que han respondido a cada forma. Si retomamos el ejemplo del concepto JACKET, tanto *chaqueta* como *americana* se registran en España, pero no es lo mismo que 102 personas, sobre un total de 166 informantes, respondan la voz *chaqueta*, que 58 personas sobre 166 respondan la voz *americana*. Aunque ambas voces se registren en un mismo país, es importante marcar o precisar de alguna manera que *chaqueta* es más representativa del español de España que *americana*.

Por ello, en nuestra propuesta de análisis demolingüístico vamos a tener en cuenta la cantidad de personas que han marcado una determinada forma léxica en relación con el total de personas encuestadas de ese país. Este ajuste se consigue mediante un cálculo ponderado que permite ofrecer un resultado más ajustado de la extensión y uso de las voces, que se traduce en términos de representatividad. Así, en esta tesis proponemos un Grado de Representatividad ponderado, que denominamos GR, con el objetivo de analizar minuciosamente la representatividad de las voces y superar un análisis más generalista que solo tenga en cuenta si una forma se registra o no en un país determinado. Este tipo de análisis es el más recomendable cuando contamos con muestras pequeñas de población puesto que nos permite valorar no solo si las voces se emplean en un país, sino en qué medida se utilizan. Para calcular el GR proponemos la siguiente fórmula:

$$g.r. ponderado = \frac{\sum_i P_i \cdot ip}{\sum_i P_i}$$

$$ip(\text{índice de ponderación}) = \frac{\text{ocurrencias}}{n.\text{informantes}}$$

donde i = lista de países encuestados; P_i = población de cada país; ip = relación entre ocurrencias de una determinada forma y el número de informantes encuestados en un determinado país.

Según nuestro cálculo demolingüístico, coincidimos con Ueda (1996) en que *saco* y *chaqueta* son las voces con mayor GR para el concepto [JACKET]. No obstante, consideramos que nuestro cálculo ofrece una visión más ajustada de la representatividad. Si tenemos en cuenta el número de respuestas que se han registrado en cada país en relación a la población total, observamos que *saco* (GR = 57,87) sigue siendo la variante más representativa; la voz *chaqueta* (GR = 21,58) aparece con mayor distancia de la variante principal *saco*, y el resto de variantes pasan a ser mucho menos representativas tal y como observamos en *americana* (GR = 5,35), *traje* (GR = 1,54) o *vestón* (GR = 1,36), que solo tendrán sentido en contextos de aprendizaje específicos.

Otro ejemplo en el que apreciamos cómo pueden variar los resultados demolingüísticos en función de la escala lo encontramos en el concepto [BLACKBOARD]. El cálculo ponderado permite ajustar mejor el uso de las voces, mostrando, por ejemplo, que *encerado* es una voz minoritaria local muy poco representativa del español y *pizarrón*, en cambio, es la voz más extendida, con un protagonismo considerablemente mayor del que apuntaba el GR general:

Tabla 6
Diferencias del análisis demolingüístico de [BLACKBOARD] con GR general y GR ponderado ordenados de mayor a menor

Concepto		GR general		GR ponderado
[BLACKBOARD]	pizarra	77,87	pizarrón	55,12
<i>Placa que se</i>	pizarrón	71,84	pizarra	30,25
<i>emplea para</i>	tablero	24,52	tablero	11,76
<i>escribir</i>	encerado	20,49	encerado	2,86

Otro ejemplo del ajuste que supone nuestra fórmula revisada de representatividad lo observamos en la distribución del concepto [BEER]. Las voces *cheves* y *birra* ni siquiera aparecen entre las cuatro voces más

representativas si no ponderamos los resultados analizados, lo que nos lleva a afirmar que el cálculo de un GR general puede distorsionar ampliamente los resultados demolingüísticos:

Tabla 7

Diferencias del análisis demolingüístico de [BEER] con GR general y GR ponderado ordenados de mayor a menor.

	GR general		GR ponderado	
Concepto [BEER]	cervecita	74,49	chela	23,99
¿Qué otro nombre	fría	61,40	birra	15,20
tiene la cerveza?	amarga	53,52	cervecita	9,83
	chela	49,38	cheves	7,97

Como se puede apreciar en los ejemplos propuestos, con un cálculo demolingüístico ponderado acotamos mejor los datos de representatividad, puesto que no solo tenemos en cuenta que una forma se emplea en un país determinado, sino que además medimos cuántas personas han contestado esa voz en relación al número total de encuestados.

Otra de las ventajas de realizar un cálculo ponderado es que solucionamos el problema de los hápax o registros únicos. Ueda (1999) se inclina por la eliminación de las formas que han sido registradas por un único informante puesto que no las considera representativas de la población total sino que las trata como errores de encuesta o usos lingüísticos particulares de informantes. Aparte del problema metodológico que conlleva la supresión de formas léxicas registradas (a pesar de que hayan sido testimoniadas por un único informante), en los datos de Varilex existen preguntas a las que han respondido únicamente 4 informantes, por lo que suprimir la respuesta de uno de ellos supone eliminar un 25 % de la muestra. Si contáramos con miles de registros en cada país, podríamos debatir acerca de la supresión de los hápax, pero dadas las dimensiones de las muestras recogidas, nos parece aventurado optar por esta premisa. En cambio, si realizamos una ponderación de los resultados, las respuestas únicas en un país tienen automáticamente una representatividad mínima, por lo que no tenemos que plantearnos esta disyuntiva. Esto no quiere decir

que puedan existir errores en los datos de Varilex, pero la supresión *a priori* de los registros únicos o hápax no está exenta de cuestionamiento.

Por todo lo expuesto consideramos que un cálculo demolingüístico ponderado de la representatividad es esencial para estudiar los datos de distribución con mayor rigor. Como veremos en el capítulo 4, gracias a los datos que ha estudiado Varilex y nuestro cálculo de representatividad, obtenemos las siguientes informaciones de las 182 nociones del *PCIC*:

- Concepto en inglés al que se refieren las nociones.
- Variantes léxicas de ese concepto.
- Países en los que se registran las variantes léxicas.
- Grado de Representatividad ponderado (GR) de cada una de las variantes.

Estos resultados permiten conocer la población que emplea activamente las variantes léxicas de cada concepto, lo que constituye una información fundamental para especialistas de diferentes campos. Además de permitir múltiples aplicaciones, nuestro cálculo demolingüístico ofrece un panorama global de la distribución de la variedad léxica que no sería posible sin los registros sistemáticos de Varilex, cuyo acceso libre permite consultar y analizar los conceptos, las variantes léxicas, las encuestas y las respuestas registradas por los informantes.

Para comparar y validar los datos, sería muy interesante realizar análisis similares con datos procedentes de otros estudios, pero Varilex es el único proyecto de investigación que recoge y publica encuestas, datos y resultados de análisis de forma periódica, que además cuentan con el prestigio y el reconocimiento de lingüistas y dialectólogos. En todo caso, tal y como veremos en el capítulo 5, en los estudios léxicos son muy habituales las discrepancias respecto a la delimitación geográfica exacta de las voces.

4.3. Materiales de estudio

En esta investigación trabajamos con tres materiales fundamentalmente: el *PCIC*, la base de datos de Varilex y el *Diccionario de americanismos*.

4.3.1. El Plan Curricular del Instituto Cervantes

El *PCIC* es la obra central de esta investigación. Remitimos al capítulo 1 de nuestra tesis, donde se presenta una descripción exhaustiva de este documento. El objetivo central de esta tesis es realizar una propuesta de revisión y ampliación de las nociones léxicas que presenta este material para integrar variedades dialectales representativas del español desde una mirada panhispánica y ecolingüística. En este sentido, nuestro trabajo demolingüístico se limita a las nociones que incorpora esta guía de enseñanza y aprendizaje, aunque podría aplicarse a cualquier otro repertorio léxico.

4.3.2. El proyecto Varilex

La base de datos de Varilex ofrece los datos de variación léxica en los distintos países hispanohablantes. Remitimos al capítulo 2 para encontrar una descripción detallada de los objetivos y características de este proyecto de investigación.

Los cálculos demolingüísticos que se presentan en esta investigación parten de los datos extraídos de las investigaciones de Varilex, sin los que nuestro estudio no sería posible. En este sentido, las nociones del *PCIC* que se han podido analizar se limitan a las nociones estudiadas por Varilex. Si contáramos con otros estudios sistemáticos de la variación podríamos calcular nuevos índices de representatividad en otras nociones léxicas pero, hasta el momento, este es el único material publicado con encuestas de uso aplicadas en un amplio periodo temporal, cuyos datos son abiertos y accesibles para los investigadores.

4.3.3. Diccionario de americanismos

El *Diccionario de americanismos* es el material de referencia que nos permite confirmar los datos recogidos por Varilex para delimitar los países donde se utilizan las voces. Este material, avalado por la Asociación de Academias de la Lengua Española, constituye un filtro fiable respecto a la distribución geográfica de las formas. Es preciso advertir que, tal y como se desprende del capítulo 5, es muy común que existan discrepancias respecto al uso de las voces en las diferentes zonas, puesto que el léxico es un inventario de la cultura humana que no puede tratarse como un todo absoluto (Ueda, 2004).

El *Diccionario de americanismos* (2010) es un diccionario *diferencial* del español que deja a un lado las palabras comunes a todos los hispanohablantes (que constituyen más del 80 %) para centrarse en describir, sin propósito normativo, un glosario de términos habituales en América.

Una de las características que engrandece este diccionario es la delimitación de los países en que se emplean las diferentes voces léxicas. Gracias a estas marcas diatópicas podremos contrastar los datos extraídos de Varilex con este documento avalado por las academias hispanoamericanas. Además, este diccionario nos permite establecer precisiones de significado ante los casos de duda que se derivan de las respuestas de algunos colaboradores, constituyendo el filtro con el que contrastar nuestros datos de estudio.

4.3.4. Otras obras consultadas

Además de los tres materiales fundamentales anteriores, para esta tesis hemos realizado consultas, fundamentalmente, en el Corpus del Español del Siglo XXI (CORPES). Este proyecto, desarrollado por la Real Academia Española, es de libre consulta en Internet e incluye más de 220 000 textos codificados, de España y América, que pretenden dar cuenta del español actual. Desde 2001, el corpus incluye una distribución geográfica especialmente interesante para nuestro estudio puesto que, a diferencia del CREA, este corpus se concibe como proyecto panhispánico con

un 70 % de textos producidos en América y un 30 % producidos en España, con formas anotadas morfosintácticamente y lematizaciones. Los textos se pueden agrupar en subconjuntos según su distribución geográfica, lo que resulta de sumo interés para nuestro estudio.

Las grandes áreas dialectales que presenta el CORPES son: zona andina, Antillas (caribeña), Caribe continental, chilena, Estados Unidos, México y Centroamérica y Río de la Plata, «mediante el cruce de criterios diferentes, entre los que figuran la población, el volumen de publicaciones, la cantidad de ediciones digitales de periódicos y revistas, etc.» (Real Academia Española, 2013). Como veremos en § 4.5, para nuestro estudio demolingüístico aplicado al léxico del *PCIC* seguiremos esta delimitación.

La posibilidad de consultar un corpus de estas características ha sido fundamental en los siguientes casos:

1. Cuando detectamos un probable americanismo que no se incluye en el *Diccionario de americanismos* o que aparece sin marcas en el *Diccionario de Lengua Española* (DLE).
2. Cuando tenemos dudas acerca de la extensión geográfica de determinadas voces porque los datos de Varilex contienen respuestas que contradicen a otras investigaciones dialectales o a nuestra intuición.
3. Cuando detectamos posibles españolismos que no se estudian en los datos de Varilex pero que sería importante marcar en el *PCIC*.
4. Cuando tenemos dudas sobre la definición o la utilización de una voz en los distintos subsistemas del español.

Además de las consultas del CORPES, para establecer diferencias de uso y matizaciones de significados, hemos consultado diferentes diccionarios, a pesar de su carácter «imperfecto por definición» (Pérez, 2005: 11). Con sus limitaciones, los siguientes diccionarios han sido fundamentales en nuestra investigación: el *Diccionario de Lengua Española* (2014), de la Real Academia Española; el *Diccionario del Español Actual* (1999), de Manuel Seco; el *Gran Diccionario de*

Uso del Español Actual (2006), basado en el corpus CUMBRE; el *Diccionario de uso del español* (2007), de María Moliner; y el *Gran Diccionario de la Lengua Española* (2008), de Francisco Rico.

4.4. Datos analizados

Para medir la representatividad de las voces léxicas del *PCIC* utilizamos como fuente de documentación la base de datos de Varilex. Mediante una serie de búsquedas en el programa MySQL, extraemos la información necesaria respecto a los conceptos que nos interesa estudiar, los datos relacionados con los informantes, las encuestas, los países en que se han realizado y los valores específicos de las variantes léxicas registradas.

4.4.1. Conceptos y variantes

En nuestra investigación hemos analizado 182 conceptos extraídos de las encuestas y resultados de Varilex. A cada uno de estos conceptos, le corresponden diversas variantes que superan un millar de voces diferentes. El Anexo II presenta todos los conceptos que hemos estudiado para nuestro análisis demolingüístico. En él incluimos las definiciones que se han ofrecido a los informantes en las encuestas de Varilex así como el concepto general en inglés al que se refieren, para evitar nombrar los conceptos en alguna de las variedades del español.

4.4.2. Países encuestados y población

A continuación presentamos los países que hemos estudiado junto con la población total que les corresponde, según los datos de las proyecciones de población de la ONU (Naciones Unidas, 2013) que publica el Instituto Nacional de Estadística (INE). Incorporamos también la abreviatura que les corresponde según el *Diccionario de americanismos* y su código ISO (*International Organization for Standardization*) que permite representar países, territorios y áreas especiales con interés geográfico. El sistema ISO 3166-1 alpha-2 es el que seguiremos para la

presentación cartográfica de nuestros resultados, como veremos en el capítulo 4. Aunque Varilex recoge también variaciones interurbanas, en este estudio nos centramos en las comparaciones internacionales².

Tabla 8

Países y ciudades encuestadas con su población y abreviatura correspondiente

País	Ciudades	Población	DAm.	ISO
Argentina	Salta, San Juan, Neuquen, Buenos Aires	42 154 900	AR	RA
Bolivia	La Paz	11 024 500	BO	BO
Chile	Arica, Santiago, Concepción, Temuco	17 924 100	CH	CL
Colombia	Medellín, Santafé de Bogotá	49 529 200	CO	CO
Costa Rica	Puerto Limón, San José	5 001 700	CR	CR
Cuba	La Habana, Santiago de Cuba	11 248 800	CU	CU
Ecuador	Quito	16 225 700	EC	EC
El Salvador	El Salvador	6 426 000	EL	SV
España	La Coruña, Santiago de Compostela, Oviedo, Santander, Salamanca, Zaragoza, Barcelona, Guadalajara, Madrid, Valencia, Sevilla, Granada, Málaga, Almería, Murcia, Santa Cruz de Tenerife, Las Palmas de Gran Canaria	47 199 100	ES	ES
Guatemala	Guatemala	16 255 100	GU	GT
Honduras	Tegucigalpa, Nacaome	8 423 900	HO	HN
México	Monterrey, Aguas Calientes, Ciudad de México, Mérida	125 235 600	MX	MX
Nicaragua	León, Managua	6 256 500	NI	NI
Panamá	Panamá	3 987 900	PN	PN
Paraguay	Asunción	7 032 900	PA	PY
Perú	Lima, Arequipa	31 161 200	PE	PE
Puerto Rico	San Juan, Dorado, Mayagüez	3 680 100	PR	PR
República Dominicana	Santiago, Santo Domingo	10 652 100	RD	DO
Uruguay	Montevideo	3 430 000	UR	UY
Venezuela	Caracas, Mérida, Valencia, Tachira	31 292 700	VE	VE

² En esta tesis, no tendremos en cuenta para los datos de representatividad las encuestas de Guinea Ecuatorial, Filipinas y Estados Unidos.

4.4.3. Informantes: edad y sexo

Las encuestas de Varilex han sido hechas a 1373 informantes, de los cuales, 690 son mujeres y 683 son hombres. En relación con la edad de los informantes, la mayoría de ellos tienen entre 20 y 60 años, aunque también hay registros de informantes mayores de 60 y de menores de 20.

El detalle de esta distribución se presenta en las figuras siguientes:

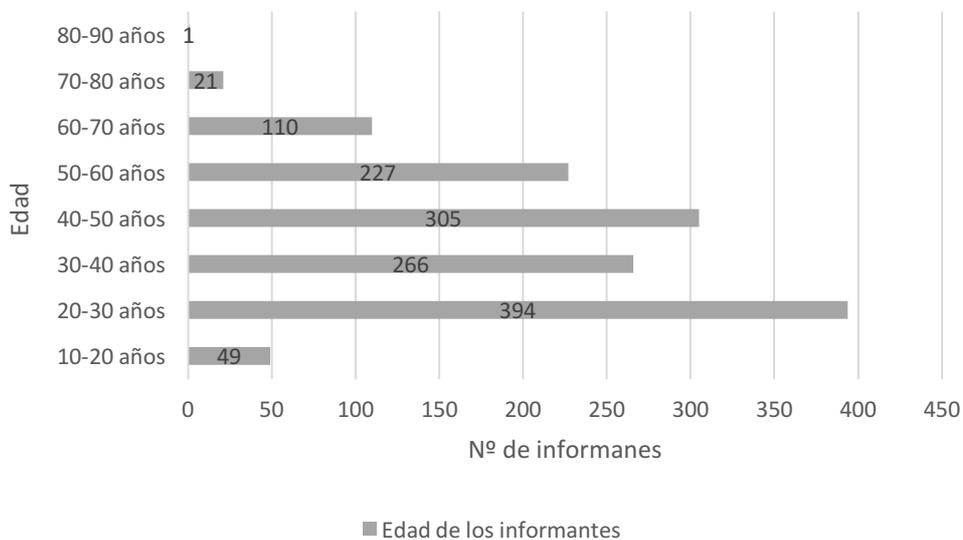


Figura 9: Distribución por edad de los informantes de Varilex

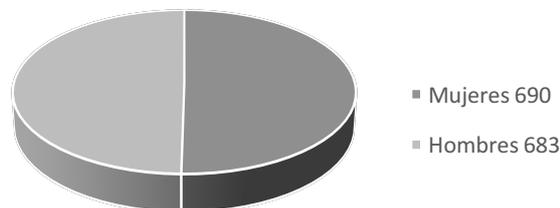


Figura 10: Distribución por sexo de los informantes de Varilex

4.4.4. Encuestas

Las datos de Varilex, recogidos en encuestas sobre variación léxica desde el año 1993, han sido fundamentales para elaborar nuestro análisis demolingüístico de los conceptos del *PCIC*. Para nuestra investigación hemos utilizado los datos de las encuestas A, B, C, D, E, F e I de Varilex, relacionadas con los siguientes campos semánticos: prendas de vestir, casa, comida, transporte, escritorio, ocio, deporte, transporte, ciudad, naturaleza y otros ámbitos del español general. En la imagen que presentamos a continuación se observa qué tipo de encuestas se han entregado a los informantes. Por lo general, se ofrece una definición y una imagen, junto al concepto en inglés.

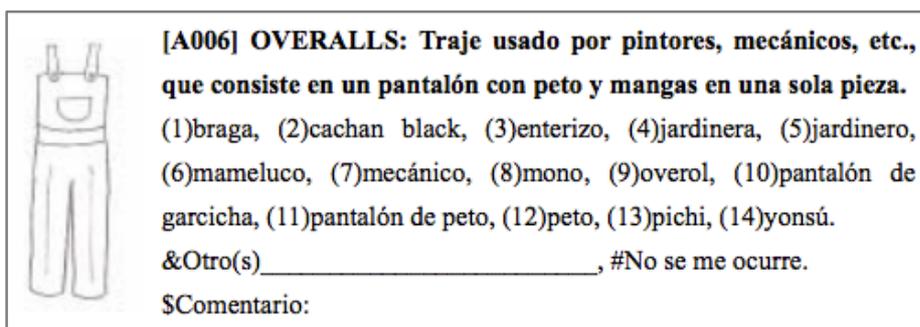


Figura 11: Ejemplo de pregunta con ilustración de la encuesta de Varilex

En los casos en los que no es posible aportar una ilustración se suele ofrecer, además, una oración que permita situar la forma en *contexto* y *cotexto*:

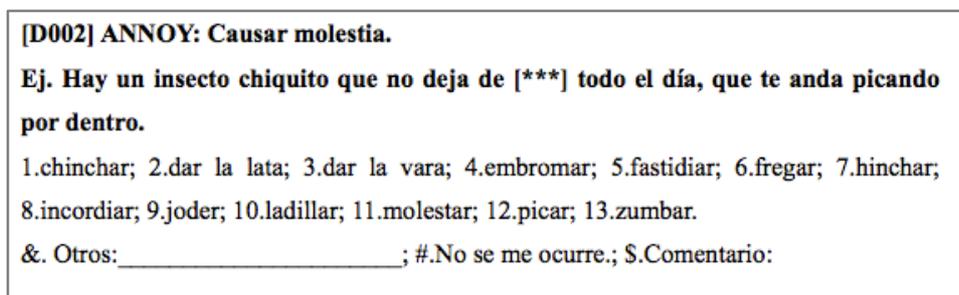


Figura 12: Ejemplo de pregunta apoyada en ejemplos de la encuesta de Varilex

Como podemos observar, Varilex ofrece una lista de opciones y permite que el hablante pueda incorporar cualquier otra voz que no se haya tenido en cuenta en las sucesivas versiones de las encuestas. Tanto el método de ejemplos como el de las ilustraciones permiten contextualizar rápidamente los conceptos, para que la respuesta de los informantes se asemeje a un contexto natural de uso lingüístico.

4.4.5. Identificación de variantes léxicas

Para seleccionar las variantes léxicas que son objeto de nuestro análisis seguimos los siguientes criterios:

1) Las distintas realizaciones fonéticas se consideran formas de una misma palabra: *chompa* y *chumpa*.

2) Las distintas marcas ortográficas se consideran formas de una misma palabra: *pulóver* y *pullover*.

3) Los cambios de género se agrupan como una sola forma: *la/el radio*.

4) Los cambios de sufijo (*computador/computadora*) y de abreviación no se agrupan como una sola forma sino como dos diferentes: *mesilla/mesita* *televisión/tele*.

Cuando las informaciones documentales, como las consultas de corpus y diccionarios, niegan la correspondencia semántica de determinadas voces, no tenemos en cuenta estas alternancias, puesto que nuestra intención es encontrar equivalencias semánticas con diferente distribución geográfica (geosinónimos).

Por otra parte, como ya hemos explicado, gracias al cálculo de representatividad ponderado, en este estudio no es necesario eliminar los hápax o registros únicos para los cálculos de representatividad. Los datos ponderados otorgan un peso muy bajo a los registros de un único informante.

4.5. Distribución en áreas dialectales

Para valorar la inclusión de variantes léxicas en el *PCIC* partimos del principio de que las nociones deben tener «un área o territorio de validez y vigencia lo suficientemente amplio como para justificar que forme[n] parte del *input* al que será sometido el aprendiz» (Andión, 2007: 27). Cuando el GR de las nociones estudiadas para la tesis no es muy elevado, necesitamos valorar la dispersión de las voces en diferentes zonas lingüísticas, con el objetivo de ofrecer una propuesta equilibrada sobre la distribución geográfica de las nociones léxicas del español.

Existen diferentes clasificaciones de las zonas geográficas del español de América. La primera clasificación del español americano la desarrolló Henríquez Ureña en 1921³. En los años sesenta, Rona se desmarca de la clasificación anterior y propone mapas lingüísticos basados en isoglosas, combinando el yeísmo, el zeísmo, el voseo y las formas verbales, estableciendo hasta 16 zonas diferenciadas. Desde entonces hasta nuestros días se han realizado múltiples clasificaciones que proponen la agrupación en otras áreas geográficas⁴.

Sabiendo que no existe una correspondencia biunívoca entre variedades lingüísticas y países, en esta investigación hemos seguido la clasificación del CORPES para delimitar las áreas geográficas del español, lo que facilita nuestras consultas al corpus cuando necesitamos comparar los registros de uso con los datos de Varilex o con los del *Diccionario de americanismos*, sobre todo en los casos que presentan discrepancias. En la Tabla 9 detallamos la distribución de las áreas con los países que les corresponden:

³ Henríquez Ureña distingue cinco áreas geográficas según criterios diversos, como la proximidad de las regiones, las relaciones políticas y culturales entre los países y el contacto de cada zona con una lengua indígena principal (Moreno de Alba, 1988: 155).

⁴ Destacan, entre otras, las clasificaciones de Resnick (1976), Cahuzac (1980) o Zamora y Guitart (1982). No obstante, el criterio de unidad y diversidad es ciertamente relativo y es necesario desarrollar investigaciones lingüísticas cohesionadas para definir la distribución de las variedades diatópicas del español en América; «hasta que llegue ese día, todos los intentos de clasificación no dejarán de ser propuestas efímeras sin valor definitivo» (De Sterck, 2000: 44).

Tabla 9
Áreas dialectales analizadas

Española	España
América Andina	Bolivia, Ecuador, Perú
Antillas (caribeña)	Cuba, Puerto Rico, República Dominicana
Caribe continental	Colombia, Venezuela
México y Centroamérica	Costa Rica, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá
Río de la Plata	Argentina, Paraguay, Uruguay
Chilena	Chile

5. Limitaciones de esta investigación

El propósito de esta investigación no es establecer inferencias causales o sociolingüísticas, ni desarrollar estudios históricos sobre la variación. Con este trabajo pretendemos mostrar la distribución geográfica de las voces léxicas de una forma exhaustiva. Para ello, es fundamental plantearse la representatividad de los datos estudiados, que proceden de las encuestas realizadas por el grupo de investigación de Varilex.

Una de las principales limitaciones de esta tesis es el número de informantes encuestados respecto a la población total de cada país, lo que exige tomar con cautela los resultados obtenidos. Necesitamos más informantes en cada país hispanohablante para considerar que nuestros datos son una muestra representativa de la población. Una muestra de 30 o 40 informantes en México orienta respecto a las preferencias léxicas pero no es verdaderamente representativa del habla de 140 millones de mexicanos. Sería de excepcional interés realizar más encuestas en los países estudiados, de manera que se amplíe la representatividad de los resultados.

Asimismo, Varilex estudia la variación léxica del español urbano en determinadas localidades del mundo hispánico. La elección de unas ciudades

frente a otras también condiciona los datos de estudio. Es importante afirmar que si aumentamos la red de encuesta, nuestros datos serán considerablemente más representativos e interesantes, sobre todo en países con zonas subdialectales internas muy marcadas.

Por último, nuestros datos están condicionados por las encuestas. En esta tesis nos planteamos inicialmente el trabajo con corpus pero esta metodología no nos permite acceder a algunas palabras que se usan exclusivamente en ciertas situaciones, sobre todo relacionadas con el habla espontánea, «when they know each other very well, in a family setting or with groups of friends» (Béjoint, 1979: 26). Además, la inexistencia de un corpus del español de España y América basado en la lengua oral no permite plantear un estudio de la variación léxica a gran escala –recordemos que estudiamos la variación léxica de más de 400 millones de hablantes. El hecho de necesitar confeccionar encuestas para mostrar la distribución geográfica del léxico exige que los cuestionarios deban ser preparados con una rigurosidad máxima⁵.

Teniendo en cuenta las limitaciones expuestas, y admitiendo que nuestros resultados no se pueden tratar de manera absoluta puesto que permanentemente se descubren nuevas voces y acepciones, consideramos que el análisis demolingüístico de esta tesis permite reflejar la presencia, ausencia y zonas de convivencia de las voces léxicas del español. Nuestra metodología cuantitativa aporta cifras de representatividad que, lejos de constituir una realidad absoluta, ofrece una imagen general de la distribución del léxico en cada país «que no contradice ni la intuición ni los conocimientos previos» (Ueda, 1995: 14). Así, los métodos estadísticos ofrecen informaciones relevantes sobre la distribución lingüística, sin que tengan que entenderse de manera literal. Por otra parte, el hecho de que los métodos cuantitativos en

⁵ Una de las grandes ventajas de las encuestas por Internet de Varilex es, precisamente, la posibilidad de modificar los errores de los cuestionarios e incluir mejoras en las sucesivas ediciones. Este sistema también permite actualizar los resultados de investigación sistemáticamente, a partir de los registros web, lo que reabre el debate sobre la ilusión metodológica del sincronismo puro en los estudios lingüísticos (Tinianov y Jakobson, 1970).

lingüística no puedan ser considerados absolutamente objetivos no significa que los datos tengan menor valor empírico, puesto que en estadística todo trabajo empírico no se hace sobre los datos brutos sino más bien sobre una imagen de la realidad (Goebel, 1981). Tal y como explicamos en el capítulo 2, hay que tener presente que el léxico no es un material cerrado sino un inventario permeable y dinámico del lenguaje, que podrá definirse y cuantificarse en mayor medida cuantos más datos se registren.

Por último, es importante señalar que el cálculo demolingüístico no se reduce a un proceso informático sino que en nuestro trabajo hemos estudiado minuciosamente los datos de análisis eliminando aquellos que presentan problemas en la distribución, en el recuento o en lo que respecta a su significado.

6. Conclusión

En este capítulo hemos presentado las características de nuestra investigación demolingüística sobre la variación léxica del español. En primer lugar, hemos explicado el origen de nuestra tesis y la justificación nuestro trabajo. La ausencia de un material de referencia sobre la distribución de la variación léxica en todo el ámbito hispanohablante implica que el tratamiento de este componente en las clases de ELE no pueda desarrollarse con la profundidad y el rigor que debería.

A continuación, hemos definido las hipótesis de investigación que orientan nuestro estudio así como los objetivos generales y específicos que pretendemos alcanzar. Tal y como hemos visto, el objetivo último de la tesis es plantear una propuesta de ampliación léxica del *PCIC*, para lo que necesitamos *operacionalizar* el concepto abstracto de representatividad. Mediante un estudio descriptivo, que parte de los datos recopilados por Varilex desde el año 1993 a partir de encuestas onomasiológicas, proponemos un análisis demolingüístico ponderado que refleje el panorama global sincrónico de la variación léxica del español urbano en el mundo.

Además, hemos definido las actividades necesarias para nuestra investigación así como los materiales esenciales con los que trabajamos. Así, el *PCIC*, la base de datos de *Varilex* y el *Diccionario de americanismos* vertebran nuestro estudio, junto con otros corpus y diccionarios que también resultan fundamentales. La descripción de los conceptos analizados incluye las características generales de distribución y procesamiento de los datos. También se han señalado las áreas dialectales que tenemos en cuenta para valorar la dispersión geográfica de las voces, junto a los datos cuantitativos de representatividad.

Por último, hemos planteado las limitaciones de esta investigación que, aunque no se puedan obviar de cara a la comprensión y extrapolación de nuestros resultados, no restan validez ni fiabilidad a la propuesta que presentaremos en los siguientes capítulos.

CAPÍTULO 5

RESULTADOS DEMOLINGÜÍSTICOS

1. Introducción

Para alcanzar el objetivo principal de nuestra investigación y ampliar la representatividad de las «Nociones específicas» del *PCIC*, vamos a calcular el GR de 182 nociones de este inventario para las que el equipo de Varilex ha identificado una serie de voces equivalentes. Se trata, por tanto, de delimitar un conjunto de geosinónimos en el territorio hispánico.

El GR de las voces léxicas se calcula mediante la fórmula de representatividad que explicamos en el capítulo 3. Se trata de un cálculo ponderado con el que medimos la extensión y el uso de cada variante, en función de las respuestas aportadas por los informantes y el peso de la población de cada país encuestado.

Los resultados que se presentan a continuación indican el GR específico de los 182 conceptos del *PCIC* que hemos analizado, así como de sus correspondientes variantes. Aplicando cálculos demolingüísticos, conseguimos operacionalizar la variable latente de la representatividad, lo que permite ordenar las voces léxicas sobre la base de su extensión en el mundo hispánico.

Para un análisis demolingüístico como el nuestro debemos tener en cuenta que las equivalencias léxicas que muestran los datos de Varilex ofrecen un panorama global de la distribución del léxico del español, que requiere, en muchos casos, una investigación adicional específica que permita interpretar los resultados adecuadamente. En todo caso, un análisis minucioso de todas las alternancias léxicas extraídas de Varilex excede los límites de la presente investigación.

Para calcular la representatividad de 182 conceptos del *PCIC*, hemos diseñado una aplicación virtual dinámica que permite conocer, a través de mapas, las tres variantes léxicas más representativas de cada concepto así como una tabla de distribución con el GR de todas las formas. Incluimos, además, una recopilación de mapas estáticos relativos a cada concepto estudiado. En este capítulo, por tanto, nos ocupamos de presentar y explicar nuestros resultados. La valoración y discusión de estos se reserva para el capítulo 6 donde propondremos, mediante un estudio minucioso de equivalencias léxicas, las posibilidades de ampliación del *PCIC*.

2. Presentación de resultados

En el ámbito de la dialectología, la variación lingüística en el espacio se ha reflejado tradicionalmente mediante mapas. Si nuestro objetivo es reflejar la distribución espacial de la variedad léxica, utilizar tablas con cientos de palabras acompañadas de un número (GR) y de enumeraciones de países no parece muy útil. Un estudio como el nuestro, que se enmarca en la geografía lingüística y se sirve de técnicas informáticas para el acceso y tratamiento de los datos, debe utilizar un sistema de representación de mapas lingüísticos que permita mostrar una imagen de la extensión del léxico en la geografía hispánica.

La importancia de los atlas lingüísticos es incuestionable en la actualidad. A pesar de que muchos de estos materiales hayan quedado relegados en las investigaciones dialectales, es importante destacar que los datos que exponen los atlas lingüísticos del siglo XX han constituido el origen de múltiples

investigaciones relacionadas con la semántica, la historia de la lengua y otros estudios aplicados de la lingüística.

En un principio, el objetivo de la dialectología era marcar en los mapas lingüísticos las isoglosas que delimitaran las áreas dialectales. Sin embargo, el concepto de isoglosa fue puesto en evidencia por la geografía lingüística, de manera que la alusión a los límites dialectales se convirtió en una cuestión virtual y genérica (Gimeno Menéndez, 1990: 22).

En realidad, la utilización de mapas en los estudios dialectológicos pretende constatar la distribución geográfica de ciertas variantes en un momento concreto, mostrando una *imagen* específica de lengua (Coseriu, 1981). Así, un estudio como el nuestro, que muestra la distribución del léxico del español según su índice de representatividad, es idóneo para cartografiarse, siempre que los resultados se hayan basado en una red de localidades fijada de antemano, con un mismo cuestionario, en una situación convenida y en un mismo nivel de uso (García Mouton, 1990).

Como vimos con más detalle en el capítulo 2, los atlas lingüísticos presentan enormes colecciones cartográficas de material lingüístico que muestran, mediante mapas, la distribución de las variantes léxicas estudiadas. De la misma manera, en esta tesis vamos a presentar cartográficamente nuestro análisis demolingüístico exponiendo con claridad la distribución léxica del español urbano, en función del GR de los países analizados.

Tras valorar diferentes alternativas, hemos diseñado un sistema de cartografiado lingüístico homogéneo en su presentación y configurado a partir de los planteamientos onomasiológicos de las encuestas de Varilex. Así, mediante mapas, mostramos la distribución espacial de las variantes léxicas más representativas de cada concepto analizado. Estos mapas permiten desarrollar hipótesis e investigaciones sobre las inferencias causales, el desarrollo histórico de los fenómenos de variación u otras cuestiones relacionadas con la distribución léxica. En todo caso, recordemos que nuestro objetivo se centra en valorar la extensión y el uso

de las formas del *PCIC* y de las correspondientes voces equivalentes más representativas del español.

Así, hemos diseñado un sistema de representación cartográfica que muestra una imagen visual de nuestros cálculos demolingüísticos, representando la distribución de las dos o tres voces más representativas de cada concepto estudiado.

3. Aplicación virtual para la representación cartográfica de los datos

Como ya hemos apuntado, en esta tesis partimos de la base de que la representación de la variación léxica mediante mapas permite comparar la representatividad y localización de las diferentes voces, así como deducir o descartar hipótesis sobre la distribución geolectal.

Una investigación léxica como la nuestra, que se sirve de cálculos demolingüísticos y de herramientas informáticas que permiten aplicar fórmulas matemáticas a gran escala, también debe hacer uso de las tecnologías de la información y de la comunicación para diseñar un sistema de cartografiado que supere las limitaciones de los atlas lingüísticos del siglo pasado. Frente a una representación geolectal que se circunscribe a un momento temporal concreto, con limitaciones relacionadas con el manejo del material, la actualización de los datos y las posibilidades de acceso, hoy en día existen herramientas tecnológicas que presentan ventajas incuestionables.

Uno de los objetivos específicos de este trabajo consiste en diseñar un sistema de cartografiado que permita visualizar con facilidad la distribución espacial de los datos. Para ello, hemos diseñado un sistema de representación que admite la intervención del usuario en un entorno digital, promoviendo consultas interactivas.

La cartografía virtual ofrece una visualización de mapas lingüísticos confiables a través de una experiencia intuitiva e interactiva que garantiza una práctica muy diferente a la que admitía la cartografía histórica. Esto permite que nuestros resultados cuenten con mayor visibilidad y alcance, lo que multiplica las posibilidades de aplicación a diferentes ámbitos del saber. De este modo, nuestro

cartografiado dinámico de los datos posibilita una lectura e interpretación de los resultados lingüísticos que es clara y accesible desde cualquier ordenador.

Tras valorar los diferentes sistemas de representación cartográfica de datos espacio-temporales, hemos optado por las oportunidades que brinda el entorno R, que se utiliza para análisis estadísticos en la investigación. En concreto, se ha utilizado el paquete Shiny, desarrollado por RStudio y que permite crear aplicaciones web interactivas.

Como ya hemos apuntado, este tipo de representación cartográfica dinámica permite que el usuario acceda con facilidad a los datos que le interesen según sean sus propósitos de investigación, pudiendo acercarse o alejar los mapas así como editar la información de búsqueda en función de sus intereses.

Por otra parte, es importante destacar que el cartografiado virtual permite realizar modificaciones sobre los mapas, incluyendo nuevos datos o rectificando los posibles errores. Además, aunque nuestra aplicación indica el GR, por país y global, de los 182 conceptos del *PCIC*, en nuestro programa podríamos añadir otras informaciones diversas, como particularidades semánticas, apoyos visuales de cada noción, especificaciones de tipo histórico, geográfico, de distribución, de uso o de zonas dialectales, entre otros.

Para acceder a nuestra aplicación hay utilizar en el siguiente enlace: <https://carmenchacon.shinyapps.io/PCIC/>. En la parte superior izquierda, podemos elegir el término en inglés que nos interesa estudiar. A continuación, aparece automáticamente un mapa de las variantes léxicas correspondientes con colores diferentes según las voces y círculos de mayor o menor tamaño, en función de su representatividad. Si pinchamos en los círculos del mapa obtenemos el GR de una variante (representada con un color determinado) en cada país.

En la tabla de resultados que aparece junto a cada mapa podemos consultar el GR total de cada voz, que es el índice que valoramos en nuestra propuesta de ampliación del *PCIC*, desarrollada en el capítulo 6. Incluimos también una lista con los países en los que se ha registrado, que complementa la información cartográfica.

En la lista de países solo aparecen aquellos en los que un mínimo de dos informantes han registrado la voz. Aunque para el cálculo del GR y su representación cartográfica hemos tenido en cuenta también estos hápax o respuestas aisladas —debido a que la técnica de la ponderación equilibra matemáticamente el peso de estos datos—, en las tablas de representación que acompañan a los mapas no se incluyen aquellos países en los que solo un informante ha registrado la voz, dado que su objetivo es mostrar una síntesis general de la distribución de la variación léxica y podrían perturbar la interpretación de los resultados.

Presentamos a continuación algunas imágenes de la interfaz gráfica que muestra, de manera interactiva, los resultados de representatividad léxica. Vamos a señalar las partes fundamentales de nuestra aplicación y las posibilidades que permite a los usuarios:

(1) Seleccionar el concepto que nos interesa estudiar: En la *Ilustración 1* mostramos cómo el usuario puede seleccionar los conceptos que quiere estudiar mediante el menú desplegable de la parte superior izquierda. Allí presentamos alfabéticamente todos los conceptos del *PCIC* que han sido estudiados en esta tesis. En este ejemplo, seleccionamos el concepto [APRON] y automáticamente se genera un mapa con la distribución geográfica del GR.

Grado de Representatividad léxica del español

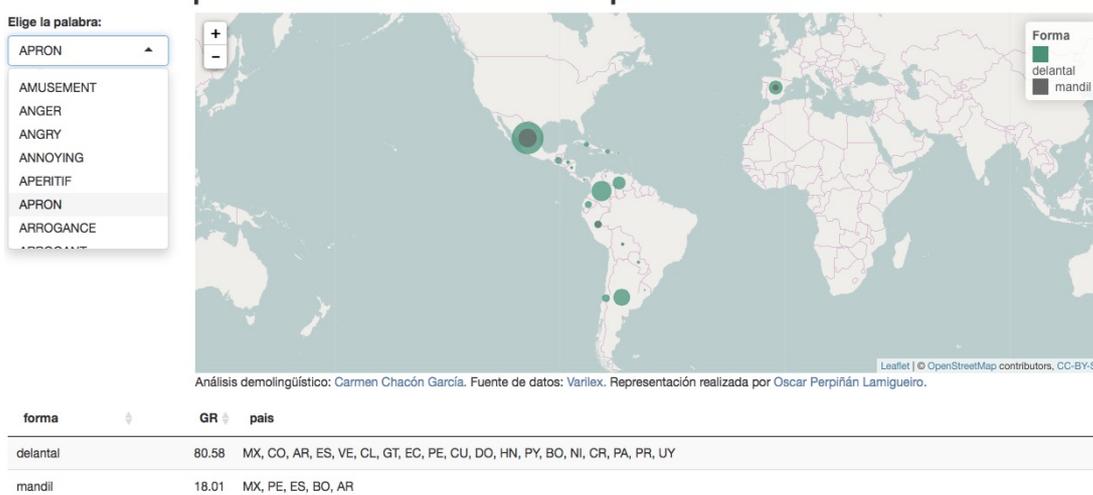


Figura 13: Cómo seleccionar un concepto de estudio en la aplicación virtual

(2) Representar en un mapa el GR de las tres formas más representativas del concepto estudiado. Cada una de las variantes tiene un color y se distribuye cartográficamente en función de su distribución geográfica. Además, cada círculo tiene un radio con un tamaño determinado en función del GR de cada forma en cada país:



Figura 14: Representación de las variantes más representativas del concepto

(3) Consultar la contribución por país respecto al GR global. Para ello, es necesario situar el cursor sobre los círculos representados:

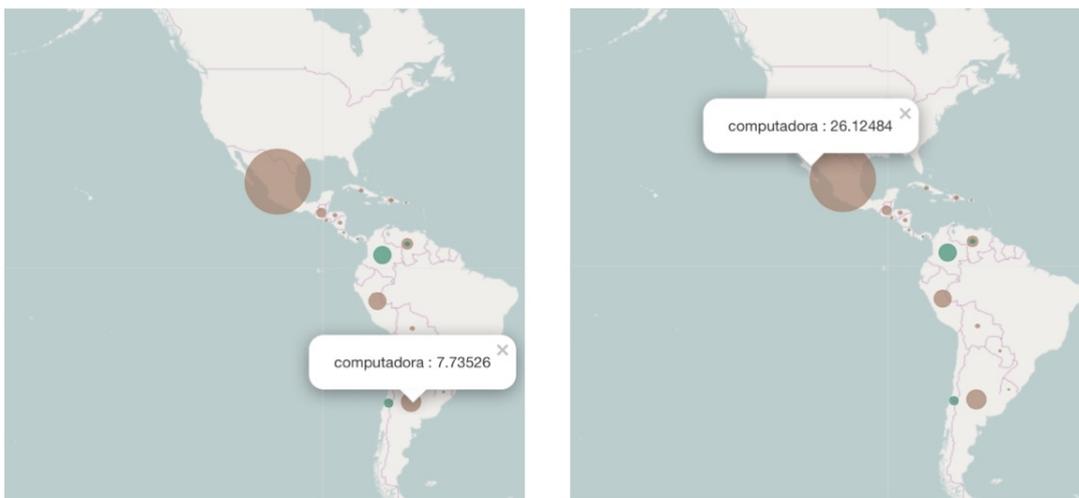


Figura 15: Resultados del GR de una variante léxica en distintos países

(4) Acceder a una tabla de resultados que incluye la distribución de las variantes léxicas de un concepto, con su respectivo GR total (que es la suma de todos los GR por país) y la lista de países en los que se ha registrado la voz.

(5) Ordenar las tablas de resultados en función de la información que busque al usuario, según las variantes estudiadas o en función de los cálculos del GR total de cada voz.

Grado de Representatividad léxica del español

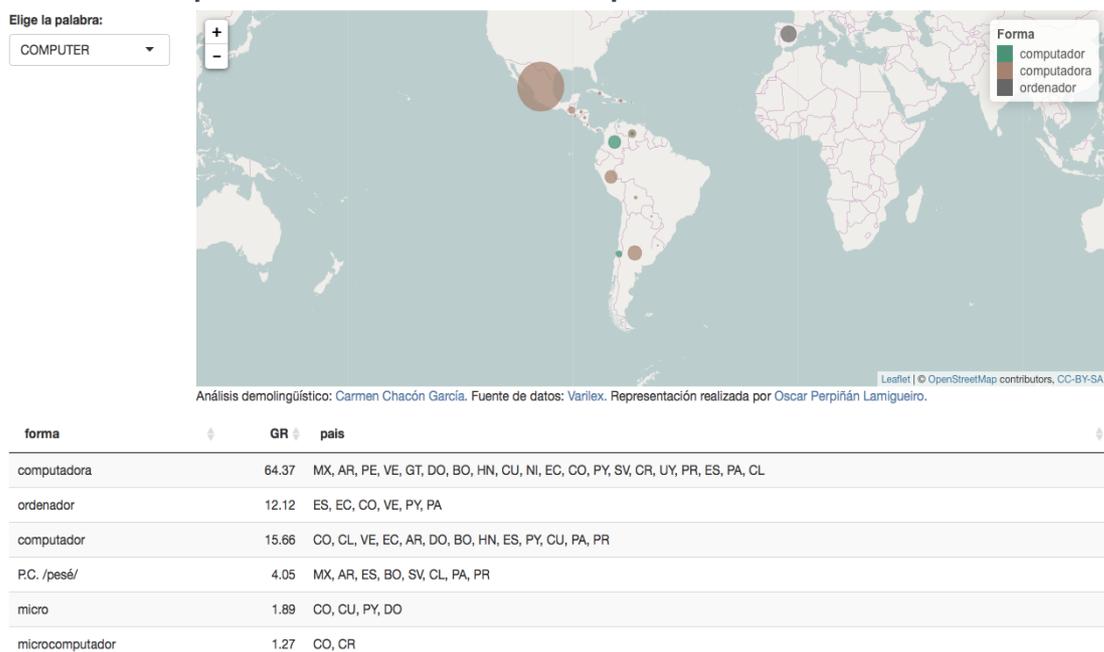


Figura 16: Consulta del mapa del concepto [COMPUTER] con la Tabla de Resultados

Somos conscientes de que nuestros datos se podrían representar a través de mapas diferentes. La representación espacial de datos permite utilizar mapas de isolíneas, de coropletas, de diagramas, de flujos, de relieve estadístico, anamórficos, mapas ráster, mapas vectoriales y mapas de figuras o cartogramas, entre otros.

En esta tesis hemos seleccionado el sistema de mapas de símbolos proporcionales puesto que, además de garantizar una correcta legibilidad por parte de cualquier usuario, resulta especialmente apropiado para exponer los resultados de nuestro estudio. Este tipo de representación se basa en seleccionar un símbolo y

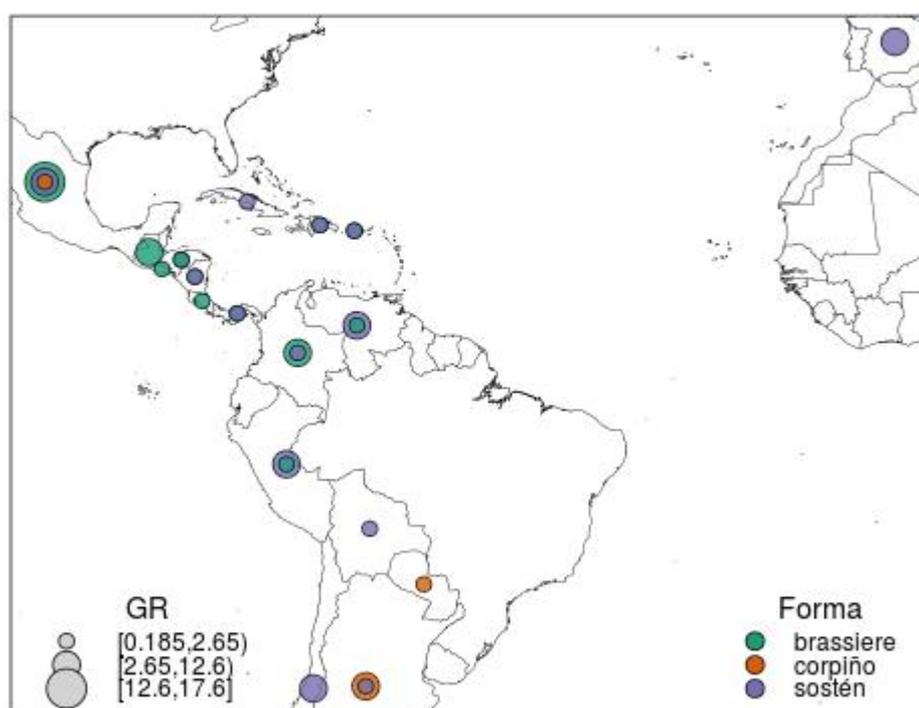
variar su tamaño en proporción a la cantidad de la variable que deseamos representar (en este caso, la representatividad, medida con el GR). Con este sistema representamos el GR de cada país de manera proporcional respecto a la representatividad global de cada voz léxica. Además, podemos señalar la utilización de diferentes variantes en un mismo país mediante círculos concéntricos de diferentes colores con sus respectivos grados representatividad. Respecto a la simbología, hemos elegido círculos frente a otras figuras geométricas (como cuadrados, esferas o cubos) o pictográficas (como diagramas de barras o dibujos) puesto que, según Slocum *et al.* (2010) y Dent (1999), el círculo es el icono más estable visualmente.

4. Mapas estáticos de la representatividad léxica

Además del diseño de este *software* específico, hemos diseñado un sistema de representación a través de mapas estáticos que incluimos en el Anexo III y que nos permite presentar nuestros resultados demolingüísticos en formato papel. Los mapas de resultados son similares a los mapas generados por la aplicación virtual y presentan una representación cartográfica de la distribución de las tres variantes con más GR de cada concepto, que son las que se tendrán en cuenta en la ampliación de las nociones del *PCIC* que presentaremos en el capítulo 6.

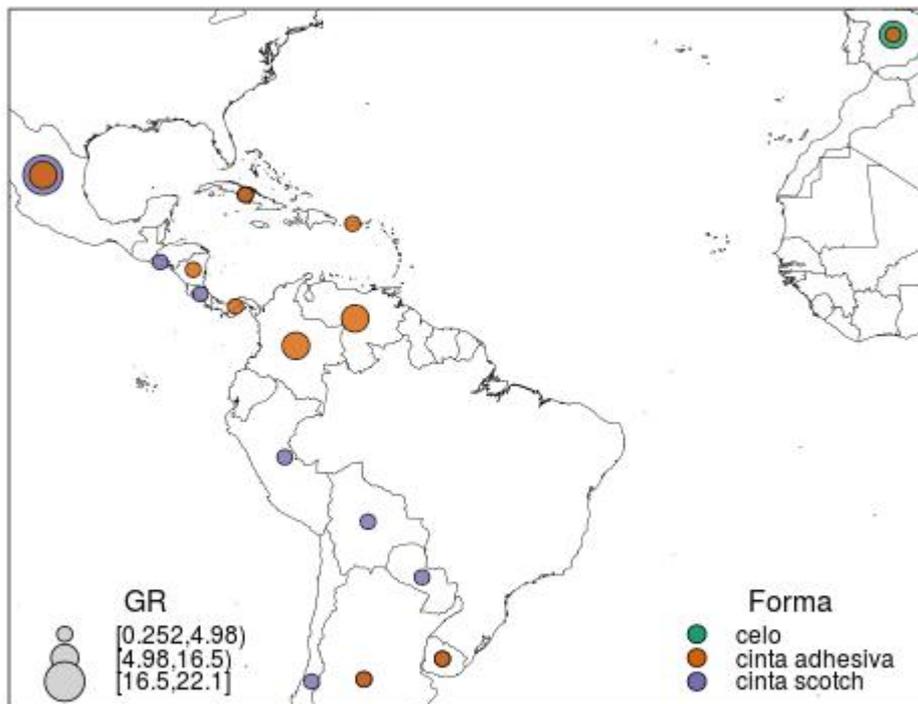
Mostramos, a continuación, dos ejemplos de nuestros mapas que corresponden al concepto [BRASSIERE] y al concepto [CELLOPHANE TAPE].

Mapa 1 *Concepto [BRASSIERE]*



forma	GR	pais
brassiere	36.23	MX, CO, GT, PE, SV, VE, NI, DO, CR, PA, HN, PR, BO, AR, CU
sostén	36.37	MX, VE, PE, CL, EC, ES, BO, CO, AR, DO, HN, NI, PA, CU, PR, SV, PY, CR
sujetador	7.04	ES, PE, CU, NI
corpiño	8.25	AR, MX, PY, BO, CL, UY, NI, ES
ajustador	2.03	CU, ES, PR
brasiel	5.32	CO, MX, HN, DO, PR, PA, NI
soutien /sutién	1.4	AR, UY
portaseno	1.37	AR, PY
soutien [sutién]	1.15	AR, UY, PY
tallador	0.49	CR

Mapa 2 Concepto [CELLOPHANE TAPE]



forma	GR	pais
cinta scotch	41.65	MX, AR, CL, PE, BO, SV, CU, PY, CR, UY, NI, PA, PR
cinta adhesiva	36.47	CO, VE, MX, ES, DO, AR, PE, NI, CU, PR, PA, BO, PY, CR, UY
celo	9.09	ES, PE
cintex	2.83	AR
celofán	0.95	ES, CU

Como podemos observar, en los mapas estáticos también se visualizan las tres voces más representativas de cada concepto mediante una representación de círculos proporcionales. Los mapas aportan información de la distribución de la variación

léxica de las voces más representativas utilizando distintos colores. También se observa el mayor o menor GR de las voces del país en función del tamaño de los círculos.

Es importante tener en cuenta que los resultados mostrados a través de mapas (tanto estáticos como dinámicos) no ofrecen explicaciones teóricas, por lo que una vez cartografiados exigen la realización de un estudio exhaustivo que permita interpretar la realidad representada. En nuestro caso, para aplicar los resultados demolingüísticos a un estudio sobre representatividad de las nociones del *PCIC*, que es el objetivo último de nuestra tesis, presentamos en el capítulo 6 un análisis detallado de la representatividad de los conceptos del *PCIC* así como de las dos o tres alternativas léxicas más representativas de cada uno.

5. Análisis de resultados

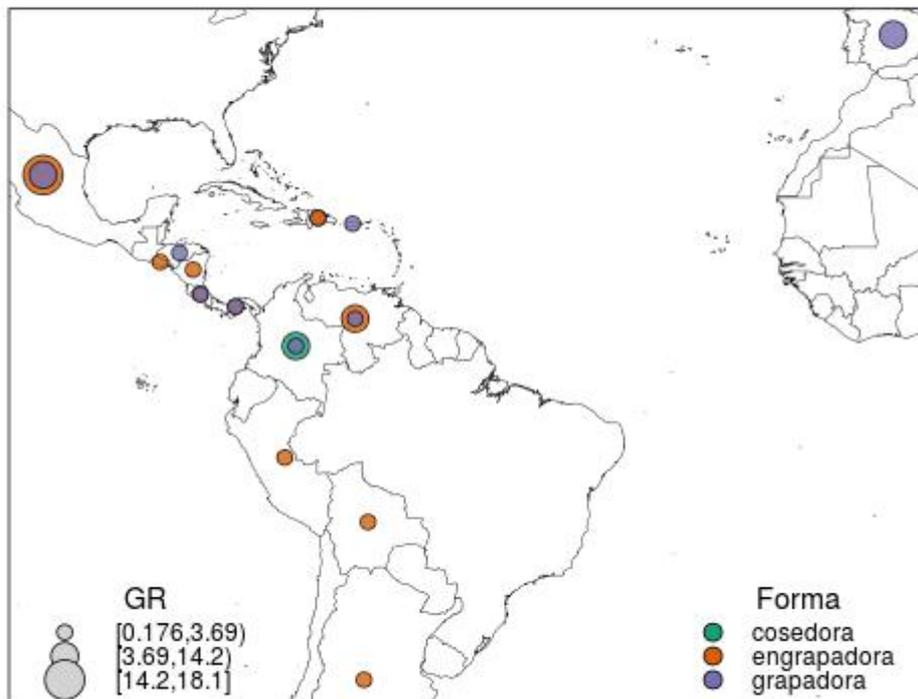
El diseño de mapas lingüísticos presenta una distribución geográfica que intuimos de manera aproximada. En ocasiones, esta distribución geográfica nos da pautas para aventurar hipótesis lingüísticas o descartarlas (García Mouton, 1990). En esta investigación corroboramos que la visualización de nuestros mapas es útil para obtener informaciones fundamentales sobre la variación léxica, como la sustitución de las voces, su existencia o inexistencia en algún ámbito geográfico, la vitalidad, la dispersión, el avance de unas formas frente a otras, la existencia de americanismos, panhispanismos, españolismos, etc.

A continuación, incluimos una síntesis de las aportaciones más significativas que ofrece la lectura de los mapas diseñados para esta tesis. Para facilitar nuestra exposición y la visualización de los resultados utilizaremos los mapas estáticos, aunque toda esta información se puede consultar con más detalle en nuestra aplicación virtual.

5.1. Sustitución de formas léxicas

La geografía lingüística nos permite estudiar la evolución de las formas léxicas a través del espacio, entendiendo la lengua como un organismo vivo que varía en función de distintos factores. En nuestros mapas de representatividad observamos, en primer lugar, que las palabras cambian en cuanto a su distribución, llegando incluso a ser reemplazadas por otras voces que, inicialmente, podían ser menos representativas, como ocurre, por ejemplo, con el concepto [STAPLER].

Mapa 3 Concepto [STAPLER]. Ejemplo de sustitución de formas léxicas



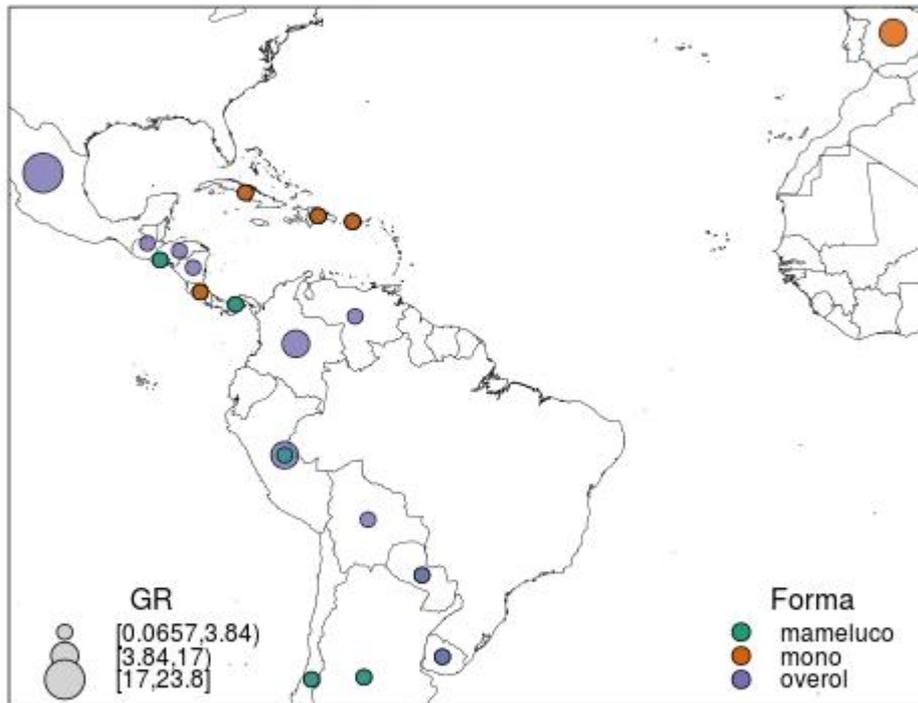
forma	GR	pais
engrapadora	33.16	MX, VE, PE, EC, SV, NI, BO, GT, CR, AR, DO, PA, HN, ES, PR
grapadora	27.77	ES, MX, CO, EC, DO, VE, HN, PR, PE, CR, PA
cosedora	8.43	CO, PA
abrochadora	6.46	AR, UY
corchetera	3.95	CL
engrampador	3.71	PE, BO, UY, PA
engrampadora	9.04	GT, MX, PE, BO, UY, VE, CR, HN, NI, SV, DO, PA
presilladora	4.03	CU, PY
abrochador	2.39	AR
engrapador	0.64	PE, SV, NI, PA

En la tabla de distribución que acompaña al mapa de [STAPLER] observamos que en España se han registrado dos variantes léxicas: *grapadora* y *engrapadora*. Según el mapa, sin embargo, en España hay solo una voz altamente representativa: *grapadora*. De hecho, si consultamos el CORPES no aparece ningún registro de *engrapadora*. Así, podemos afirmar que, tal y como se desprende de nuestros mapas, *grapadora* ha sustituido finalmente a la voz *engrapadora*, que todavía se incluye en el *Diccionario de Lengua Española* (2014).

5.2. Americanismos

La visualización cartográfica de la distribución de las formas léxicas también permite dar cuenta de la existencia de americanismos. Como sabemos, se trata de voces comunes a la mayoría de países americanos que no se han registrado en España. En nuestros mapas se puede observar su extensión geográfica, la coexistencia con otras variantes del español en regiones de Hispanoamérica, así como la ausencia de estas voces en territorio peninsular. Veamos, por ejemplo, el concepto [OVERALLS]. Tanto en el mapa demolingüístico como en la tabla de distribución, observamos que la voz *overol* es compartida por 19 de los 20 países estudiados. Se trata de un caso muy claro de americanismo que pone de manifiesto la existencia de voces que responden a la oposición clásica España / América.

Mapa 4 Concepto [OVERALLS]. Ejemplo de americanismo



forma	GR	pais
overol	60.01	MX, CO, PE, EC, GT, AR, BO, HN, CU, CL, NI, SV, VE, DO, CR, PA, UY, PR, PY
mono	7.74	ES, CU, DO, CR, AR, CL, PR
brega	5.97	VE
jardinero	4	AR, BO, UY
peto	3.92	ES, MX
mameluco	7.45	AR, CL, PY, PE, DO, PR, PA, UY, SV, HN, CR, NI, ES
pantalón de peto	2.83	MX, ES, DO
jardinera	3.24	CL, AR, PY, PE, BO, UY
pantalón de garcicha	0.89	GT
enterizo	2.29	CO, PE, BO, AR, CU, DO

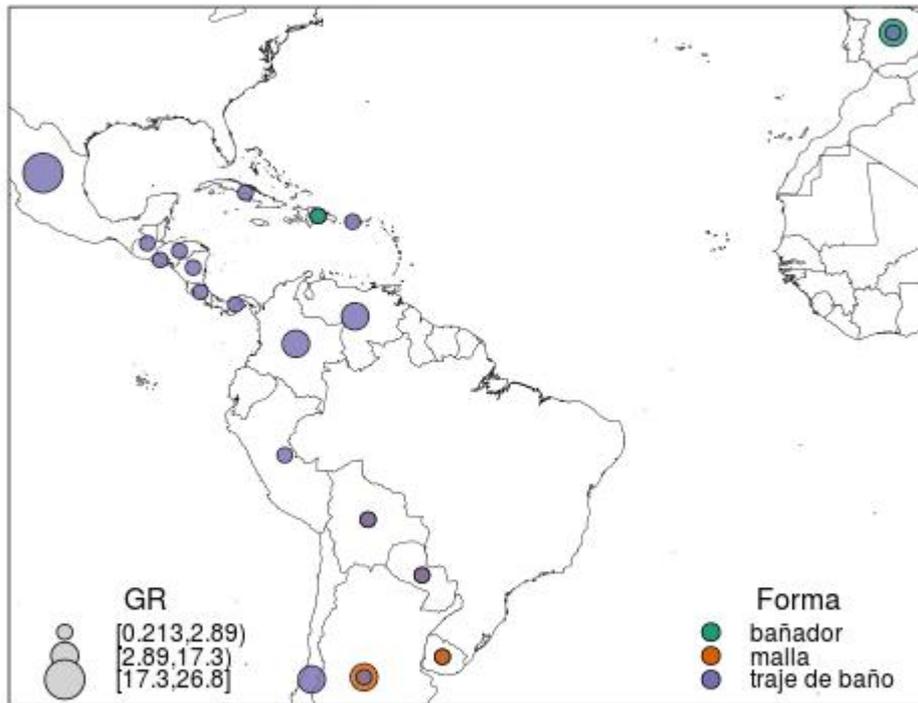
5.3. Españolismos

El análisis demolingüístico de nuestra tesis constata la existencia de españolismos o nociones exclusivas del español de España. Se trata de voces peninsulares que no se han registrado en países americanos o que se han registrado de manera muy puntual, con pocos informantes y en escasas zonas de la geografía hispánica.

A continuación presentamos dos mapas que muestran la distribución de las voces *bañador* y *frigorífico*. Podemos observar que *bañador* es una voz altamente representativa del español de España pero que no se registra en el resto de la geografía hispánica.

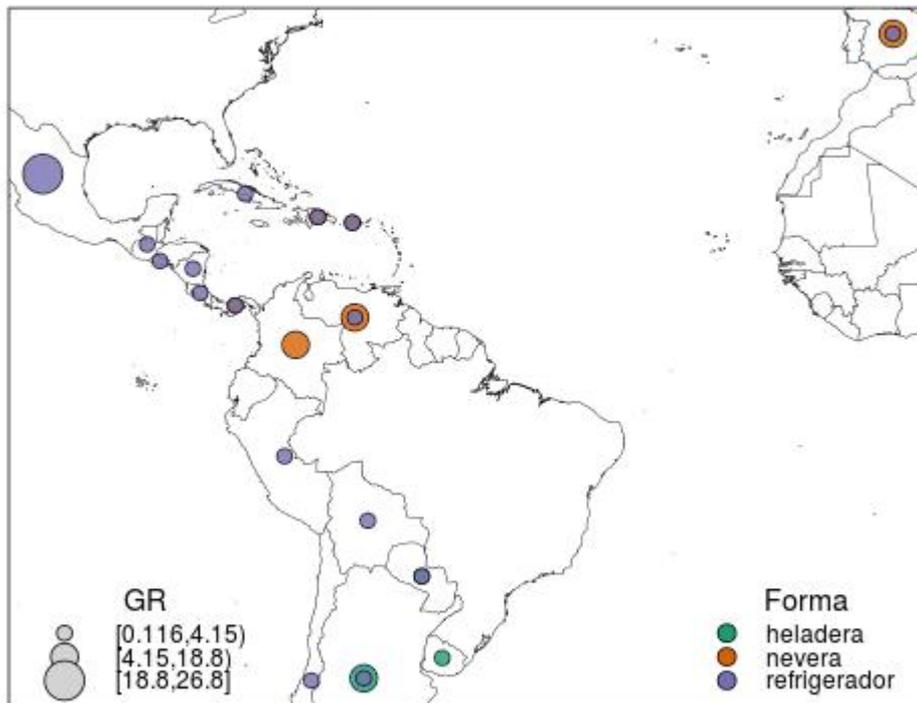
Lo mismo sucede con *frigorífico*, que es una voz fundamentalmente peninsular que no se utiliza en las regiones americanas. En cambio, *refrigerador* es la voz más representativa del español y *nevera* también aparece en gran parte del mundo hispanohablante. Esta distribución tiene consecuencias directas para la enseñanza de ELE puesto que la inclusión de *frigorífico* en un manual de enseñanza no respetaría los criterios de rentabilidad y extensión en la selección del vocabulario, puesto que es una voz exclusivamente peninsular que además convive con los dos panhispanismos mencionados.

Mapa 5 Concepto [SWIMSUIT]. Ejemplo de españolismo



forma	GR	pais
traje de baño	60.59	MX, VE, CO, CL, EC, AR, ES, DO, HN, GT, NI, PR, BO, PE, PY, UY, PA, CR, SV, CU
bañador	8.36	ES, PE, DO, NI, PR
vestido de baño	7.03	CO, CR, PA, NI
ropa de baño	5.43	PE, ES
malla	7.68	AR, BO, PY, PE, UY, DO
trusa	2.28	CU, PR
calzoneta	3.01	GT, SV, NI
malla de baño	2.41	AR, BO, PE, PY
bikini	3.2	MX, CO, ES, AR, CL, BO, DO

Mapa 6 Concepto [REFRIGERATOR]. Ejemplo de españolismo



forma	GR	pais
refrigerador	42.76	MX, CL, GT, PE, BO, CU, VE, AR, HN, CR, NI, SV, ES, DO, PY, PA, PR
nevera	25.76	CO, VE, ES, DO, MX, PR, PA, CR, CU
heladera	10.33	AR, PY, UY, BO
frigorfico	4.57	ES, PR
refrigeradora	10.6	EC, PE, HN, GT, NI, SV, CR, PA, AR, PY, DO, PR
frigidier /friydér/	2.54	PE, BO, CL
frío	0.64	CU, ES
frigidaire [frihidáire]	1.06	PE, BO, CU, CL
frigidier [friydér]	0.68	PE, CL, ES
frigo	0.54	ES, PE

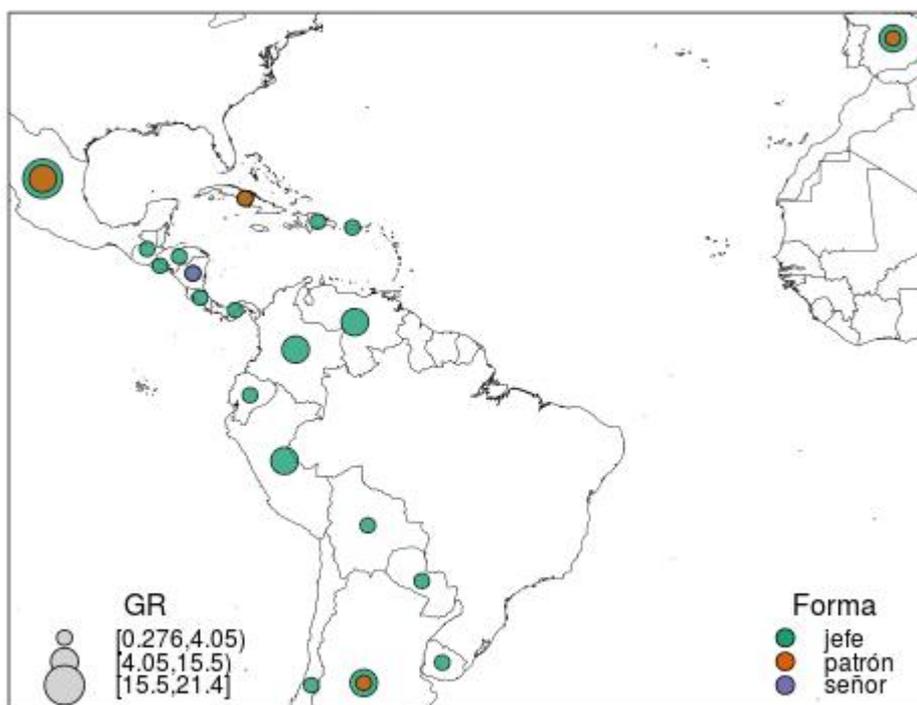
5.4. Panhispanismos

Nuestros mapas de resultados muestran aquellas palabras que se utilizan a lo largo de toda la geografía hispanohablante, tanto en España como en América, y que denominamos panhispanismos o nociones panhispánicas.

Estas voces son fundamentales para la enseñanza de español puesto que garantizan la comunicación con el mayor número de hablantes y cumplen con los criterios de extensión, uso y rentabilidad en el aprendizaje. Estas palabras son utilizadas o conocidas por cualquier hispanohablante sea cual fuere su procedencia y, por este motivo, no deberían faltar en el currículo de un curso de español.

Incluimos, a continuación, el mapa relativo al concepto [BOSS]. Podemos observar que la variante *jefe* está plenamente extendida en todo el ámbito del español, aunque convive con otras voces representativas como *patrón* o *señor*. No obstante, *jefe* es la voz que más se utiliza y conoce en todo el territorio hispanohablante. Podemos afirmar, por tanto, que *jefe* es una noción panhispánica.

Mapa 7 Concepto [BOSS]. Ejemplo de panhispanismo



forma	GR	pais
jefe	79.57	MX, ES, CO, PE, AR, VE, CL, GT, EC, BO, CU, DO, SV, PY, HN, NI, CR, PR, PA, UY
patrón	13.75	MX, AR, CO, VE, ES, EC, DO, BO, CU, PY, HN, CR, PA, UY, NI, PR
don	2.41	CO, CL, ES
señor	4.27	MX, GT, EC, AR, HN, NI, CU, ES

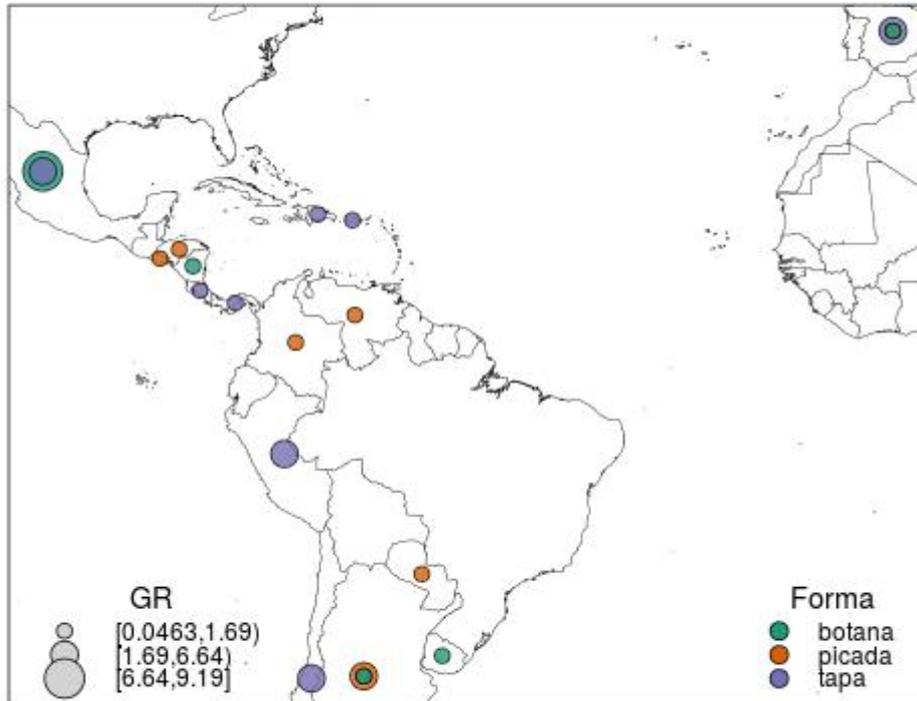
5.5. La vitalidad de las voces

Nuestros resultados también permiten observar la vitalidad de las voces, valorando su extensión y representatividad en el conjunto de la geografía hispanohablante. Por lo general, las voces más extendidas aparecerán en el mapa de nuestra aplicación, puesto que estas formas suelen tener mayor Grado de Representatividad.

La tabla de distribución respectiva ofrece información sobre la representatividad del resto de voces, señalando los países en los que se registra cada noción. El estudio del conjunto de nuestros datos puede hablarnos sobre la vitalidad de las formas léxicas. Así, un alto GR también puede venir asociado a que una determinada voz es la preferida de una región, sin que coexista con otras variantes equivalentes.

Esto sucede, por ejemplo, con la voz colombiana *pasante*. Así, para el concepto [TIDBITS], el colombianismo *pasante* presenta un $GR = 2,6$ que es bastante elevado para tratarse de una voz local. Sin embargo, esta voz no aparece en el mapa, puesto que existen otras palabras más representativas de este concepto, como *botana*, *picada* o *tapa*. Sin embargo, podemos acceder a todas estas informaciones gracias a la tabla de distribución que acompaña a los mapas en nuestra aplicación virtual y que permite mostrar detalles sobre los resultados de nuestro análisis demolingüístico en cada variante analizada.

Mapa 8 Concepto [TIDBITS]. Ejemplo de voz local con gran vitalidad en un territorio



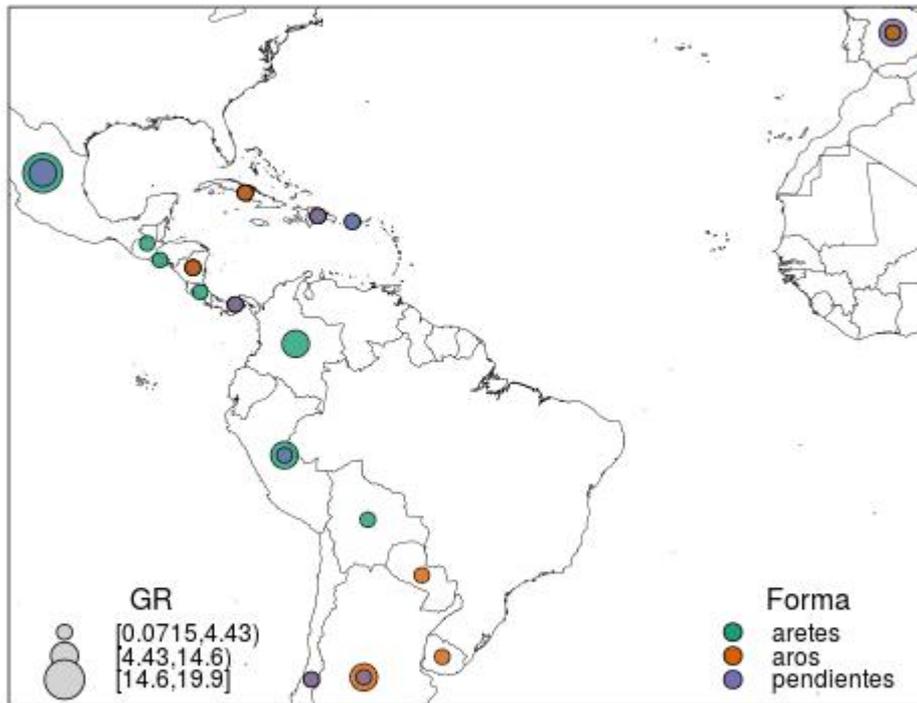
forma	GR	pais
botana	11.85	MX, NI, AR, UY, ES, PA
tapa	15.61	ES, MX, CL, PE, CR, AR, DO, VE, PA, PR
pasabocas	3.51	CO, ES
boquita	4.68	MX, HN, CO, UY, SV, PA, CU
pasapalo	2.68	VE, PA
pasante	2.6	CO
tentempiés.	2.63	MX, SV, CU, ES, PA
entremés	6.47	AR, VE, CO, MX, CU, ES, CL, PR, PA
picada	6.57	AR, CO, VE, PY, BO, SV, ES, HN, CU, PR
aperitivo	6.08	MX, ES, AR, PY, PE, CU, VE, HN, DO, PA, PR

5.6. La convivencia de varias voces

Una de las ventajas de nuestros mapas de representatividad léxica, tanto estáticos como dinámicos, es que muestran la utilización de diversas voces en un mismo territorio, gracias al sistema de representación de círculos proporcionales. En la versión interactiva, además, se puede consultar el porcentaje de uso de cada voz en cada región, si pinchamos con el cursor en cada uno de los círculos.

Si acudimos, por ejemplo, al concepto [EARRINGS], observamos en el mapa de nuestra aplicación que existen tres variantes posibles muy extendidas en la geografía hispánica: *aretes*, *aros* y *pendientes*. Una de ellas, *aretes* (GR [Mx] = 19,94), es la más representativa, que coincide con la preferida del resto del continente americano y convive con la alternativa más peninsular, *pendientes* (GR [Mx] = 5,28).

Mapa 9 Concepto [EARRINGS]. Ejemplo de convivencia de varias formas léxicas



forma	GR	pais
aretes	48.89	MX, CO, PE, GT, BO, EC, CU, DO, CR, ES, SV, VE, PA, NI, HN, CL, UY, PR
pendientes	16.52	ES, MX, AR, PE, CU, BO, CL, NI, DO, PY, PA, UY, PR
aros	13	AR, CL, PY, DO, ES, CU, PA, UY, NI, PR
zarcillos	8.39	VE, ES, DO, PY, NI, CU, PA, PR
arracadas	2.53	MX, ES
argollas	3.36	EC, CU, PE, NI, DO, SV, PA, CR
aritos	5.47	CO, HN, AR, SV, CL, DO
pantallas	0.69	PR, DO
caravanas	0.67	UY, AR
chapas	0.42	NI

5.7. La dispersión de las voces

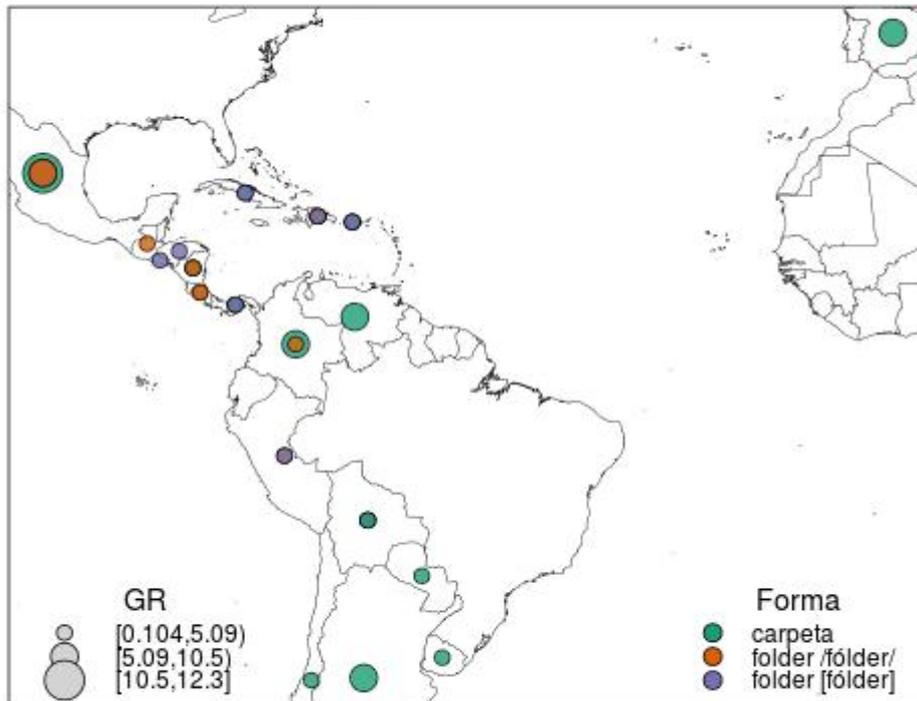
Los resultados de nuestro análisis también permiten interpretar la dispersión de las voces en el territorio hispanohablante. Puede ocurrir que una voz se limite a una única región o que, por el contrario, se extienda por varias áreas geográficas del español. Esto sucede, por ejemplo, con el concepto [FOLDER] o con [FLAT], cuyos mapas respectivos mostramos a continuación.

Si observamos las alternativas léxicas que corresponden al concepto [FOLDER] podemos concluir que el anglicismo *fólder* se extiende por varias áreas geográficas, al tiempo que convive con la variante más representativa del español, *carpeta*. A pesar de ser una voz secundaria en términos de representatividad, se trata de una voz muy extendida en el mundo hispanohablante, que ocupa cuatro áreas dialectales del español.

Lo mismo sucede con la voz *departamento*, que corresponde al concepto [FLAT (US: APARTMENT)] de Varilex. Al igual que sucedía con *fólder*, la voz *departamento* ocupa todas las áreas del dominio hispánico excepto la chilena. Tiene, por tanto, un alto nivel de dispersión.

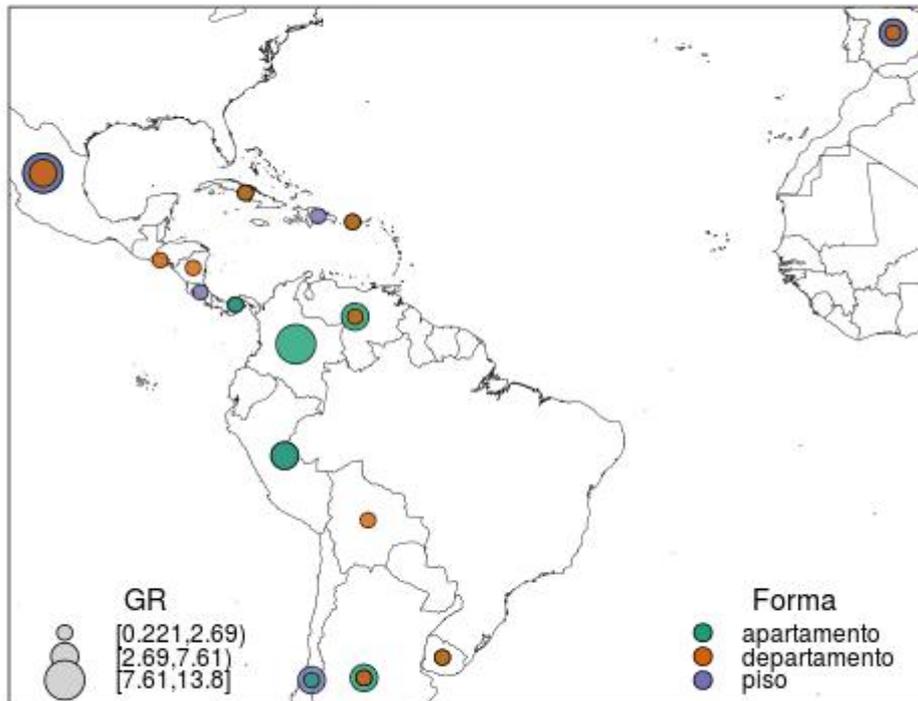
El análisis de la dispersión de las voces es fundamental en los estudios de variación léxica, puesto que algunos términos pueden ser altamente representativos sin ser los más utilizados, precisamente por ser conocidos en todas las áreas dialectales.

Mapa 10 Concepto [FOLDER]. Ejemplo de dispersión en la geografía hispánica



forma	GR	pais
carpeta	56.59	MX, CO, ES, AR, VE, EC, CL, PY, CU, DO, GT, UY, BO, PE, NI, PA, PR, CR
folder /fólder/	17.69	MX, PE, GT, CO, BO, DO, PR, PA, CR, NI
folder [fólder]	13.76	MX, HN, PE, BO, SV, CR, NI, DO, CU, PA, PR
portafolios	3.67	ES, AR, PE, NI, CU, DO, PY, CR, CL, PR, PA
archivador	2.49	ES, VE, CL, PE, AR, BO, NI
sobre	1.53	CO, AR, CU, DO, PR
cartapacio	1.63	SV, AR, PA, ES, NI, PR
file [fáil]	1.74	PE, MX, CU, BO
file /fáil/	0.64	PE, CU, PA, PR
carpetilla	0.26	ES

Mapa 11 Concepto [FLAT]. Ejemplo de dispersión en la geografía hispánica



forma	GR	pais
apartamento	46.88	MX, CO, AR, ES, VE, PE, CU, CL, DO, PR, UY, PA, CR
piso	26.84	MX, PE, ES, CL, AR, DO, HN, CR, PA
departamento	16.65	MX, VE, ES, BO, AR, SV, NI, HN, CU, UY, PR
condominio	0.94	BO, ES, PR

6. Conclusión

Los resultados demolingüísticos de esta tesis ofrecen informaciones relevantes respecto a la variación léxica del español. En el contexto de nuestra investigación podemos afirmar que estos resultados son fundamentales para estudiar la representatividad de las nociones del *PCIC*, haciendo posible contrastar el GR de los términos de referencia con el de otras voces equivalentes del español. Así, es posible presentar una propuesta de ampliación del léxico del *PCIC* con voces altamente representativas del español.

Para la presentación de nuestros resultados hemos diseñado una herramienta virtual que expone cartográficamente la distribución geográfica de las variantes léxicas con mayor GR para cada concepto. Además, en la aplicación incluimos una tabla dinámica que muestra el GR de todas las variantes léxicas así como los países en los que se ha registrado la forma. Además de esta aplicación, hemos diseñado mapas estáticos con las tres variantes con mayor GR de cada concepto estudiado, que incluimos en el Anexo III.

Los resultados que muestran nuestros mapas virtuales y estáticos ofrecen variadas aplicaciones para investigaciones específicas sobre variación. En este sentido, presentan ventajas indiscutibles para el estudio de la variación diatópica, dando cuenta de la estructuración del léxico, la localización de los términos, la vitalidad de las voces, la innovación, el conservadurismo y las relaciones interdialectales. En lo que respecta a la confiabilidad de la información, consideramos que nuestro análisis es fiable puesto que se basa en criterios objetivos y coincide con los datos aportados por otras fuentes de información léxica. Para nuestra propuesta de ampliación del *PCIC* es necesario interpretar exhaustivamente los resultados y realizar una revisión por expertos, con el objetivo de verificar los resultados e identificar aquellas áreas léxicas que requieran precisiones semánticas específicas para garantizar una rigurosa ampliación del material.

CAPÍTULO 6

PROPUESTA DE REVISIÓN Y AMPLIACIÓN DE LAS «NOCIONES ESPECÍFICAS» DEL *PCIC*

1. Introducción

El objetivo central de esta investigación es hacer una revisión crítica de las «Nociones específicas» del *PCIC* y elaborar una propuesta de ampliación con voces representativas de todo el dominio hispánico. Tal y como presentamos en el capítulo 1, el *PCIC* es la obra de referencia para la enseñanza de ELE. El hecho de que no exista un inventario léxico que integre la variación dialectal hace necesario presentar un repertorio nocional que seleccione y secuencie rigurosamente el léxico del español. En el capítulo 2 de esta tesis mostramos cómo este vacío bibliográfico se debe a la complejidad que entraña este componente lingüístico en lo que se refiere a su sistematización y delimitación geográfica. En el capítulo 3 hemos propuesto un cálculo demolingüístico que permite operacionalizar el concepto de representatividad léxica en una variable cuantificable, que puede medirse y expresarse numéricamente. En el capítulo 4 hemos mostrado nuestros resultados demolingüísticos, estudiando la distribución de 182 conceptos del español y calculando su Grado de Representatividad (GR) y el de sus correspondientes variantes léxicas, mediante una aplicación demolingüística con mapas geográficos y tablas de distribución, que constata el alcance y las posibilidades que brinda el estudio de representatividad léxica que hemos defendido a lo largo de esta tesis.

En las siguientes páginas incluimos una propuesta personal de revisión y ampliación de las nociones léxicas del *PCIC*, integrando voces de otras variedades dialectales. Nuestra aportación cumple con los criterios de rentabilidad y extensión en la selección léxica para la enseñanza de ELE y aporta una visión más completa y representativa del conjunto del léxico hispánico. No obstante, una ampliación de estas características debe contrastarse con un equipo de especialistas que aporte nuevos datos y reflexiones acerca de la representatividad de las voces planteadas y la pertinencia de su inclusión en un material especializado para la enseñanza de español. Esta tesis constituye un primer acercamiento a la variación léxica del *PCIC*, fundamentado en los datos científicos recogidos por el equipo de Varilex, y a través de una metodología demolingüística que analiza rigurosamente los datos de variación. Nuestro deseo es que esta propuesta genere reflexiones y nuevas aportaciones científicas para que, junto a un equipo coordinado de especialistas, se proponga una ampliación léxica definitiva de las «Nociones específicas» del *PCIC*.

El método de trabajo que hemos seguido para revisar la representatividad de las «Nociones específicas» comienza por identificar el GR de los términos de referencia del *PCIC*. Ponemos especial atención en los términos que tienen un GR bajo porque indican la posible existencia de otras voces con mayor representatividad en el mundo hispánico. No obstante, no siempre encontramos *una voz* especialmente representativa; a veces ocurre que existen varias voces dispersas, en diferentes países, sin que ninguna de ellas alcance alta representatividad. También puede ocurrir que, para un término de referencia con un GR bajo, no podamos hacer un estudio de variación, puesto que nuestra tesis se circunscribe a los datos que ha estudiado el equipo de Varilex.

Cuando encontramos voces alternativas a las del *PCIC* que cuentan con mayor representatividad, proponemos su incorporación al repertorio léxico propuesto por el Instituto Cervantes. Presentamos, por tanto, voces equivalentes semánticamente pero que son más representativas en los términos demolingüísticos planteados en el capítulo 3. Estas voces pueden ser *panhispánicas*, empleadas en España y América, o específicamente *americanas*, palabras que no existen en el

español peninsular. En todo caso siempre proponemos voces representativas del español y evitamos los términos regionales o locales.

El objetivo último de esta tesis es revalorizar las nociones del *PCIC*, ofreciendo el primer repertorio léxico del español, secuenciado en niveles de referencia, que incorpora las nociones más representativas de toda la población hispanohablante. Esta ampliación no está exenta de dificultades: si la variación es en sí misma una característica intrínseca al uso de la lengua, cuando además establecemos relaciones entre los múltiples subsistemas lingüísticos del español, la clasificación del léxico se vuelve realmente compleja.

Tal y como advertimos en el capítulo 2, las particularidades léxicas de las diferentes zonas geográficas son muy difíciles de sistematizar. No es factible elaborar una teoría intralingüística de la variación, puesto que el léxico constituye un sistema abierto, inacabado y potencialmente infinito, que no está estructurado (Chaudenson, 1993). De hecho, en la mayoría de los casos, incluso las voces léxicas locales, que denominamos *regionalismos* (Rona, 1969), también se utilizan —o al menos, se conocen— en otras zonas geográficas, lo que en nuestra opinión impide hablar de regionalismos *strictu sensu* (Rabanales, 1953).

Ante la dificultad de sistematizar el componente léxico y la complejidad de definir áreas dialectales en función del vocabulario, se necesitan herramientas de análisis que permitan establecer categorías en función de una medición ordinal. Los cálculos demolingüísticos expuestos en el capítulo 3 nos permiten medir un concepto abstracto como la representatividad, en función de criterios geográficos y demográficos. A pesar de todo, es preciso reflexionar sobre la complejidad que encontramos a la hora de proponer equivalencias léxicas en el *PCIC*, y sobre los tipos de oposición que pueden surgir en los estudios de variación.

2. Cuestiones previas sobre las equivalencias léxicas

Las nociones que se proponen en esta ampliación permiten determinar que para una forma *x* del *PCIC*, existe otra *y* con una distribución geográfica específica.

Ambas palabras conforman lo que denominamos *oposición geolectal simple x / y*. No existen matices ni precisiones de significado, sino que para un mismo concepto, conviven diversas voces según las regiones. Es el caso de la oposición *vaqueros* (España) – *jeans* (México), por ejemplo.

Pero más allá de las equivalencias biunívocas entre dialectos, el estudio de la variación léxica exige tener en cuenta, además, los subsistemas léxico–semánticos de cada una de las variedades, así como los valores connotativos y denotativos de las variantes (Ávila, 1997). Es posible que una misma palabra cambie su significado según la región en la que se utilice o que, en determinadas zonas, lleve asociados rasgos de uso relacionados con el registro, el grupo social o sus connotaciones históricas. También puede ocurrir que dentro de un mismo dialecto se produzca una oposición interna, es decir, que convivan diferentes voces para referirse a un mismo concepto. En este caso generalmente una de las variantes lleva asociada algún tipo de marca: histórica, si una de las dos voces está en desuso; diastrática, cuando una variante es exclusiva de un grupo social; o diafásica, cuando una de las voces se emplea solo en determinados registros.

Así pues, en la comparación léxica dialectal existen diferentes tipos de relaciones¹. Vamos a resumirlas en tres grandes grupos:

(1) Oposiciones *diatópicas simples*:

- a. *falda* (Es) – *pollera* (Ar)
- b. *ordenador* (Es) – *computadora* (Mx)

Las oposiciones incluidas en (1) constituyen oposiciones diatópicas simples. Son los llamados *geosinónimos*. El ejemplo (1a) señala que la voz española *falda* tiene el mismo significado que la voz argentina *pollera*. De la misma manera, según (1b), para referirse al concepto [COMPUTER], en España se utiliza el término *ordenador* mientras que en México se prefiere *computadora*. La oposición entre

¹ Para comprender en detalle todas las posibles oposiciones que podemos encontrar en el estudio de la variación léxica del español recomendamos la lectura imprescindible de *Variación léxica: connotación, denotación, autorregulación* de Ávila (1997).

ambas voces se basa en su distribución geográfica. Este es el tipo de oposición que pretendemos incluir en nuestra ampliación de las «Nociones específicas» del *PCIC*, aunque, como veremos a continuación, no siempre existen estas equivalencias biunívocas simples.

(2) Oposiciones *connotativas*:

- a. *sujetador / sostén* (Es) – *sostén* (Ch)
- b. *gafas / lentes* (Es) – *lentes / anteojos* (Mx)

En el segundo grupo de oposiciones léxicas mostramos una alternancia de voces que trasciende la variación diatópica simple. En los ejemplos de (2) observamos que, dentro de un mismo subsistema lingüístico, pueden coexistir dos voces léxicas, originando oposiciones tanto a nivel interno como externo. En estos casos, alguna de ellas suele adoptar valores connotativos específicos, en función del sexo, el grupo de edad, el registro o el marco histórico, siendo excepcionales los casos de sinonimia absoluta. Si en el español de España coexisten *sujetador* y *sostén* o *gafas* y *lentes*, buscaremos algún tipo de matiz sociolingüístico que pueda explicar estas alternancias.

Según (2a), para el concepto [BRASSIERE] de Varilex, existe una oposición interna en el español de España (*sujetador / sostén*), que a su vez se opone externamente a la voz *sostén* en Chile. En nuestro estudio es fundamental señalar tanto la alternancia externa entre España y Chile como la convivencia de dos voces en el español de España, puesto que el análisis de las voces léxicas —activas y pasivas— en las diferentes regiones, nos permitirá encontrar semejanzas dialectales interregionales.

En el caso (2b), observamos que la voz *lentes* se utiliza tanto en España como en México, lo que la convierte en una voz representativa del español. No obstante, la voz *lentes* alterna con otras voces dentro de los subsistemas lingüísticos de ambos países. En España, la variante *lentes* se entiende y se utiliza, aunque en contextos más restringidos; la voz preferida es *gafas*. En cambio, en México, *lentes* es la voz preferida, alternando con *anteojos* en contextos de uso minoritarios. Así, las voces

gafas y *anteojos* se oponen a nivel externo entre España y México, mientras que el término *lentes* pertenece a un español compartido por ambos países.

Aunque puedan coexistir con otras voces a nivel interno, las palabras compartidas por varios países hispánicos son las más representativas del español, puesto que garantizan la comunicación en una mayoría de regiones y contextos. Es importante destacar que las voces más representativas no se corresponden necesariamente con las más frecuentes, sino con las más conocidas. Esta es una de las fortalezas de los datos de Varilex: la capacidad de medir el léxico pasivo de los hispanohablantes. La posibilidad que ofrecen las encuestas de marcar varias voces dentro de una lista específica de palabras, permite registrar tanto los términos que se utilizan como los que se conocen pero no son preferentes en el uso.

(3) Oposiciones *denotativas*:

- a. *lapicero* (Es) – *lapicero* (Pe)
- b. *papalote* (Mx) – *papalote* / *coronel* (Cu)
- c. *boleto* (Mx) – *boletín* / *pasaje* / *entrada* (Cu)²

El último tipo de oposición fundamental para el estudio de la variación léxica se refiere a las oposiciones *denotativas*, que presentan voces idénticas con significados distintos en los diversos subsistemas del español. Así, según (3a), la voz *lapicero* se refiere a realidades distintas en España y Perú. El *lapicero* en Perú hace referencia al utensilio de escritura cargado con tinta, que en España se denomina *bolígrafo*.

También puede ocurrir que los términos respondan a oposiciones connotativas debido a una segmentación diferente del contenido. Así, (3b) indica que el ‘juguete de papel o tela que se eleva contra el viento, se sostiene por medio de una cuerda y lleva generalmente una cola larga’ se denomina *papalote* tanto en México como en Cuba, aunque en Cuba se refiere únicamente a una cometa de tamaño pequeño; si este juguete tiene gran tamaño, no se llamará *papalote* sino *coronel*. Otro ejemplo de este tipo aparece en (3c), según el cual, el ‘papel de tamaño pequeño que

² Los ejemplos (3b) y (3c) los tomamos de Ávila (1997).

sirve de pasaje o entrada' se denomina *boleto* en México, mientras que en Cuba existen tres variantes: el *boletín*, para viajar en tren; el *pasaje*, para viajar en avión; y la *entrada*, para acudir a un espectáculo como el cine o el teatro.

Las voces compartidas que establecen oposiciones denotativas entre dialectos son las más difíciles de percibir para los hispanohablantes, y las más complejas de incorporar en la didáctica de ELE. El profesor de español debe conocer las estrategias de *autorregulación* fundamentales que permiten a los hablantes, nativos y no nativos, superar las posibles dificultades comunicativas.

Una vez explicados los tres grandes tipos de oposición que generan las alternancias léxicas, es importante expresar que nuestra propuesta de ampliación de las «Nociones específicas» pretende incorporar, primeramente, aquellas variantes panhispánicas que equivalen a los términos del *PCIC* pero que son más representativas del español global. En estos casos de convivencia de varias voces a nivel interno pueden aparecer oposiciones connotativas o denotativas entre las variantes, como se reflejaba en el ejemplo (2): si en España se utilizan tanto *sujetador* como *sostén*, o tanto *gafas* como *lentes*, es importante marcar las diferencias entre estas equivalencias.

En relación a la incorporación de voces alternativas americanas, nuestra tesis propone incluir oposiciones diatópicas simples, como las que recoge el ejemplo (1). Según Ávila (1997), la mayoría de las oposiciones que recoge Varilex se corresponden con este tipo de oposición. En los casos en los que encontremos oposiciones connotativas o denotativas relevantes, también lo mencionaremos.

No descartamos que puedan existir más oposiciones, internas o externas, en alguno de los subsistemas lingüísticos del español. Un análisis exhaustivo y pormenorizado del comportamiento de las variantes léxicas en cada uno de los países donde se habla español, implicaría la realización de numerosas monografías dialectales. Un trabajo de este tipo sería fundamental para comprender e interpretar el fenómeno de la variación dentro de cada sistema, y entre los diferentes dialectos,

pero, además de ser una empresa prácticamente inabarcable, no permitiría plantear una mirada de conjunto interdialectal sobre la variedad en tantos países.

Nuestra tesis, en cambio, se enmarca en el ámbito de la geografía lingüística y pretende tratar la variación léxica como fenómeno global, presentando la distribución de cientos de formas a lo largo de miles de kilómetros. La investigación que presentamos en este trabajo muestra un panorama general de la distribución de la variedad, con enormes cantidades de datos, que no habría posible desde la dialectología tradicional. Los resultados que hemos presentado en el capítulo 4 son fruto de un análisis de la variación más relacionado con las pretensiones de los atlas lingüísticos de gran dominio, que permiten estudios de gran magnitud como el que presentamos a continuación. En todo caso, no podemos perder de vista que el análisis de las variantes léxicas es un asunto verdaderamente complejo, para el que faltan aún muchas investigaciones por emprender. Esta tesis pretende ser un acercamiento novedoso al respecto y se circunscribe al ámbito de la geolingüística.

3. Acercamiento a la variedad léxica

Antes de presentar nuestra propuesta de ampliación veamos un ejemplo del procedimiento que hemos seguido para valorar la posible inclusión de variantes léxicas en las «Nociones específicas» del *PCIC*.

Si acudimos al apartado «12.2. Ropa, calzado y complementos» del nivel A2 de las «Nociones específicas» del *PCIC*, encontramos la voz *calcetines*, que corresponde al concepto [SOCKS] de Varilex y se define como ‘Prenda de punto que recubre el pie’. La base de datos de Varilex registra tres variantes léxicas representativas para este concepto (además de otras que no tendremos en cuenta por su baja representatividad): *calcetines*, *medias* y *soquetes*.

La palabra que incluye el *PCIC* es la que tiene mayor GR: *calcetines* (GR = 47.23). Esta noción se considera representativa del español pero observamos que existe una voz con una representatividad similar, *medias* (GR = 38.79), proporcionada especialmente por informantes americanos.

Con el objetivo de validar los datos de Varilex y determinar la extensión y el significado del término, buscamos la palabra *medias* en el *Diccionario de americanismos*:

medias f. *ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur*. Calcetín, prenda que cubre el pie y llega hasta la pantorrilla.

Gracias a esta herramienta de referencia, podemos corroborar que *medias* se corresponde con el concepto [SOCKS] y se utiliza en 16 países del mundo hispanohablante, distribuyéndose a lo largo de todas las áreas dialectales americanas: Centroamérica [–Mx], Antillas, Caribe continental, zona andina, zona rioplatense y Chile³.

Si pretendemos dotar de una mirada integradora de la variedad léxica al inventario de «Nociones específicas» del *PCIC*, consideramos que la palabra *medias* debería aparecer junto al término de referencia *calcetines*.

En relación con la tercera voz registrada por Varilex, *soquetes* (GR = 6,29), no resulta pertinente considerar su inclusión por dos motivos: en primer lugar, tiene un GR muy inferior a las otras opciones; por otro lado, si acudimos al *Diccionario de americanismos*, observamos que *soquetes* conforma una oposición denotativa respecto a *calcetines* y *medias*, puesto que tiene una diferencia de significado. Un *soquete* es un tipo de calcetín corto, que cubre el pie hasta el tobillo, por lo que esta voz alternativa no está al mismo nivel que *calcetines* – *medias* y no nos interesa para la ampliación.

4. Descripción general de nuestra propuesta de revisión y ampliación

En este apartado vamos a presentar nuestra propuesta de revisión y ampliación de las «Nociones específicas» del *PCIC*. A través del análisis de los datos de Varilex, y mediante los cálculos demolingüísticos definidos en el capítulo 3,

³ Considerar la distribución geográfica de las variantes léxicas es fundamental para nuestro trabajo puesto que, como ya hemos explicado, el léxico que presentamos a los estudiantes debe cumplir los criterios de extensión, rentabilidad y frecuencia, sobre todo en los primeros niveles de aprendizaje (Chacón García, 2014).

proponemos una ampliación del repertorio léxico del *PCIC* que se basa en los siguientes ejes:

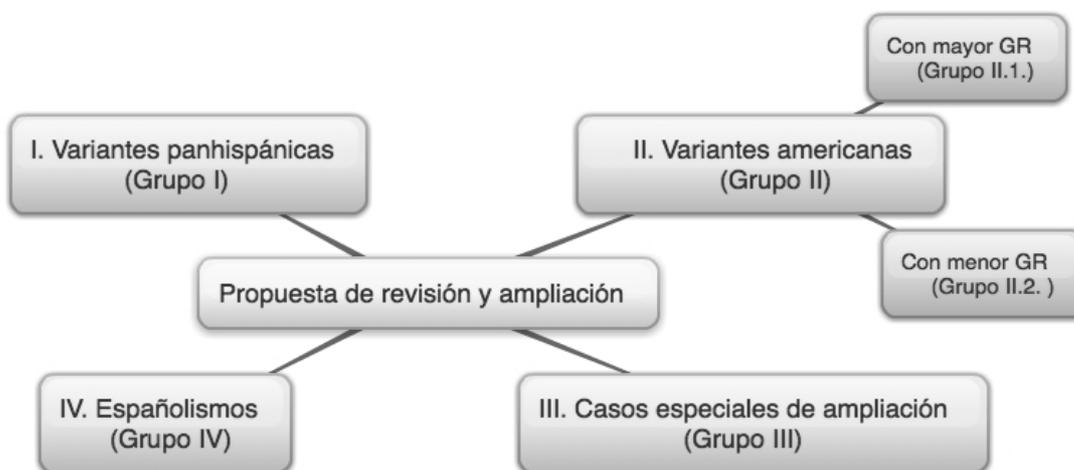


Figura 17: Propuesta de revisión y ampliación de las «Nociones específicas» del *PCIC*

Nuestra aportación comienza incorporando una serie de voces panhispánicas que tienen mayor GR que los términos del *PCIC* (Grupo I). El objetivo es proponer, junto a los términos de referencia, otras voces complementarias que son conocidas tanto en España como en América y que gozan de mayor representatividad en la geografía global del español. Estas nociones son las fundamentales para un estudiante de ELE puesto que constituyen la base común del idioma, que le permitirá comunicarse con una mayoría de hablantes en un elevado número de contextos.

En segundo lugar, proponemos voces específicamente americanas. Por un lado, incluimos los americanismos que tienen mayor GR que algunos términos del *PCIC* (Grupo II.1). La incorporación de estos americanismos con alta representatividad es fundamental para cualquier usuario de este plan curricular.

Por otro lado, presentamos americanismos que, aun teniendo un GR inferior al de los términos del *PCIC*, constituyen una muestra de español global, representativo y rentable para los estudiantes (Grupo II.2.). En esta investigación no pretendemos hacer del *PCIC* un tratado exhaustivo de la variación léxica hispánica, con regionalismos o palabras privativas de cada zona lingüística. Nuestro interés se

basa en aportar palabras que gocen de gran vitalidad y uso en la geografía hispánica. Por ello, la inclusión de estos americanismos secundarios, con menor GR que los términos de referencia, se justifica en función de la extensión geográfica, siendo necesario que las voces se utilicen, al menos, en tres áreas dialectales. En caso contrario tendremos que justificar las ventajas de su inclusión.

En cada uno de los grupos anteriores (I y II) aportamos, además, una tabla con voces que podríamos considerar dudosas para la ampliación, cuya incorporación es cuestionable debido a su bajo GR. No obstante, quedan ahí recogidas por si satisfacen el interés particular de algún investigador.

Por último, hemos seleccionado aquellas palabras que constituyen casos aislados o excepcionales que no pueden incluirse en ninguno de los grupos anteriores. Se trata de nociones que merecen una explicación específica debido a particularidades de significado o distribución (Grupo III).

Además de incluir las ampliaciones mencionadas, en esta tesis consideramos imprescindible señalar también aquellas palabras del inventario de «Nociones específicas» que son exclusivas de la variedad española, que denominamos españolismos (Grupo IV). Estas nociones se tratan específicamente en el §5 de este capítulo. La noción *borde*, por ejemplo, del nivel C, se utiliza exclusivamente en España con el significado de persona impertinente o antipática (DLE, 2015). Al no existir ningún diccionario de españolismos, y dado que muchos manuales de ELE adoptan la variedad de España como variedad preferente, es fundamental para los usuarios del *PCIC* localizar aquellas nociones del inventario léxico que son exclusivas del español de España.

Una vez explicados los ejes que seguimos para la ampliación de las «Nociones específicas», presentamos a continuación nuestra propuesta, cuyo objetivo último es ampliar el potencial del inventario léxico del *PCIC*.

4.1. Voces panhispánicas (Grupo I)

Las palabras panhispánicas son aquellas que tienen el mismo significado tanto en España como en América. Así son la mayoría de las palabras del español, comunes y compartidas, aunque coexisten generalmente con otras voces sinónimas en las diferentes regiones.

Para un estudiante de ELE, los panhispanismos son fundamentales en el aprendizaje. Se trata de palabras que no pueden faltar en los manuales de español puesto que constituyen la médula del idioma, el español común más rentable y representativo que facilita la comunicación con los hispanohablantes en la mayoría de contextos posibles. Se trata del verdadero estándar.

Aunque la labor del docente de español es adaptar las nociones léxicas a las necesidades del grupo de aprendizaje y al contexto de enseñanza, las palabras panhispánicas con alto GR son indispensables. Es preferible evitar —sobre todo en niveles inferiores— las palabras locales en favor de nociones léxicas que se ajusten al criterio de rentabilidad pedagógica (Chacón García, 2014).

Sabemos que los términos de referencia de las «Nociones específicas» se corresponden con la variedad centro norte peninsular española. Según nuestro análisis demolingüístico de los datos de Varilex, encontramos que para algunos de estos términos, existen voces sinónimas panhispánicas que cuentan con mayor GR. Así, junto a los términos de referencia, presentamos otras alternativas que también pertenecen a la variedad peninsular española pero con mayor extensión geográfica.

Tengamos en cuenta que cuando dos variantes léxicas coexisten en un mismo territorio suele darse una variación de tipo connotativo, con marcas de uso de los hablantes. Así, una de las voces podrá ser más habitual que la otra en un grupo social determinado, en ciertos registros de uso o en un periodo temporal concreto. Podría ocurrir también que uno de los dos términos presentara particularidades semánticas. En ese caso se trataría de una variación denotativa con cambio de significado, por lo que descartaríamos su incorporación.

4.1.1. Propuesta específica de panhispanismos

Teniendo como referencia las respuestas de Varilex, consideramos que una noción léxica es un panhispanismo *representativo* del español cuando:

(a) Se registra tanto en España como en América.

(b) Aparece, al menos, en cuatro de las siete zonas geográficas del español: España, Centroamérica, Antillas, Caribe continental, zona andina, zona rioplatense y Chile.

(c) Ocupa la primera o la segunda posición en términos de representatividad (GR) respecto a todas las respuestas registradas en la base de datos de Varilex.

Los criterios anteriores garantizan que la noción panhispánica que se proponga para el *PCIC* goce de gran extensión en el mundo hispanohablante (criterio *b*) y forme parte del léxico activo o pasivo actual de los hispanohablantes (criterio *c*).

A continuación presentamos nuestra propuesta de ampliación, con 23 voces panhispánicas que tienen mayor GR que los términos de referencia del *PCIC*. La Tabla 10 incluye:

- El término de referencia del *PCIC*, con su GR.
- La voz alternativa panhispánica más representativa, con su GR.
- El concepto en inglés que atribuye Varilex para referirse a ambas voces.
- La definición aportada a los informantes en las encuestas de Varilex.

En relación con la ortografía, en este trabajo seguimos el *Diccionario panhispánico de dudas* (2005).

Tabla 10
Propuesta de ampliación de variantes panhispánicas en el PCIC

GRUPO 1. Variantes panhispánicas con mayor Grado de Representatividad					
Término de Referencia	GR	Alternativa panhispánica	GR	Concepto Varilex	Definición de Varilex
A1-A2					
jersey	9,11	suéter	51,46	SWEATER	Prenda de vestir de tejido de punto, con mangas que cubre aproximadamente hasta la cintura.
gafas	26,13	lentes	40,29	GLASSES	Anteojos usados corrientemente que se apoyan en la nariz y en las orejas. (No son de color oscuro para el sol.)
bañador	8,3585	traje de baño	60,59	SWIMSUIT	Traje de baño femenino.
nevera	25,76	refrigerador	42,75	REFRIGERATOR	Aparato que mantiene alimentos a baja temperatura.
bocadillo	15,9	sándwich ⁴	49,11	SANDWICH	Pan que lleva jamón, queso, etc. adentro.
salón	2,03	sala	45,07	LIVING ROOM	Sala para estar cómodo, para ver la televisión, etc.
sujetador	7,04	sostén	36,37	BRASSIERE	Prenda interior femenina que se utiliza para sujetar los senos.
reservado	4,38	retraído	19,75	MEEK	El que está siempre callado y no trata con nadie.
enfadado	8,39	enojado	26,62	ANGRY	Que tiene un disgusto muy fuerte.

⁴ Seguimos la ortografía recomendada por el *Diccionario panhispánico de dudas* (2005).

B1-B2						
					TURN (TO THE RIGHT)	Cambiar de dirección.
girar	8,13	doblar	27,61	TURN (TO THE RIGHT)		Cambiar de dirección.
matrícula	14,83	placa	68,47	NUMBER PLATE		Placa visible con el número de registro oficial.
torta	1,33	bofetada	33,6	SLAP		Golpe dado en la cara con la mano abierta sobre la cara.
nota	13,64	puntuación	33,5	MARK		Nota que se pone a un ejercicio.
arrogancia	12,11	soberbia	25,88	ARROGANCE		Orgullo excesivo o sentimiento de superioridad frente a los demás.
valentía	12,23	valor	31,42	COURAGE		Cualidad de valiente.
carrera	1,26	licenciatura	50,49	DEGREE		Grado académico que permite ejercer una profesión al terminar los estudios en la universidad.
callado	3,58	retraído	19,75	MEEK		El que está siempre callado y no trata con nadie.
enfadarse	11,7	enojarse	27,45	GET ANGRY		Enojarse, enfadarse: Cuando supo la verdad, [***] mucho.
C1-C2						
coraje	23,16	valor	47,68	COURAGE		Cualidad de valiente.
atiborrarse	0,59	hartar, hartarse	8,53	EAT TOO MUCH		Hartarse de comer: El domingo, después del partido se fueron todos a comer chuletas y [***]
forofos	0,41	aficionados	44,41	FANS		Conjunto de personas que van regularmente a ver una competición deportiva.
tener buen saque	2,68	comer como una vaca	51,38	EAT LIKE A HORSE.		Comer mucho: No sé cómo no revientan de tanto comer. Estos muchachos (***)

La tabla anterior muestra pares de voces sinónimas que se podrían emplear indistintamente dentro del subsistema lingüístico del español de España. Junto a la noción propuesta por el *PCIC*, anotamos variantes léxicas panhispánicas que son más representativas según el porcentaje de GR, calculado a partir de los datos de Varilex y del peso demográfico de cada país.

Si tomamos, por ejemplo, la alternancia *jersey* – *suéter*, observamos que aunque ambas palabras se usan en España, *jersey* (GR = 9,11) parece la voz más frecuente en el español peninsular. Si consultamos el CORPES, confirmamos nuestra intuición: *jersey* aparece en 949 documentos, de los cuales, 813 corresponden a la variedad lingüística de España. En cambio, *suéter* (GR = 51,56) se documenta en todas las áreas dialectales del español, lo que le otorga más representatividad en el dominio hispánico. En este caso, no encontramos diferencias significativas entre ambas voces, siendo intercambiables en los mismos contextos, por lo que la incorporación de *suéter* junto a *jersey* en el *PCIC* contribuye a mejorar la representatividad de las «Nociones específicas».

Sin embargo, como ya hemos apuntado, cuando en un dialecto encontramos dos variantes para designar un mismo concepto, puede ocurrir que una de ellas esté marcada connotativamente. De este modo, una de las voces se considera el término no marcado frente a otra que se utiliza en contextos específicos. Por ejemplo, *gafas* y *lentes* en España son intercambiables, aunque *lentes* se reserva a un ámbito culto o literario.

Incluir en el *PCIC* estas voces sinónimas panhispánicas, más rentables y extensas en el ámbito de la lengua española tiene gran importancia, sobre todo porque, a veces, la voz propuesta por el Instituto Cervantes (que frecuentemente coincide con las voces incluidas en los manuales de ELE) es exclusiva del español de España, como veíamos en la oposición *jersey* – *suéter*. Si existe una voz sinónima válida para el español peninsular, que es conocida también en todo el ámbito hispanohablante, nos parece importante anotarla. Otro ejemplo similar lo encontramos en la alternancia *bañador* – *traje de baño*. La voz recogida por el *PCIC*,

bañador (GR = 8,35), se limita al español de España y tiene un GR considerablemente inferior a la noción *traje de baño* (GR = 60,59), válida en España y América y que, además, es más transparente para el estudiante de ELE. Recordemos la importancia del criterio de transparencia en la selección léxica, que recoge el *MCER* (2001) en el §6.4.7.3.

La inclusión de voces panhispánicas en el currículo de la enseñanza de español es ineludible. Gracias a nuestro análisis podemos conocer aquellas voces compartidas por la mayoría de hispanohablantes e introducirlas en manuales de referencia y repertorios léxicos como el que nos ocupa. Recordemos que no siempre se trata de incorporar las voces más utilizadas sino las más conocidas en el conjunto de la geografía hispánica.

4.1.2. Otros panhispanismos sujetos a valoración

Además de la ampliación de palabras panhispánicas que proponemos en la Tabla 10, hemos encontrado otras voces comunes a España e Hispanoamérica que, según nuestros cálculos, cuentan con un GR ligeramente superior a los términos de referencia del *PCIC*. Dado que estas palabras no gozan de una representatividad significativamente mayor, su posible inclusión en el inventario de «Nociones específicas» debe valorarse de forma individual. La Tabla 11 muestra nuestros resultados:

Tabla 11*Otras nociones panhispánicas, con GR ligeramente superior a las del PCIC*

Valoración	Concepto	Término <i>PCIC</i>	GR	Panhispanismo alternativo	GR
Palabras <u>descartadas</u> para la ampliación	BATHTUB	bañera	32,51	tina	36,72
	STAPLER	grapadora	22,77	engrapadora	33,16
	LAZY	vago	21,16	flojo	27,41
	TREMENDOUS	brillante	4,28	bárbaro	10,43
	VALISE	maletín	36,73	portafolios	43,48
	PANTY HOSE	medias	4,5	pantis	8,17
	THIN (SKINNY)	escuálido	4,96	flacucho	6,72
Palabras <u>propuestas</u> para la ampliación	YOUNG ADULT	joven	25	muchacho	26,7
	OLD MAN	anciano	21,84	viejito	25,41

4.1.2.1. Palabras panhispánicas ligeramente más representativas que descartamos para la ampliación

Aunque la serie *tina*, *engrapadora*, *flojo*, *bárbaro* se conoce en el español de España, las variantes *bañera*, *grapadora*, *vago*, *brillante* son las más habituales, llegando a desplazar a sus alternativas panhispánicas. La voz *engrapadora*, por ejemplo, aunque se incluye en el *Diccionario de Lengua Española* (2014), no se registra en ningún documento del CORPES publicado en España. Por este motivo, y teniendo en cuenta que no hay una gran diferencia de GR entre ellas, no nos parece pertinente incluirlas en el *PCIC*.

El caso de *maletín* – *portafolios* es muy interesante puesto que pone de manifiesto cómo los diccionarios no siempre se corresponden con el uso lingüístico de los hablantes. Si consultamos el *Diccionario de la Lengua Española* (2014) encontramos que *maletín* y *portafolios* tienen acepciones distintas. Cabe preguntarse por qué en los datos de Varilex responden al mismo concepto. Si consultamos otros

diccionarios de uso como el *Gran Diccionario de la Lengua Española* (1996) o el *Diccionario del Español Actual* (2011), confirmamos que estas palabras se utilizan, en realidad, como voces sinónimas en español, tal y como presentamos a continuación.

Tabla 12

Oposición maletín – portafolios

DLE	maletín	m. Especie de maleta pequeña.
	portafolios	m. Cartera de mano para llevar libros, papeles, etc.
Gran Diccionario de la Lengua Española	maletín	m. Maleta pequeña, por lo general rectangular, para guardar o transportar papeles, documentos u otros útiles de uso profesional.
	portafolios	m. Cartera de mano, por lo general rectangular, utilizada para guardar o transportar papeles, documentos u otros útiles de uso profesional.

Respecto a la oposición *medias – pantis*, observamos que el término *medias* (GR = 4,5), propuesto por el PCIC, tiene menor representatividad que el término *pantis* (GR = 8,17). No obstante, incluir la variante *pantis* en el PCIC tampoco resulta de gran interés puesto que ambas variantes tienen un GR bastante bajo. Como veremos en la Tabla 13, existe una alternativa americana, *pantimedias* (GR = 31,61), que podría acompañar al término *medias* debido a su extensión y representatividad.

Por último, la oposición *escuálido – flacucho* tampoco resulta pertinente porque *flacucho* deriva de *flaco* y constituye una forma marcada gramaticalmente por el sufijo *—ucho* que adopta los rasgos [+ coloquial, + despectivo].

4.1.2.2. Palabras panhispánicas ligeramente más representativas propuestas para la ampliación

Existen, en cambio, dos términos del mismo campo semántico con GR ligeramente superior que sí proponemos para la ampliación: *viejito*, con un sufijo

lexicalizado que se ha extendido de forma significativa tanto en América como en España; y, dentro del mismo campo semántico, *muchacho*, por su vitalidad en todo el territorio hispánico.

4.1.3. Voces panhispánicas excluidas

Antes de terminar este primer apartado dedicado a la búsqueda de nociones panhispánicas representativas, señalamos que, en nuestras consultas de datos de Varilex, hemos encontrado algunas oposiciones léxicas que se refieren a realidades diferentes desde un punto de vista semántico, por lo que son oposiciones que no nos interesan para la ampliación. Por ejemplo, para el término de referencia *pañuelo*, incluido en el *PCIC*, encontramos la oposición *pañuelo* – *chal*, que no hemos valorado puesto que incluye un cambio de significado relacionado con la forma. Según los diccionarios consultados, *chal* se caracteriza por tener forma rectangular (‘mucho más largo que ancho’, cita el DLE) mientras que *pañuelo* tiene, por lo general, forma cuadrada. Descartamos la oposición por no constituir una alternancia diatópica.

Asimismo, la oposición *tormenta* – *aguacero* también conlleva un cambio de significado según el DLE: el *aguacero* se caracteriza por su corta duración y la *tormenta* por venir ‘acompañada de aparato eléctrico’, según el DLE.

También se han descartado las oposiciones panhispánicas que son muy cercanas desde el punto de vista formal, cuya ampliación no resulta significativa para el *PCIC*, por ejemplo, la alternancia *tener estrella* – *nacer con estrella*.

4.2. Voces americanas con mayor GR que los términos de referencia del *PCIC* (Grupo II.1.)

El segundo grupo que proponemos para la ampliación del *PCIC* está formado por voces americanas. Una de nuestras hipótesis de investigación era verificar la existencia de americanismos generales o *panamericanismos*, que no se emplean en

España. Coincidiendo con Haensch (2002), nuestras consultas en la base de datos de Varilex constatan la existencia de estas voces.

Si tenemos en cuenta que el continente americano concentra un 90 % de los hablantes de español, estos panamericanismos constituyen una ausencia fundamental del *PCIC*. Como ya hemos explicado, en nuestra tesis proponemos una ampliación de las «Nociones específicas» que aporte una mirada ecuánime ante la variedad lingüística del español, según los criterios de rentabilidad y extensión del léxico para la enseñanza de ELE.

La inclusión de variantes léxicas americanas resulta incuestionable, sobre todo, cuando las voces americanas tienen mayor representatividad que los términos de referencia del *PCIC*. Estas voces americanas con mayor GR son las que se incluyen en este apartado. No pretendemos cambiar las nociones del *PCIC* por otras voces alternativas americanas pero sí nos parece fundamental que el docente pueda acceder a otras variedades del español —que además son más representativas— para que integre y secuencie el componente léxico con rigor y exhaustividad en la clase de ELE.

Así, las voces americanas propuestas son palabras representativas del español, que conforman una oposición diatópica simple respecto a los términos del *PCIC*, como explicamos en §2. Por tanto, la diferencia entre las voces está ligada a su distribución en el espacio, sin que cambien valores de uso ni de significado.

Somos conscientes de la dificultad que conlleva proponer voces *americanas*, cuando lo más frecuente es que existan numerosos términos locales en el repertorio léxico de cada país. No obstante, gracias a los datos de Varilex y al cálculo matemático del GR, hemos encontrado voces comunes a un gran número de países hispanohablantes que son fundamentales en la enseñanza de español —lo que no impide que estas palabras puedan convivir con otras voces locales más o menos frecuentes.

En todo caso se trata de voces idóneas para el aprendizaje de español atendiendo al criterio de rentabilidad y extensión de las variantes. Esto no quiere

decir que no se enseñen regionalismos o palabras locales de acuerdo con el contexto de enseñanza, pero existen palabras, como las que mostramos a continuación, que son imprescindibles en la enseñanza de ELE. De hecho, sería deseable que estas palabras se incluyesen en los manuales didácticos, para garantizar el tratamiento integrado y sistemático de las voces más representativas del español.

Para confirmar el significado y la extensión de estas voces americanas, utilizaremos el *Diccionario de americanismos* (2010), que constituye una obra de referencia esencial para estudiar la dispersión de las voces americanas. Utilizamos esta herramienta para validar los datos de Varilex (o para advertir de discrepancias cuando sea el caso). Este diccionario considera que los americanismos pueden ser:

- Lexemas autóctonos de América y, en caso de haberlos, sus derivados.
- Creaciones originales americanas.
- Criollismos morfológicos.
- Lexemas de procedencia española con cambio o especificación de contenido semántico.
- Arcaísmos españoles vivos en América.
- Lexemas procedentes de otras lenguas tanto antiguos en el español americano (afronegrismos, por ejemplo) como recientes.

4.2.1. Propuesta específica de americanismos

En nuestra ampliación consideramos que una noción léxica es un americanismo representativo del español cuando:

(a) Aparece en el *Diccionario de americanismos* (2010).

(b) Ocupa la primera o la segunda posición en términos de representatividad respecto a todas las respuestas registradas en la base de datos de Varilex.

Presentamos a continuación la Tabla 13 con 40 voces americanas que complementan las nociones del *PCIC*. Todas ellas están avaladas por el *Diccionario de americanismos* (2010) y aportan léxico de gran extensión y representatividad para el estudiante de español. Esta ampliación léxica plurinormativa contribuye a aumentar la solidez referencial del *PCIC* y permite adaptar las voces léxicas a los diversos contextos de aprendizaje.

La Tabla 13 incluye:

- El término de referencia del *PCIC*, con su GR.
- La voz alternativa americana más representativa, con su GR.
- El concepto en inglés que atribuye Varilex para referirse a ambas voces.
- La definición aportada a los informantes en las encuestas de Varilex.

En algunos americanismos aparece la marca [!] con la que señalamos que la palabra común en América también existe en el español de España, pero con un significado diferente que podría dar lugar a equívocos, por ejemplo, cuando las dos voces hacen referencia a realidades cercanas semánticamente pero con significados distintos a ambos lados del Atlántico.

Por otra parte, 32 de las 40 voces americanas tienen, como mínimo, 10 puntos más de GR que los términos del *PCIC*. Este dato garantiza que en nuestra ampliación solo incluimos palabras verdaderamente representativas, evitando los términos locales. Los americanismos que tienen una representatividad superior a las nociones del *PCIC* pero con menos de 10 puntos de diferencia respecto a su GR, se marcan con color gris. Su incorporación en nuestra propuesta se justifica en §4.2.2, en base a la extensión geográfica que representan.

En relación con la ortografía, en el caso de los extranjerismos se han seguido las recomendaciones ortográficas del *Diccionario Panhispánico de Dudas* (2005).

Tabla 13
Propuesta de ampliación de americanismos en el PCIC

Grupo II.1 . Americanismos con mayor Grado de Representatividad						
Voz PCIC	GR	Voz americana	GR	Concepto de Varilex	Definición de Varilex	
A1-A2						
cazadora	11,26	chamarra	25,87	WINDBREAKER	Chaqueta fuerte y cómoda que se lleva para hacer actividades fuera de casa.	
brega	9,89	calzón [!]	16,46	PANTIES	Prenda interior femenina que recubre la parte inferior de la cintura.	
bolso	22	cartera [!]	47,56	HANDBAG	Bolsa de cuero, tela o similar que llevan las mujeres con sus cosas personales cuando salen de casa.	
lavabo	34,63	lavamanos	48,72	WASHBASIN	Lugar donde se lavan las manos.	
coche	22,08	carro	39,31	CAR	Vehículo destinado al transporte de personas.	
ordenador	12,12	computadora	64,36	COMPUTER	Máquina para procesar los datos de información.	
sello	17,61	estampilla [!]	53,81	STAMP	Estampita emitida por una administración postal al servicio de correos.	
cocina	26,17	estufa [!]	36,2	COOKER (STOVE)	Aparato donde se cocina.	
billete	10,17	boleto	55,61	NOTE (US:BILL)	Papel de tamaño pequeño que sirve de pasaje de un tren.	
pizarra	30,25	pizarrón	55,12	BLACKBOARD	Planta que se emplea para escribir.	
salón	2,03	living	13,58	LIVING ROOM	Sala para estar cómodo, para ver la televisión, etc.	
armario	18,27	clóset	44,54	WARDROBE	Armario donde se guardan ropas.	

medias	4,5	pantimedias	31,61	PANTY HOSE	Combinación de panty y medias.
carne de conducir	13,17	licencia de conducir	45,92	DRIVING LICENCE	Documento oficial que autoriza a su poseedor a conducir un vehículo de motor. Ej. <i>Saqué el(**)** a los dieciocho años pero todavía no tengo coche.</i>
anuncio	5,88	cartel	37,39	POSTER	Anuncio escrito de un evento, colocado en lugares públicos.
camarero	12,6	mesero	41,53	WAITER	Persona que se dedica a servir comidas o bebidas en un establecimiento.
patata	34,95	papa	56,35	POTATO	Tubérculo comestible, de forma redonda o alargada y de color marrón por fuera y blanco o amarillo por dentro, que se usa como alimento.
susponder	5,71	reprobar	23,07	GIVE A FAILING GRADE	Dar la pésima nota.
caña	2,53	chela	23,99	BEER	¿Qué otro nombre tiene la cerveza?
B1-B2					
escaparate	20,29	vitrina	41,64	SHOP WINDOW	Papel de tamaño pequeño que sirve de pasaje de un tren.
conducir	16,55	manejar	46,30	DRIVE	Llevar el control de un automóvil
alianza	4,5	argolla	11,55	RING	Aro que se pone en los dedos de la mano como adorno.
mono	7,74	overol	60,01	OVERALLS	Traje usado por pintores, mecánicos, etc., que consiste en un pantalón con peto y mangas en una sola pieza.
grifo	18,57	llave (de agua)	56,09	TAP (US:FAUCET)	Dispositivo para abrir o cerrar el paso del agua.
lavavajillas	9,87	lavavajilla	14,1	DISHWASHER	Máquina para lavar platos.
pizarra	30,25	pizarrón	55,12	BLACKBOARD	Planta que se emplea para escribir.

celo	9,09	(cinta) <i>scotch</i>	41,65	CELLOPHANE TAPE	Cinta transparente, con una sustancia adhesiva en una de sus caras, que se emplea, por ejemplo, para fijar papel a otra superficie. Ej. <i>¿Tienes un poco de (***) para pegar este papel en el tablón?</i>
pincho	0,72	botana	11,85	TIDBITS	Alimento ligero que se sirve como acompañamiento de una bebida.
manta	14,26	cobija	36,12	BLANKET	Trozo rectangular de un tejido grueso que se usa para abrigar en la cama.
quedarse en blanco	0,03	írsele la onda	9,25	CANNOT REMEMBER	No poder recordar algo.
C1-C2					
pluma estilográfica	6,22	pluma fuente	29,45	PEN	Utensilio para escribir con tinta.
repetidor	18,14	repitente	19,28	REPEATER	El que ha obtenido muy malas notas y no ha podido pasar al grado superior.
colérico	1,19	bravo	8,58	ANGRY	Que tiene un disgusto muy fuerte.
catear	3,64	aplazar	10,66	GIVE A FAILING GRADE	Dar la pésima nota.
compadre	3	viejo; cuate	8,93; 8,67	OLD BOY	Forma de llamar a un amigo.
amigote	0,94	yunta; cuate	4,4; 4,27	CLOSE FRIENDO	El mejor amigo.
desagradable	1,21	fregado	17,05	ANNOYING	Que causa disgusto.
tener enchufe	6,37	tener palanca	53,72	HAVE PULL.	Gozar de la protección de alguien: El jefe no se atreve a llamarle la atención, porque (***)
histérico	3,34	bravo	6,17	ANGRY	Que tiene un disgusto muy fuerte.

Podemos observar que algunos de los términos propuestos, además de tener mayor GR que los términos de referencia, son más transparentes que las nociones del PCIC, por ejemplo, *lavabo* (Esp) – *lavamanos* (Am)⁵.

En otros casos, la inclusión de la voz americana también es importante porque el término de referencia de España tiene otro significado diferente en la mayoría de los países americanos, como sucede con gran número de voces, por ejemplo, con *media*, *billete* o *suspender*.

Como ya hemos advertido, la marca [!] advierte de posibles interferencias semánticas en los siguientes casos:

braga (Esp) – *calzón* (Am) [!]

La ‘prenda interior femenina que recubre la parte inferior de la cintura’ [PANTIES] se denomina *braga* (GR = 9,89), en el español de España, y *calzón* (GR = 16,46), en gran parte de la geografía hispánica (Gu, Ho, Ni, CR, Co, Ec, Pe, Bo, Ch; Ar, Ur, obsol.) En España se utiliza *calzón* para referirse a la prenda interior masculina, que se lleva debajo de los pantalones. Es cierto que *calzón* antiguamente designaba la prenda interior femenina (GDUEA, 2006), pero con el tiempo el término *braga* empezó a sustituir a *calzón* y este último hoy ya no se utiliza.

bolso (Esp) – *cartera* (Am) [!]

En los países americanos la ‘bolsa de cuero, tela o similar que llevan las mujeres con sus cosas personales cuando salen de casa’ se denomina *cartera* (Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur.). En cambio, en España se prefiere el término *bolso*. El término *cartera* en el español europeo designa una realidad cercana semánticamente pero que se refiere a otro concepto: ‘Funda de material flexible (piel, tela, etc.) con compartimientos en su interior, usada para guardar billetes y otros objetos planos y finos (papeles, tarjetas de crédito, carné de identidad, etc.)’ (GDUEA, 2006).

⁵ Ya comentamos la importancia del criterio de transparencia en §4.1.1.

sello (Esp) – *estampilla* (Am) [!]

Una *estampilla*, en la mayoría de países americanos (Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, Co, Ve, Ec, Pe, Ch, Py, Ar, Ur; Bo, pop + cult → espon.), es el sello fiscal o de correos. En España la *estampilla* es el sello o la plancha de caucho con la firma o rúbrica de alguien en relieve (GDUEA, 2006).

cocina (Esp) – *estufa* (Am) [!]

La *estufa*, en varios países americanos (Mx, Gu, Ho, Pa, RD, Co), es el aparato que sirve para cocinar, con hornillo o fuegos y a veces horno. Para este concepto, en España se prefiere la voz *cocina*; la voz *estufa* se refiere al utensilio doméstico que calienta mediante la combustión de gas, leña, carbón, etc. convirtiendo la electricidad en calor.

4.2.2. Otros americanismos sujetos a valoración

Aunque en la mayoría de los casos las voces americanas propuestas tienen 10 puntos más de GR que los términos del *PCIC*, existen algunas voces americanas con una diferencia menos significativa. En estos casos es preciso valorar su inclusión según el criterio de extensión geográfica, en función de la dispersión de las voces en las áreas dialectales americanas y el número total de países en los que se registra la voz. Para incorporar o descartar su inclusión seguimos los siguientes criterios:

(a) Todas las palabras alternativas americanas deben utilizarse al menos en 3 zonas dialectales del español.

(b) Todas las palabras alternativas americanas deben utilizarse al menos en 6 países americanos.

El sentido de estas condiciones es evitar la inclusión de voces americanas excesivamente locales, puesto que nuestra propuesta presenta elementos léxicos compartidos por la mayoría de hispanohablantes.

Dado que en Varilex se incluyen registros léxicos relativos a distintos perfiles de informantes, es muy útil trabajar con una herramienta como el *Diccionario de*

americanismos, que aporte validez y fiabilidad a los datos de representatividad. Este diccionario, avalado por la Asociación de Academias de la Lengua Española, es el filtro adecuado para confirmar el uso de las nociones en los distintos países.

A pesar de las posibles controversias respecto a la veracidad de los datos dialectales y teniendo en cuenta la dificultad de determinar taxativamente la existencia o ausencia de voces léxicas en los distintos países, el *Diccionario de americanismos* constituye un material de estudio acreditado que nos sirve para determinar la extensión dialectal de las nociones léxicas.

En la Tabla 14 incluimos americanismos con un GR ligeramente superior a los términos del *PCIC*, cuya ampliación debe ser cuidadosamente valorada. Mostramos primeramente las palabras propuestas para la ampliación —que ya aparecían en la Tabla 13 en color gris— y, a continuación, las palabras descartadas para el objetivo de esta tesis, pero que no queremos dejar de mencionar. La extensión geográfica de las voces, en función de los países y zonas que las representan, es la que justifica la incorporación o el descarte proponiendo, como decíamos, las voces que aparezcan en un mínimo de tres zonas dialectales y seis países americanos.

Tabla 14*Valoración de otros americanismos con GR ligeramente superior*

Valoración	Voz	Extensión Dicc Am	Zonas	Definición del DA
Palabras <u>propuestas</u> para la ampliación	calzón	Gu, Ho, Ni, CR, Co, Ec, Pe, Bo, Ch; Ar, Ur, obsol.	5	Braga, prenda interior femenina.
	aplazar	Ho, ES, Ni, Ec, Pe, Bo, Py, Ar, Ur; Pa, Ve, p.u.	3	No aprobar una persona a alguien en un examen o en una prueba, principalmente académica.
	viejo	Mx, Ho, ES, Ni, Cu, B o, Ch.	4	Amigo íntimo, compañero inseparable.
	yunta	Cu, Pe, Bo, Ch, Py; Ec . p.u. juv.	3	Amigo íntimo, inseparable.
	cuate	Mx, Gu, Ho, Ni, PR, Ve, Bo, Py.	5	Camarada, amigo íntimo. pop + cult → espon.
	bravo	Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Bo, Py.	5	<i>Referido a persona,</i> irritada, enojada u ofendida.
	argolla	Mx, Gu, Ho, ES, Ni, Pa, Cu, RD, PR, Co, Ec, Bo, Ch, Ar; Ur.	5	Anillo de compromiso o de matrimonio.
	írsele onda	Mx, Ho, ES, Ni, CR, Cu, Co, Bo, Ch.	4	Perder alguien momentáneamente la ilación, u olvidarse de algo. pop.

Palabras descartadas para la ampliación	torta	Mx	1	Pan blanco partido longitudinalmente, que se rellena con diversos alimentos.
	agüitado	Mx, ES	1	<i>Referido a persona, abatida, triste, melancólica. pop.</i>
	lavavajilla	Mx, Ni, Ch, Py, Ar.	3	Lavavajillas, máquina para lavar platos, cubiertos y otros enseres de cocina.
	botana	Mx, Gu, Pa; Ho, Ni, p.u.	1	Bocadito.
	lonchería	Mx, Ho, Co, Ve.	2	Lugar donde se venden comidas ligeras.

Por los motivos expuestos, *torta*, *agüitado*, *lavavajilla*, *botana*, *lonchería* no se incluyen en nuestra propuesta de ampliación, aunque tengan un GR ligeramente superior a los términos del *PCIC*. Las voces americanas propuestas que se consideran suficientemente representativas —y que ya adelantamos en la Tabla 13— son: *calzón*, *aplazar*, *yunta*, *cuate*, *bravo*, *argolla*, *írsele la onda*.

4.3. Voces americanas secundarias, con menor GR pero con alcance y extensión significativa (Grupo II.2)

Además de los americanismos propuestos en §4.2., también nos parece importante incluir algunas voces americanas que permitan incorporar en las «Nociones específicas» variedades dialectales representativas del español, aunque no tengan mayor GR que los términos de referencia del *PCIC*. Teniendo en cuenta que este plan curricular está orientado a la docencia de ELE, tampoco tiene sentido presentar todas las variantes léxicas de los distintos subsistemas lingüísticos. Se trata de incluir únicamente una serie de *americanismos secundarios* con gran rentabilidad y extensión que permitan, tanto al docente como al estudiante, conocer variantes diatópicas significativas e incorporarlas en función del nivel y el contexto de aprendizaje.

Incluir estas voces en el *PCIC* no significa que los docentes tengan que utilizar todas estas variantes en el aula, ni mucho menos los estudiantes; pero teniendo en cuenta que no existe ningún manual exhaustivo que secuencie en niveles de aprendizaje la variación léxica en la enseñanza de español, resulta de gran interés hacer del inventario de «Nociones específicas» del *PCIC* un repertorio léxico riguroso para la orientación y consulta en la enseñanza del léxico, incorporando nociones representativas del español que complementen la variedad peninsular escogida como preferente. Así, cuando el contexto de enseñanza y el nivel de aprendizaje lo requieran, el docente, el planificador de exámenes o el creador de materiales didácticos podrán consultar la extensión de ciertas voces americanas importantes para el aprendizaje.

Al incluir estas voces pretendemos, ante todo, presentar un repertorio léxico coherente, constituido por los vocablos de mayor difusión panhispánica (Moreno Fernández, 2000), que incluyen el léxico del área dialectal de España pero también el léxico fundamental de otras regiones hispánicas.

4.3.1. Propuesta específica de voces americanas secundarias

Para proponer *voces americanas secundarias* en el *PCIC* seguimos dos criterios:

(a) La voz aparece registrada en el *Diccionario de americanismos*

(b) Si tiene un GR inferior a 10 puntos, la voz deberá contar con una extensión mínima de 5 países que se distribuyan, al menos, en 3 zonas dialectales.

Aunque las nociones propuestas no tengan mayor GR que los términos de referencia —por eso se denominan *secundarias*—, todas ellas gozan de una extensión geográfica significativa y se registran en el *Diccionario de americanismos*, complementando a las nociones del *PCIC* desde la perspectiva panhispánica que venimos defendiendo a lo largo de esta tesis. La Tabla 15 presenta nuestra propuesta de voces americanas secundarias e incluye:

- El término de referencia del *PCIC*, con su GR.
- La voz alternativa americana secundaria, con su GR.
- Los países en los que se registran según el *Diccionario de americanismos*.
- Las zonas dialectales en las que se distribuyen las voces.
- La definición está incluida en el *Diccionario de americanismos*.

En algunas voces americanas observamos el símbolo [!], que ya hemos utilizado previamente para marcar la existencia de diferencias semánticas que podrían dar lugar a equívocos, por tratarse de realidades cercanas pero con significados distintos. Asimismo, algunas palabras de la tabla incorporan la marca [*], con la que señalamos que esas variantes americanas deberían incluirse en niveles superiores de aprendizaje del *PCIC*, bien por sus marcas de uso, por su bajo índice de representatividad o por la existencia de otras variedades alternativas más representativas en ese mismo nivel.

Puede ocurrir que los americanismos secundarios se propongan para términos del *PCIC* que ya cuentan con voces alternativas (americanas o panhispánicas) propuestas en las secciones anteriores. En estos casos, la Tabla 15 refleja, primeramente, el término del *PCIC*; a continuación, el primer término alternativo (panhispánico o americano); por último, la voz americana secundaria. Respecto a la ortografía, seguimos las indicaciones del *Diccionario de Lengua Española* (2014).

Tabla 15
Variantes americanas secundarias

Grupo III: Variantes americanas secundarias con rentabilidad y extensión significativas						
Términos de Referencia	GR	Voz americana secundaria	GR	Países	Zonas	Definición del Diccionario de americanismos
A1–A2						
abrigo (PCIC)	56,45	sobretudo (DA)	4,33	Cu, RD, Co, Ec, Pe, Bo, Py, Ar, Ur; Pa, Ch, obsol.	5	sobretudo. Prenda de vestir de manga larga y que llega hasta más abajo de la rodilla, usada como abrigo.
falda (PCIC)	88,26	pollera (DA)	8,16	Co: N, Ch, Py, Ar, Ur.	5	pollera. Falda o parte del vestido de mujer que cae desde la cintura.
calcetines (PCIC)	47,23	medias (DA)	38,79	ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur.	6	media. Calcetín, prenda que cubre el pie y llega hasta la pantorrilla.
sujetador (PCIC)	7,04			Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Py, Ur.	6	
sostén (Panhisp)	36,37	brasier	36,23			brasier. Prenda interior femenina para sujetar el pecho y darle forma.
braga (PCIC)	9,89			Mx, Ho, Ni, Pa, PR, Ve; f. pl. RD, Co: N.	3	
calzón (Panh)	16,46	pantaleta (DA)	12,68			pantaleta. Braga, prenda interior femenina.
gafas (PCIC)	26,14			Mx, Ho, Ni, CR, Cu, PR, Co, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur.	6	
lentes (Panh)	40,29	anteojos (DA)	29,84			anteojos. Lentes para corregir defectos de visión, o para proteger los ojos.

CAPÍTULO 6

bañador (<i>PCIC</i>) traje de baño (Panhisp)	8,35 60,59	10	10	Bo, Py, Ar, Ur; Ch, obso.	3	mall . Bañador.
maleta (<i>PCIC</i>)	63,81	20,71	20,71	Mx, Gu, Ho, ES, Ni, CR, Bo, Py, Ar.	3	valija . Maleta o bolso grande para llevar el equipaje en un viaje.
nevera (<i>PCIC</i>) refrigerador (Panhisp)	25,76 42,75	10,6	10,6	Gu, ES, Ni, CR, Pa, Ve, Ec.	3	refrigeradora . Nevera, electrodoméstico.
azafata (<i>PCIC</i>)	64,03	31,67	31,67	Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur, p.u.	6	aeromoza . Auxiliar de vuelo, miembro de la tripulación, encargada de atender a los pasajeros durante el vuelo.
ascensor (<i>PCIC</i>)	64,03	27,25	27,25	Mx, Gu, Ho, ES, Pa, Cu, RD, PR, Ec, Bo, Ch, Py, p.u.	5	elevador . Ascensor.
bolígrafo (<i>PCIC</i>)	46,22	14,01	14,01	Gu, ES, Ni, CR, RD, Co:O,SO, Ve, Pe, Bo, Ch.	5	lapicero . Bolígrafo.
sello (<i>PCIC</i>) estampilla (DA)	17,61 54,81	22,07	22,07	Mx, Ni, Pa, RD, Ec, Bo, Ur.	4	timbre . Sello postal.
piscina (<i>PCIC</i>)	69,74	9,83	9,83	RD, Bo, Py, Ar, Ur.	3	pileta . Piscina.

		picada* (DA)	6,57	Pa, Co, Ec, Bo, Py, Ar, Ur.	4	picada. Conjunto de alimentos, como trozos pequeños de carne, papa, maíz, patacones, que se sirven en reuniones informales como aperitivo o acompañamiento de bebidas alcohólicas.
tapa (PCIC)	15,61	bocadito* (DA)	4,12	Ho, Ni, Pa, Cu, RD, PR, Ec, Pe, Bo, Py, Ar, Ur.	4	bocadito. Alimento elaborado con ingredientes variados, que pueda ingerirse en un solo bocado y que se sirve generalmente en aperitivos para acompañar bebidas.
anuncio (PCIC)	5,88			Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ar, Ur; Mx, p.u.	5	afiche. Cartel que obedece a fines informativos, publicitarios o decorativos.
cartel (Panhis)	37,39	afiche (DA)	27,76			
cerdo (PCIC)	45,36	chanchó (DA)	9,11	Mx, Gu, Ho, ES, Ni, CR, Pa, Ve, Ec, Pe, Bo, Ch, Py, Ar, Ur; RD, PR, rur.	5	chanchó. Cerdo, animal mamífero.
mosquito (PCIC)	61,41	zancudo (DA)	29,15	Mx, Gu, Ho, ES, Ni, CR, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Ar; Ur, p.u.	6	zancudo. Mosquito, insecto díptero de hasta 10 mm de longitud, de color negro, a veces con manchas blancas; la hembra se alimenta de sangre y el macho de néctar.
		churro* (DA)	8,26	Co, Pe, Bo, Ar, Ur; Ch, obsol.	4	churro. Persona bien parecida y atractiva. pop.
		regio* (DA)	5,25	Mx: N, Ni, PR, Ec, Pe, Bo, Ch, Py, Ar, Ur	5	regio. Muy hermoso o lindo. pop + cult → espon.
guapo (PCIC)	21,97	mango* (DA)	4,64	Mx, Gu, Ho, ES, Ni, Pa, Cu, Bo.	3	mango. Persona atractiva físicamente. pop.

gripe (PCIC)	19,83	gripa (DA)	12,24	Mx, Pa, Co, Ec.	3	gripa. Gripe.
negocio (PCIC)	7,52			ES, Ni, Pa, RD, Co, Ec, Pe, Bo, Ch, Py, Ar, Ur; Ve. obsol.	6	almacén. Tienda donde se vende al por menor una gran variedad de mercancías.
tienda (Panhispp)	55,5	almacén (DA)	21,46			
español (PCIC)	69,81	gallego* (DA)	12,37	Cu, RD; Bo, Py; Ar, Ur, pop.	3	gallego. 1. Relativo a España. 2. <i>Referido a persona</i> , de ascendencia española.
B1-B2						
carpeta (PCIC)	56,59	fólder ⁶ (DA)	31,45	Mx, Gu, Ho, ES, Ni, CR; Pa, Cu, RD, Ec, Pe, Bo, Ch, Py, Ar.	5	fólder. Carpeta de cartulina, plástico u otro material, para archivar papeles o documentos.
interruptor (PCIC)	38,63	switch (DA)	20,84	Mx, Gu, Ho, Ni, CR, RD, PR, Ec, Pe, Bo, Ch.	4	switch. Interruptor eléctrico.
bizcocho (PCIC)	22,74	bizcochuelo (DA)	14,27	Co, Ve, Bo, Ec, Py, Ar, Ur.	3	bizcochuelo. 1. Pastel que tiene como base harina, huevos y azúcar batidos. 2. Masa esponjosa que se prepara con harina, huevos y azúcar batidos.
hueso (PCIC)	29,26	pepa* (DA)	18,35	Ho, Pa, CR, PR, Pe, Bo.	3	pepa. 1. Semilla grande y dura de algunas frutas. 2. Semilla grande de algunos frutos como el chayote.
nota (PCIC)	13,64					
puntuación (Panhispp)	33,5	puntaje (DA)	15	Mx, Ho, ES, Ni, CR, Pa, RD, Co, Ve, Ec, Pe, Bo, Ch, Ar, Ur;	6	puntaje. Conjunto de puntos obtenidos en algún tipo de prueba.

⁶ Seguimos la ortografía del *Diccionario panhispánico de dudas* (2005).

desempleo (PCIC) paro (Panhisp)	56,3 22,91	desocupación (DA)	9,46	Mx, CR, Cu, Co, Ec, Pe, Bo, Ch, Py, Ar, Ur.	6	desocupación. Desempleo.
de pie (PCIC)	56,31	parado (DA)	34,99	Mx, Gu, ES, Ni, CR, Pa, Cu, RD, PR, Ve, Ec, Pe, Ch, Py, Ar, Ur; Bo, pop.	6	de pie. Referido a persona, que está de pie.
despedir (PCIC) echar (Panh)	35,42	botar* (DA)	3,31	Mx, Gu, Ho, ES, Ni, Pa, Cu, RD, PR, Co, Ve, Ec, Pe, Bo, Ch, Ar.	5	botar. Despedir a alguien de un empleo.
cerveza de barril (PCIC)	n. r. ⁷	chela (DA)	23,99	Mx, Gu, Ho, ES, Ni, Co, Ec, Pe, Bo, juv.	3	chela. Cerveza.
C1-C2						
ataúd (PCIC)	69,5	cajón (DA)	10,95	Co, Pe, Bo, Ch, Py, Ar, Ur; Mx, Gu, Ho, ES, Ni, Pa, rur; pop; Ve, Ec	4	cajón. Ataúd.
látigo (PCIC)	48,58	fuete (DA)	18,67	Mx, Gu, Ho, ES, Ni, Pa , Cu, RD, PR, Co, Ve, Ec, Pe; Bo, p.u; CR,	4	fuete. Azote largo, delgado y flexible, de cuero, cuerda u otra materia, con que se aviva y castiga especialmente la caballería.
becario (PCIC)	53,64	becado (DA)	37,66	Gu, Ho, Ni, CR, Cu, R D, PR, Co, Ve, Ch, Ur.	5	becado. Persona que estudia con una beca, becario.
gordínflón (PCIC)	12,31	chanchó (DA)	6,18	Ho, ES, Ni, CR, Pe, Bo, Ch, Py, Ar.	4	chanchó. Persona muy gorda. pop + cult → espon ^ desp.

⁷ La voz no está registrada en los datos de Varilex.

En la Tabla 15 incluimos variantes americanas que, aun siendo representativas del español, tienen menor GR que los términos de referencia del *PCIC*. A veces ocurre que las voces americanas que añadimos tienen un GR muy inferior puesto que se oponen a hiperónimos panhispánicos. Esto sucede, por ejemplo, con *abrigo*, *azafata* o *ascensor*, que alternan con otras voces secundarias según las regiones: *sobretudo*, *elevador* o *aeromoza*.

Como ya hemos mencionado, cuando aparece el símbolo [!] advertimos de diferencias semánticas importantes. Es el caso de la oposición *bolígrafo* – *lapicero*, puesto que *lapicero* en gran parte del territorio hispanohablante se refiere al utensilio de madera que sirve para escribir con una barra de grafito en su interior, aunque en Guatemala, El Salvador, Nicaragua, Costa Rica, República Dominicana, Colombia, Venezuela, Perú, Bolivia y Chile se refiere al utensilio que se utiliza para escribir con tinta.

Por otra parte, con la marca [*] se advierte de la necesidad de resituar estas palabras en otros niveles del *PCIC*. Se trata de los casos siguientes:

tapa (Esp) – *picada**, *bocadito** (Am)

Este es un caso interesante para reflejar la complejidad que conlleva la aplicación de nuestros resultados demolingüísticos a las «Nociones específicas» del *PCIC*. El concepto [TIDBITS] que se define como ‘alimento ligero que se sirve como acompañamiento de una bebida’ presenta, junto a la voz *tapa*, tres voces alternativas: *botana* (11,85) (que no aparece en la tabla por tratarse de un regionalismo), *picada* (6,57) y *bocadito* (4,12); entre otras más locales que no son relevantes para nuestro estudio.

En principio, si atendemos a nuestro cálculo de GR, consideramos que la palabra *botana* es la primera alternativa posible. Las voces *picada* y *bocadito* tienen índices de representatividad más bajos pero, en cambio, según los datos del *Diccionario de americanismos* son voces con un alto nivel de dispersión que se distribuyen a lo largo de cuatro áreas dialectales del español. En este sentido, a pesar del alto índice de representatividad que se concede a *botana*, siguiendo el criterio de

rentabilidad y extensión en los repertorios léxicos para el aprendizaje de español, las voces *picada* y *bocadito* serían las primeras candidatas para incluirse en el PCIC.

El alto GR de *botana* viene motivado —además de por ser la voz preferida de México— por ser una variante muy consolidada en los países en que se ha contestado, es decir, que en los países en los que se registra, la mayoría de personas la contestaron como voz preferida, lo que le aporta una fuerte representatividad. El GR más bajo que se atribuye a *picada* y *bocadito* podría explicarse por ser voces secundarias, conocidas en muchos países, aunque siempre a la sombra de otras formas con mayor representatividad.

En todo caso, existen discordancias respecto a la distribución del concepto [TIDBITS] entre los datos del *Diccionario de americanismos* y los registros de Varilex. En la tabla siguiente mostramos, a modo de ejemplo, la distribución de *bocadito*. Marcamos con ‘1’ los países en los que se ha registrado la voz y con ‘0’ los países en los que no se ha registrado. La letra ‘H’ en los registros de Varilex se refiere a los *hápx* o respuestas aisladas ofrecidas por un único informante, cuya utilización se considera dudosa.

Tabla 16

Distribución de la voz bocadito en el Diccionario de Americanismos y en Varilex

bocadito	Ho	Ni	El	Pa	Cu	RD	PR	Ve	Ec	Pe	Bo	Ch	Ar	Ur	Py
DA	1	1	0	1	1	1	1	0	1	1	1	0	1	1	1
VRLX	0	0	H	0	1	0	0	H	0	1	1	H	1	0	H

Dado que el *Diccionario de americanismos* pretende mostrar voces habituales en de los países de América, sorprende, por ejemplo, que *bocadito* se registre en Honduras, Nicaragua, Panamá, República Dominicana, Puerto Rico, Ecuador y Uruguay pero que ningún informante de esos países la haya marcado en las encuestas de Varilex. El caso contrario, el de las voces recogidas por Varilex que no se incluyen en el *Diccionario de americanismos*, es más fácil de entender puesto que las personas que marcaron *bocadito* en El Salvador, Venezuela y Chile son *hápx* o respuestas aisladas que no representan el habla de la población.

guapo (Esp) – *churro**, *regio**, *mango** (Am)

Como suele ocurrir con los adjetivos valorativos, la riqueza léxica que se genera en los subsistemas lingüísticos del español es enorme. Para hacernos a la idea de cuantísimas expresiones recoge Varilex para el concepto [ATTRACTIVE MAN] citamos, a modo de ejemplo, algunos adjetivos: *atractivo*, *apuesto*, *guapísimo*, *guapetón*, *elegante*, *papacito*, *bárbaro*, *impresionante*, *bellísimo*, *cuerísimo*, *galán*, *hermoso*, *guapérrimo*, *papirriqui*; o las expresiones idiomáticas: *estar como un tren / de vicio / de pecado / cañón / de morirse / rebien / que quita el hipo...* entre muchas otras.

De entre todas ellas, destacan tres variantes americanas con alto GR: *churro*, *regio* y *mango*. Aunque no se incluyan en el nivel A del *PCIC* —por ser menos generales y, en consecuencia, poco rentables para la enseñanza de ELE— proponemos incorporarlas en el apartado «1.2. Características físicas» del nivel C.

español (Esp) – *gallego** (Am)

Para referirse a la persona de origen español, además de otras voces como *gachupín*, *gaita* o *godo*, el término más extendido en la geografía americana es *gallego* con un alto GR y utilizado, según el *Diccionario de americanismos*, en Cuba, República Dominicana, Bolivia, Paraguay, Argentina y Uruguay. Sorprende que este diccionario no incluya Colombia puesto que tanto Varilex como el DLE señalan su uso en esta región.

Además de los países mencionados, los informantes de Varilex también marcaron la voz *gallego* en Puerto Rico, Venezuela y México. El *Diccionario de americanismos* explica que la voz se utiliza en esos países pero con un sentido despectivo. En casos como este confirmamos que aunque Varilex recoge enormes cantidades de registros léxicos en todo el mundo, es fundamental ordenar y

estructurar los datos incorporando marcas de uso específicas, como hace el *Diccionario de americanismos*⁸.

hueso (Esp) – *pepa** (Am)

La ‘parte interior de algunos frutos como el melocotón o la ciruela’ a la que se refiere el concepto [STONE (US: PIT)] de Varilex se denomina *pepa* en algunos países de América. Este concepto, más restringido que el concepto *hueso* del nivel B, podría introducirse en el nivel C1 junto a la serie ‘*pepita, gajo, pulpa*’ del apartado «5.3. Alimentos».

despedir/ echar del trabajo (Esp) – *botar** (Am)

Para el concepto [EXPEL] que hace referencia a ‘echar a una persona del trabajo’, la voz más representativa es *despedir*, que se incluye en el nivel B1 del *PCIC*. Más adelante, en el nivel C1 se incluye la forma *echar del trabajo*, que tiene menor GR y es propia de contextos más restringidos. Junto a ella, proponemos incluir la voz *botar*, marinerismo de enorme extensión en los países de América, tal y como confirma el *Diccionario de americanismos*.

4.3.2. Americanismos descartados para la ampliación

Al contrastar el inventario de «Nociones específicas» con los datos léxicos de Varilex se han registrado otras variantes americanas secundarias que no hemos incluido en nuestra ampliación por no cumplir los requisitos expuestos en §4.3.1. Se trata de palabras con baja extensión, para las que ya se han propuesto otras variantes más significativas. Existen otros casos en los que el descarte está justificado porque las voces entran en oposición denotativa, designando realidades distintas.

⁸ Nuestras consultas en el CORPES tampoco han conseguido resolver las discrepancias respecto a la extensión del término pero confiamos en que, según aumente y perfeccione el tamaño y los sistemas de etiquetación léxica, consultas de este tipo podrán quedar ser resueltas. Hemos comprobado la existencia de un registro de la voz *gallego* en Puerto Rico referido a persona de origen español pero no hemos registrado la existencia de la voz en otros países como Venezuela o México.

Tabla 17
Americanismos descartados para la ampliación

Voz	Extensión	GR	Motivo de descarte
recámara (<i>dormitorio</i>)	Mx; Ho, Ni, Co, Ec, Bo, Ch, p.u, esm; Pa, pop + cult → esm.	13,79	La voz es casi exclusiva de México. Aunque se conoce en otros países, lleva la marca de <i>poco usado</i> . Además, ya se incluyen en el <i>PCIC</i> las voces <i>dormitorio</i> y <i>habitación</i> , mucho más representativas.
alberca (<i>piscina</i>)	Mx; Gu, Ho, Ni, Pa, rur; Bo, p.u.	20, 44	Voz exclusiva de México. Conformar una oposición denotativa respecto a <i>piscina</i> puesto que se refiere a un ‘estanque para la natación y otros deportes’. Además, <i>alberca</i> también tiene otro significado en España.
pantijós (<i>medias</i>)	ES	3,59	Voz local.
trusa (<i>calzoncillo</i>)	Mx, Pe	14, 41	Voz local.
interior (<i>calzoncillo</i>)	RD, Ve	7,4	Voz local.
bañadera (<i>bañera</i>)	Cu, Rd, Pe, Py, Ar, Ur; Bo: E.	10,25	No es necesaria añadirla en el <i>PCIC</i> puesto que ya incluimos dos alternativas: <i>bañera</i> y <i>tina</i> .
chalina (<i>pañuelo</i>)	Mx, ES, Co, Ec, Ch, Py, Ar, Ur; Ho, Pa, p.u.	18,88	Oposición denotativa: conceptos diferentes.
lavarropas (<i>lavadora</i>)	Ar, Ur; m–f. Py; f. Bo.	7,64	Voz léxica de muy poca extensión que convive con una única voz alternativa mayoritaria.
chaparro (<i>bajo</i>)	Mx, Gu, Ho, ES, Ni, Pa, PR; Bo, pop + cult → espon; CR, p.u.	15,27	Término de relativa extensión que además tiene otro significado en otras áreas dialectales (<i>rechoncho</i>). Se podría incluir en nivel C1, junto a otras caract. como <i>robusto</i> , <i>corpulento</i> ...
cheves (<i>cerveza</i>)	EU, Mx, Gu, E S. juv.	7,97	Voz local.
cuadrar(se) (<i>ligar</i>)	intr. prnl. Co, Ve. juv. Cu.	12,89	Voz local.

Propuesta de revisión y ampliación de las «Nociones específicas» del PCIC

egresado (<i>licenciado</i>)	Mx, Gu, Pa, Cu, RD, Co, Ve, Bo, Py; sust/adj. Ho, ES, CR, Ec, Pe, Ch, Ar, Ur.	12,08	Oposición denotativa: conceptos diferentes.
altanería (<i>arrogancia</i>)	Ho, Ni, Pa, Co, Ve, Bo, Ur.	10,1	Oposición denotativa: conceptos diferentes.
matado (<i>empollón</i>)	Mx, Gu.	5,33	Voz local.
nerdo (<i>empollón</i>)	CR, PR, Co, Ec; Gu, Ho, ES, Ni, est.	4,4	Oposición denotativa: conceptos diferentes.
mano (<i>compadre</i>)	Gu, Ni, Bo; Pa, rur.	7,83	Voz local.
mameluco (<i>mono, overol</i>)	Pa, RD, PR, Ec, Pe, Ch, Ar, Ur; Bo, Py, pop; Cu, obsol.	7,45	Voz léxica de muy poca extensión que convive con una voz alternativa mayoritaria.
chirimoya (<i>coco</i>)	Gu, Ho, ES, Ni, RD, Ve; Mx, Bo, pop + cult → espon ^ fest; Cu, obsol; CR, p.u.	3,78	Voz léxica de poca extensión empleada que convive con una única voz alternativa claramente mayoritaria.
empacar (<i>atiborrarse</i>)	Mx, Co	6,27	Voz local.
atorar (<i>atiborrarse</i>)	Gu, Ho, ES, Ni	5,73	Voz local.
maceta (<i>cabeza</i>)	Bo; Gu, ES, Ni, Pa, pop; Mx. fest.	6,34	Voz local.

4.4. Casos particulares de ampliación

Además de los casos ya mencionados (panhispanismos, americanismos y americanismos secundarios) presentamos a continuación algunas apreciaciones respecto a las nociones del *PCIC* y la variedad que no se ajustan a los grupos anteriores.

Tabla 18

Casos particulares de ampliación

autobús	Varilex indica la variación léxica del concepto [BUS] con los siguientes registros: <i>autobús, micro, colectivo, bus, microbús, ómnibus, camión, banderita, guagua, flota, taxibús, camioneta, ruta</i> . Si observamos el GR de estas voces, ninguna de las alternativas de la voz <i>autobús</i> (GR = 32,39) parece lo suficientemente representativa como para incluirse en el <i>PCIC</i> . La alternativa más extendida sería <i>ómnibus</i> (GR = 9,80), que también se registra en España y en el <i>Diccionario de Lengua Española</i> y que, según los datos de Varilex, se distribuye así en el territorio hispánico: Ar (35,5 %), Py (37,5 %), Uy (87,5 %), Pe (60 %), Cu (30,75 %) y, en menor porcentaje, RD, Bo y Esp.
bar	Todas las variantes léxicas representativas de la noción <i>bar</i> (<i>taberna, cantina, tasca</i>) se incluyen en otras secciones del <i>PCIC</i> . No es necesario añadirlas.
dinero	Aunque no aparece en el <i>Diccionario de americanismos</i> , tal y como señala el <i>Diccionario de Lengua Española</i> y como se confirma también con los datos de Varilex, la forma <i>plata</i> es un americanismo con gran vitalidad en América que se refiere a ‘dinero o riqueza’. Proponemos incluir esta voz.
pendientes	Aunque no aparece en el <i>Diccionario de americanismos</i> , los datos de Varilex confirman que <i>aretes</i> es un americanismo que corresponde a la noción [EARRINGS]. <i>Aretes</i> ya se ha estudiado como americanismo en general en América, menos en algunos países donde se usan otras voces peculiares: <i>aros</i> en Py, Ch y Ar; y <i>zarcillos</i> en Ve (Ueda, 1996).

5. Marca [Esp]: voces preferentes del español de España o españolismos

En esta sección señalamos las voces del *PCIC* que son preferentes del español de España y que podemos denominar *españolismos*. Este apartado no constituye, por tanto, una ampliación léxica del *PCIC* sino la incorporación de una marca específica para determinadas voces.

Nuestro objetivo es señalar, mediante la marca [Esp], los españolismos del *PCIC*, puesto que las palabras privativas del español de España pasan inadvertidas frecuentemente tanto en diccionarios como en manuales de enseñanza de español. Partimos de la hipótesis de que es esperable que entre las páginas del *PCIC* existan este tipo de nociones, puesto que la institución escoge abiertamente la variedad centro norte peninsular española para establecer los Niveles de Referencia del español. Partiendo de un análisis minucioso de la base de datos de Varilex, hemos localizado, primero, las voces registradas por informantes de España que son desconocidas o poco utilizadas en otras variedades de español. A continuación, hemos constatado que algunas de estas voces están presentes en el *PCIC*.

Los españolismos localizados en las «Nociones específicas» del *PCIC* suelen tener un GR bajo, pues se utilizan exclusivamente en una zona dialectal: la española. Sería deseable que de la misma manera que existen diccionarios de mexicanismos, argentinismos, colombianismos etc., existiera también un repertorio de españolismos. Localizar y señalar estas voces es especialmente importante para la enseñanza de ELE puesto que, como ya hemos comentado, las voces locales solo serán importantes en niveles superiores o en contextos específicos de aprendizaje. Es fundamental que el docente conozca y sea consciente de la extensión limitada de algunas nociones del *PCIC*, especialmente si pensamos en la frecuente elección de la variedad española como variedad preferente de los manuales de español.

En ningún caso consideramos oportuno eliminar estas voces del *PCIC*. No obstante, es fundamental que el usuario de esta obra de referencia conozca y sea

consciente de que algunas de las voces que se incluyen en el inventario de «Nociones específicas» —y en muchas de las obras dedicadas a la enseñanza de ELE— son exclusivas de España, lo que contradice los criterios de rentabilidad y extensión en la selección léxica y no responde al carácter panhispánico y plurinormativo que observamos en otras secciones del documento.

Esta información es ineludible para un tratamiento adecuado de la variación léxica en el currículo y resulta muy útil para la preparación de clases, materiales didácticos o, incluso, para el diseño de pruebas de certificación que tengan validez internacional. Creemos que incluir la marca de españolismo en las «Nociones específicas» amplía notablemente el uso y la validez del *PCIC* como guía de referencia para la enseñanza del léxico.

5.1. Propuesta de la marca [Esp]

Para marcar como españolismo una voz del *PCIC*, acudimos a la base de datos de Varilex y consideramos dos posibilidades:

(a) Que la voz haya sido contestada exclusivamente por hablantes procedentes de España (*españolismos puros*).

(b) Que la voz haya sido contestada preferentemente por los informantes de España y de forma minoritaria en algún punto de la geografía hispánica (*españolismos de preferencia*).

5.1.1. Españolismos puros

En la Tabla 19, presentamos 22 nociones del *PCIC* que podemos considerar españolismos puros según los datos de Varilex. Como se puede apreciar, se trata de palabras con un GR bajo, puesto que han sido contestadas exclusivamente por informantes de España. A veces, se registran respuestas aisladas o hápax en otros países que no se pueden considerar representativos de la población encuestada.

En la Tabla 19 marcamos:

- El término de referencia del *PCIC* que consideramos españolismo con su GR.
- El porcentaje de informantes españoles que han contestado cada voz.
- Los informantes de otros países que han contestado esas formas, puesto que a veces se registran respuestas aisladas o hápax.
- El concepto en inglés al que se refieren los términos.
- La definición aportada a los informantes en las encuestas de Varilex.

Tabla 19
Españolismos puros del PCIC

Grupo IV: Españolismos puros						
Españolismo PCIC	GR	% Esp	% Otros	Concepto	Definición ofrecida en las encuestas de Varilex	
A1-A2						
sujetador	7,04	63 %	Hápax: Ni, Cu, Pe	BRASSIERE	Prenda interior femenina que se utiliza para sujetar los senos.	
bragas	9,89	89 %	Hápax: PR, Pe, Ar	PANTIES	Prenda interior femenina que recubre la parte inferior de la cintura.	
caña	2,53	24 %	0%	BEER	¿Qué otro nombre tiene la cerveza?	
salón	2,03	16 %	Hápax: Pe, Ch	BEER	Sala para estar cómodo, para ver la televisión, etc.	
billete	10,17	81 %	Hápax: Cu, Ar, Co	NOTE (US: BILL)	Papel de tamaño pequeño que sirve de pasaje de un tren.	
B1-B2						
torta	1,33	12 %	0 %	SLAP	Golpe dado en la cara con la mano abierta sobre la cara.	
colega	2,11	11 %	Hápax: PR, Cu, Pa, RD, Co	OLD BOY	Forma de llamar a un amigo.	
pincho	0,72	7 %	0 %	TIDBITS	Alimento ligero que se sirve como acompañamiento de una bebida.	
celo	9,09	80 %	Hápax: Pe	CELLOPHANE TAPE	Cinta transparente, con una sustancia adhesiva en una de sus caras, que se emplea, por ejemplo, para fijar papel a otra superficie. Ej. <i>¿Tienes un poco de (***) para pegar este papel en el tablón?</i>	

C1-C2						
estar como un tonel	0,88	4 %	Hápax: Cu, Co	FAT (PLUMP)	Muy gordo.	
irse al otro barrio	3,91	5 %	Hápax: Mx, Cu	DIE	¿Qué dice cuando una persona deja de existir?	
borde	1,31	13 %	0 %	COARSE	Que tiene malos modos.	
hecho polvo	0,36	3 %	0 %	SAD	Triste: Con la noticia del accidente de su primo, se quedó [***].	
atiborrarse	0,59	5 %	0 %	EAT TOO MUCH	Hartarse de comer: El domingo, después del partido se fueron todos a comer chuletas y [***].	
tener buen saque	2,68	25 %	0 %	EAT LIKE A HORSE	Comer mucho. No sé cómo no revientan de tanto comer. Estos muchachos [***].	
ponerse morado	0,59	3 %	0 %	EAT TOO MUCH	Hartarse de comer: El domingo, después del partido se fueron todos a comer chuletas y [***].	
empollón	6,93	64 %	Hápax: PR, Py	GRIND	El que siempre está estudiando (con el sentido peyorativo).	
catear	3,64	15 %	0 %	GIVE A FAILING GRADE	Dar la pésima nota.	
gandul	1,81	7 %	0 %	LAZY STUDENT	El que no estudia mucho.	
forofos	0,41	4 %	0 %	FANS	Conjunto de personas que van regularmente a ver una competición deportiva.	
hacer novillos	1,38	13 %	Hápax: PR	PLAY TRUANT	Faltar a las clases sin avisar ni pedir permiso.	
tener enchufe	6,37	48 %	Hápax: RD, Py	HAVE PULL.	Gozar de la protección de alguien: El jefe no se atreve a llamarle la atención, porque (***).	

En la Tabla 19 observamos la lista de españolismos del *PCIC* según los datos recogidos de Varilex. Además de las 11 respuestas dadas exclusivamente por informantes españoles (*caña, torta, pincho, borde, hecho polvo, atiborrarse, tener buen saque, ponerse morado, catear, gandul, forofos*), hay otras 11 palabras que también han marcado aisladamente informantes de otros países (*sujetador, bragas, salón, billete, colega, celo, estar como un tonel, irse al otro barrio, empollón, hacer novillos, tener enchufe*). Estas respuestas aisladas o hápax pueden aparecer como consecuencia de un uso personal o de un error (Ueda, 1996c).

Así por ejemplo la palabra *sujetador* es la voz preferida en España para denominar la ‘prenda interior femenina que se utiliza para sujetar los senos’ puesto que un 70 % de encuestados en España la seleccionaron. La voz no se registra en otros países con la excepción de tres informantes aislados: uno en Nicaragua, otro en Cuba y otro en Perú. Si hubiera al menos dos informantes que afirmaran el uso de una voz podríamos valorar su utilización en esos países, pero tratándose de casos tan aislados, podemos considerar que *sujetador*, aunque se pueda comprender en determinados puntos de la geografía hispánica, es una voz preferentemente española.

Si atendemos al porcentaje de españoles que han registrado las formas podemos afirmar que hay voces léxicas peninsulares que constituyen la variante más representativa de algunos conceptos, mientras que otros españolismos conviven con otras formas alternativas válidas en España. Los españolismos con mayor presencia en el español peninsular son: *sujetador* (63 %), *bragas* (89 %), *celo* (80 %), *empollón* (64 %) y *tener enchufe* (48 %).

5.1.2. Españolismos de preferencia

Además de los españolismos *puros*, existen 4 voces en el *PCIC* que son preferentes del español peninsular, pero que también se han registrado en algunos países americanos en menor medida. Estos españolismos *de preferencia* son *jersey, bañador, ordenador, frigorífico*. En la Tabla 20 mostramos:

- El término de referencia del *PCIC* que consideramos españolismo, con su correspondiente GR.
- Los países en los que se ha marcado la voz, según los datos de Varilex.
- El número de informantes que han registrado la voz en cada país.
- El número total de encuestados en cada país.

Tabla 20
Distribución de españolismos de preferencia

Españolismos		España	Mx	Ni	Pa	RD	PR	Ec	Pe	Co	Ch	Ar	Py
jersey	resp.	114	1	1	1	1	3		1		1	1	
	total	161	35	15	15	24	40		21		33	45	
bañador	resp.	125		1		2	1		1				
	total	168		16		22	29		23				
ordenador	resp.	73			2			1		1			1
	total	87			11			3		1			8
frigorífico	resp.	77					2						
	total	177					35						

Podemos afirmar que la voz *jersey* es de preferencia española, aunque también la han registrado tres informantes de Puerto Rico. Además existen otros casos aislados o hápax en Panamá, República Dominicana, Nicaragua, Chile, Argentina, Perú y México.

La voz *bañador* es la preferente de la región de España, aunque también la han registrado dos informantes de República Dominicana. Además, constituye un hápax en Puerto Rico, Nicaragua y Perú.

Para el concepto [COMPUTER] la voz *ordenador* es la que marcaron prácticamente la totalidad de informantes de España, aunque también se registró en Panamá y, de manera aislada en Ecuador, Colombia y Paraguay.

Por último, *frigorífico* también se registra de forma preferente en España, donde convive con la voz *nevera*. En el resto del territorio hispánico, tan solo dos informantes de Puerto Rico identificaron esta palabra con su variedad dialectal.

5.2. Otros posibles españolismos sujetos a valoración

Además de los españolismos que hemos localizado a través de nuestro estudio sistemático de la base de datos de Varilex, existen en el *PCIC* otras palabras que también se circunscriben al subsistema lingüístico del español de España. Dado que la base de datos de Varilex estudia la variación de una serie limitada de nociones, consideramos importante señalar otras voces que, según fuentes secundarias, también se consideran exclusivas del español de España. En todo caso, las palabras que incluimos en la Tabla 21 exigen otros estudios o investigaciones sobre variación que nos permitan conocer mejor su distribución geográfica, para afirmar o descartar la marca españolismo.

Presentamos las referencias que hemos encontrado en la bibliografía sobre españolismos —todavía muy incipiente— y remitimos a futuras investigaciones para confirmar o descartar el uso preferentemente peninsular de estas voces.

Tabla 21
Otros españolismos según fuentes secundarias.

Grupo 5. Posibles españolismos del PCIC según otras fuentes bibliográficas		
Españolismo	Referencias bibliográficas	Comentarios
A1–A2		
zumo	Moreno de Alba (1998) Haensch (2002) Torres Torres (2001)	Para referirse al líquido que se obtiene de una fruta los hablantes americanos utilizan <i>jugo</i> .
B1–B2		
albornoz	Haensch (2002)	Para denominar la bata larga de tela de toalla que se utiliza para salir del baño en Hispanoamérica se utiliza <i>bata de baño</i> o <i>salida de baño</i> .
ambulatorio	Ávila (2005)	<i>Ambulatorio</i> es sinónimo de ‘clínica’ y se refiere el establecimiento sanitario donde se pasa consulta, pero sin camas para ingresar enfermo. No encontramos un equivalente común en el español de América.
escayola, escayolar	Ávila (2005)	La masa de yeso y agua que es moldeable y cuando se seca se queda dura se denomina <i>yeso</i> . El verbo correspondiente es <i>enyesar</i> .
parador	Ávila (2005)	Un <i>parador</i> es un hotel de gran calidad que respeta las instalaciones originales y mantiene las tradiciones del lugar donde se encuentra.
guisantes	Moreno de Alba (1998)	Los <i>guisantes</i> de España son <i>chícharos</i> en México, Nicaragua, Cuba, Honduras y República Dominicana.
cremallera	Moreno de Alba (1999)	Aunque Moreno de Alba lo cita como españolismo, según Varilex <i>cremallera</i> es la voz preferente en España pero también en Colombia (100 % de respuestas). Se han recogido además otros testimonios

			del término en Argentina, Paragua y Puerto Rico, aunque minoritarios.
descapotable	Moreno de Alba (1998)		Los registros de Varilex apuntan a que <i>descapotable</i> (31,86) es la respuesta de mayor extensión frente a <i>descapotable</i> (2,86). <i>Descapotable</i> no aparece ni en el <i>Diccionario de Lengua Española</i> ni en el <i>Diccionario de Americanismos</i> .
pantano	Haensch (2002)		Para la voz <i>pantano</i> Haensch propone el término <i>embalse</i> , que no se incluye en el <i>PCIC</i> .
C1–C2			
patio de butacas	Moreno de Alba (1998)		La planta baja de un teatro se denomina <i>luneta</i> en Costa Rica, Panamá, Cuba, Colombia, Ecuador, Bolivia, El Salvador y Nicaragua, según el <i>Diccionario de americanismos</i> .
entresuelo	Moreno de Alba (1998)		La alternativa <i>mezanine</i> para <i>entresuelo</i> se utiliza según el <i>Diccionario de americanismos</i> en México, Honduras, Costa Rica, Panamá, Cuba, Colombia, Ecuador y Bolivia.
albarán	Haensch (2002)		Frente a <i>albarán</i> para la nota de entrega de una mercancía, Haensch propone <i>boletín de entrega</i> .
chabola	Haensch (2002) Torres (2001)	Torres	Para denominar las viviendas pequeñas con pocos recursos económicos existen múltiples denominaciones en Hispanoamérica: villa miseria o barriada, cantegril, callampa, villa, chozas, casuchas, tugurios, ranchos, barracadas, llega y pon, chamizos...
archivador	Ávila (2005)		Según los datos de Varilex, la carpeta en la que se guardan documentos y papeles se corresponde con la voz <i>archivador</i> únicamente en España y en Chile.
centollo	Ávila (2005)		Según Ávila (2005) <i>centollo</i> no tiene un equivalente en México.
aparcamiento	Torres Torres (2001)		Alterna con múltiples denominaciones americanas: estacionamiento, playa de estacionamiento, parqueo, garaje, parking, parqueadero, playa, cochera.
orfanato	Moreno de Alba (1998)		Moreno de Alba propone la voz <i>orfanatorio</i> que según el <i>Diccionario de americanismos</i> se registra en México, Nicaragua, Costa Rica, República Dominicana, Puerto Rico, Ecuador, Perú, Bolivia y Chile.

6. Conclusión

En este capítulo hemos presentado nuestra propuesta de revisión y ampliación de las «Nociones específicas» del *PCIC*. El objetivo era doble. Por un lado, incorporar nociones léxicas que gocen de mayor representatividad en la geografía hispánica. Por otro, identificar aquellas nociones del *PCIC* que son exclusivas del español de España. Debido a la complejidad inherente al tratamiento de la variedad léxica hemos incluido, primeramente, algunas aclaraciones respecto al tema de las equivalencias léxicas, explicando los distintos tipos de oposiciones que encontramos en nuestra investigación. Por otra parte, las frecuentes discrepancias en los registros de las voces se deben al carácter abierto y de extensión impredecible del léxico, que traspasa fronteras nacionales y cambia rápidamente en función del contacto lingüístico que hayan tenido los distintos hablantes. La falta de sistematicidad de estudios léxicos a gran escala y la carencia de un corpus representativo de la lengua española dificultan la delimitación geográfica de los términos.

En lo que respecta al estudio de las nociones del *PCIC*, consideramos que algunas de sus voces podrían coaparecer con otras palabras sinónimas que son más representativas en español. En algunos casos hemos propuesto nociones panhispánicas, válidas en España y América, que cuentan mayor GR por lo que son fundamentales para la enseñanza de español. Estas palabras garantizan la comunicación con una mayoría de hablantes en mayor número de contextos. En otros, proponemos la incorporación de americanismos.

Nuestra hipótesis sobre la existencia de americanismos generales se verifica con la presencia de nociones panamericanas que no existen en el español de España, y cuya incorporación es fundamental para la enseñanza de español. Asimismo, hemos presentado algunos americanismos secundarios que son también importantes porque permiten integrar otras variedades del español en el inventario, lo que revaloriza y amplía las posibilidades del *PCIC* desde una mirada ecuánime ante la variedad.

En cuanto al tratamiento de los españolismos, consideramos que señalar todas las nociones de este inventario que son exclusivas de la variedad española es incuestionable. Adoptamos así una mirada panhispánica ante las variedades dialectales del español, algo que, por otra parte, está ya presente en otros capítulos del *PCIC*, así como en la mayoría de las publicaciones lingüísticas de nuestra época.

CONCLUSIONES

El objetivo principal de esta investigación era determinar la representatividad de las voces propuestas en las «Nociones específicas» del *PCIC* y proponer una solución ante la ausencia de marcas dialectales en este inventario, aportando el primer repertorio léxico del español secuenciado en niveles de referencia con las nociones más representativas de toda la población hispanohablante.

Así pues, la aportación fundamental de este trabajo consiste en el diseño de una propuesta de ampliación de las «Nociones específicas» con vocabulario representativo de la globalidad del mundo hispánico. Nos hemos apoyado en la idea de que el *PCIC* constituye un primer nivel de concreción lingüística, que exige niveles de ajuste y adaptación por parte de sus usuarios en función de las distintas situaciones de enseñanza y aprendizaje. Por ello, resulta fundamental proponer voces propias de otras regiones geográficas para garantizar la selección léxica que mejor se ajuste a cada contexto.

Esta tesis demuestra que, aunque el *PCIC* concede gran importancia al fenómeno de la variación diatópica e incluye especificaciones lingüísticas de diferentes regiones del español en la mayoría de sus inventarios, en el repertorio de «Nociones específicas» se presenta un léxico exclusivamente peninsular sin marcas de variación geolectal. La complejidad inherente a los estudios de variación léxica, condicionados por el concepto controvertido de sinonimia, el uso de cuestionarios

para estudiar la distribución del vocabulario o la inexistencia del *Diccionario de americanismos* en la época de elaboración del *PCIC*, son algunos de los factores que pueden explicar la ausencia de marcas dialectales en el repertorio léxico de esta obra de referencia para la enseñanza del español.

Antes de elaborar nuestra propuesta, hemos presentado un estado de la cuestión crítico sobre las investigaciones lexicográficas existentes. Además de considerar las posibilidades que ofrecen tanto los atlas lingüísticos como las monografías dialectales, los diccionarios, las bases de datos o los corpus lingüísticos, también hemos reflexionado sobre los mecanismos de incorporación de palabras al lexicón mental, las actividades didácticas que favorecen este proceso y, sobre todo, los criterios para la selección léxica en la enseñanza de lenguas.

Teniendo en cuenta todos estos factores, en este trabajo hemos fijado unas líneas de ampliación de las «Nociones específicas» que siguen la orientación plurinormativa del *PCIC*. Nuestra propuesta se ha centrado en tres ejes: la búsqueda de voces panhispánicas, válidas en España y América, que garanticen la rentabilidad léxica; la incorporación de americanismos altamente representativos, que consideramos incuestionables cuando tienen mayor representatividad que las nociones peninsulares; y, por último, la asignación de la marca ‘españolismo’ en los términos que son exclusivos o preferentes del español de España.

Para fundamentar nuestra aportación, hemos propuesto una fórmula de análisis demolingüístico del léxico hispánico que tiene por objetivo medir la extensión y la frecuencia de las voces para garantizar la incorporación de variantes representativas, extensas y rentables en el *PCIC*. Nuestro cálculo ponderado de la representatividad ha sido posible gracias a las informaciones recogidas por el proyecto Varilex, que nos ha concedido libre acceso y uso de su base de datos. Así, mediante un estudio descriptivo de los datos, hemos defendido un análisis que pretende reflejar la distribución de las voces del español, en función del concepto multidimensional de representatividad, asignando un índice numérico a cada una de las variantes.

El análisis demolingüístico propuesto en nuestro trabajo presenta ventajas indiscutibles para el estudio de la variación diatópica y ofrece información importante relacionada con la estructuración de las voces, la localización, la vitalidad, la innovación, el conservadurismo o las relaciones interdialectales de las variantes léxicas. A la luz de nuestros resultados, hemos logrado ratificar las hipótesis de trabajo planteadas en nuestra tesis, de las que extraemos las siguientes conclusiones:

H1: Las nociones léxicas del PCIC son una muestra del español de España que no incorpora voces representativas de otras variedades de español.

Nuestro análisis de las «Nociones específicas» confirma esta hipótesis. La variedad geolectal que toma el *PCIC* para la selección léxica es exclusivamente peninsular. Así, las «Nociones específicas» no incluyen vocabulario relativo a otras variedades geolectales, a pesar de que sí lo hacen otros de sus inventarios.

H2: En el PCIC existen variantes léxicas que son exclusivas del español peninsular (españolismos).

Nuestro trabajo exhaustivo con los datos de Varilex confirma que en las «Nociones específicas» aparecen voces exclusivas del español peninsular, que no se utilizan en América (*sujetador, pincho, colega, borde...*) Aunque no es un número muy elevado de nociones, estas palabras, regionales y minoritarias en el mundo hispánico, se incluyen frecuentemente en manuales y currículos del español como segunda lengua o lengua extranjera sin ninguna aclaración o marca de extensión.

De acuerdo con Werner (2001) y Stork (2006), consideramos fundamental incluir la marca de españolismo en los repertorios lexicográficos. Tal y como se constata en el capítulo 2, resulta muy complicado acceder a fuentes de consulta fiables y contrastadas sobre variación léxica que permitan conocer las equivalencias semánticas de los términos en las dos direcciones, distinguiendo las palabras que se emplean exclusivamente en Hispanoamérica o en España e identificando el término

equivalente en otras regiones geográficas, junto a las precisiones de significado oportunas cuando sea necesario.

Nuestro análisis de las nociones del *PCIC* confirma la existencia de 22 españolismos. Se trata de voces que han sido aportadas solo por informantes de España. Además, se han localizado 4 voces que, aunque también se hayan registrado en otros países americanos, son voces preferentes del español peninsular. Por último, hemos identificado 17 nociones del inventario que han sido consideradas españolismos según otras obras bibliográficas pero que exigen investigaciones más profundas para confirmar o descartar su uso preferentemente peninsular.

H3: En el léxico del español podemos encontrar variantes panhispánicas que son comunes a una mayoría de hispanohablantes (panhispanismos) que deben formar parte del léxico de ELE.

Nuestra propuesta demolingüística confirma la existencia de panhispanismos o voces que garantizan la comunicación entre los hispanohablantes de cualquier procedencia. Hemos defendido que estas palabras son esenciales para los estudiantes de español y su presencia en manuales y repertorios léxicos debe estar garantizada, puesto que representa el verdadero estándar. Estas nociones tienen gran extensión en el mundo hispánico y forman parte del léxico pasivo o activo actual de los hispanohablantes.

En nuestro estudio de las «Nociones específicas» hemos localizado 23 voces para las que existen alternativas panhispánicas con mayor Grado de Representatividad (*suéter, traje de baño, refrigerador, bofetada...*). Su inclusión, junto a los términos del subsistema lingüístico del español de España, otorga mayor representatividad a la selección léxica del inventario. Además de estas voces, incluimos una lista con otros panhispanismos sujetos a valoración, que requieren una investigación más profunda respecto a las connotaciones de las equivalencias léxicas y su distribución en el mundo hispánico.

H4: En el léxico del español existen variantes comunes a las regiones de Hispanoamérica que no existen en el español peninsular (americanismos) pero que deben formar parte del léxico de ELE.

Nuestro análisis coincide con Haensch (2002) y demuestra la existencia de americanismos generales o panamericanismos que se oponen a la variedad del español de España. En nuestra investigación hemos defendido la inclusión de estas variantes (*computadora, carro, pizarrón, clóset...*). No podemos obviar que el continente americano concentra un 90 % de los hablantes de español. Se ha considerado que esta ampliación con voces americanas resulta incuestionable, sobre todo, cuando las voces tienen mayor Grado de Representatividad que los términos de referencia del *PCIC*. Así, hemos planteado la inclusión de 40 americanismos, que constituyen una aportación fundamental en términos de extensión y representatividad para las «Nociones específicas». Todas las voces incluidas en esta propuesta están avaladas por el *Diccionario de americanismos* y contribuyen a aumentar la solidez referencial del *PCIC* desde una mirada plurinormativa, tal y como se hace en otras secciones de este documento.

H5: La identificación de las variantes léxicas definidas en H2, H3 y H4 permite proponer una ampliación del PCIC representativa del español urbano actual.

El capítulo 6 de esta tesis ratifica esta posibilidad, proponiendo una ampliación y marcación de las «Nociones específicas» con variantes léxicas representativas del español urbano actual.

Una vez ratificadas estas hipótesis, y a la luz de nuestros resultados demolingüísticos, podemos afirmar que esta tesis cumple con el objetivo central de nuestra investigación y aporta, con el capítulo 6, una propuesta personal razonada de

ampliación y marcación de las «Nociones específicas» del *PCIC* con léxico representativo del español. Gracias a la metodología rigurosa de Varilex, que ha resultado adecuada para nuestro estudio de la variación, hemos presentado el Grado de Representatividad de 182 nociones específicas, valorando de manera sistemática la inclusión de geosinónimos representativos del español actual.

En relación con los objetivos específicos de investigación, nuestra tesis aporta, en primer lugar, una fórmula matemática que permite medir la representatividad de las voces léxicas del español. El cálculo ponderado del Grado de Representatividad ha sido una de las aportaciones teóricas fundamentales de este trabajo puesto que, a diferencia de otras propuestas anteriores, nuestra fórmula demolingüística se aproxima con mayor rigor al fenómeno de la variación.

Por otra parte, el análisis y procesamiento de los registros de Varilex, mediante una base de datos relacional de creación propia, nos ha permitido implementar la fórmula ponderada de la representatividad, que tiene en cuenta los datos de estimación demográfica y las respuestas registradas por los informantes. Todo ello hace posible la identificación de panhispanismos y americanismos altamente representativos del español.

Otro de los objetivos específicos de la tesis consistía en el diseño de un sistema de representación cartográfica de los resultados. Este aspecto ha tomado especial relevancia en el transcurso de la investigación, superando notablemente nuestras expectativas iniciales. Desde los albores de nuestro trabajo, la representación de los resultados de variación ha constituido una prioridad. Ante la necesidad de diseñar un sistema de representación que tuviera en cuenta las posibilidades que ofrecen las tecnologías actuales, finalmente optamos por implementar una aplicación virtual dinámica que permite una excelente lectura de nuestros resultados demolingüísticos. Además, se pueden consultar desde cualquier ordenador, lo que expande el alcance y la visibilidad de este trabajo y acrecienta el interés y la relevancia de esta investigación.

Además de la creación de la aplicación virtual, con el objetivo de mostrar nuestros resultados en formato papel, hemos generado un conjunto de mapas estáticos de los conceptos estudiados que ha resultado muy conveniente para la representación de datos demolingüísticos. El sistema de mapas de símbolos proporcionales, que se utiliza igualmente en la aplicación web, permite una correcta legibilidad de los resultados y hace posible representar cartográficamente el uso de diferentes voces léxicas dentro de un mismo país. Así, mediante círculos de diferentes tamaños en función de la mayor o menor representatividad de las voces, hemos incluido en el Anexo III un mapa para cada uno de los 182 conceptos estudiados del *PCIC*, que presenta la distribución de las tres variantes léxicas con mayor Grado de Representatividad de cada uno de ellos.

Nuestra valoración final de esta investigación es que la posibilidad de plantear un tratamiento riguroso de la variación léxica que tenga en cuenta la extensión de las voces y el número de hablantes que las conocen y utilizan, no solo es posible sino que es, además, estrictamente necesario para la enseñanza del léxico del español. A pesar de la reiterada complejidad que se atribuye al estudio de la variación léxica, y sin obviar las dificultades que plantea la investigación en este campo de estudio, los investigadores tenemos dos opciones: abandonar la misión de estudiar el vocabulario, por la complejidad que lleva implícita esta tarea, o explorar nuevas aproximaciones al fenómeno de la variación léxica que tengan en cuenta los datos y recursos de las investigaciones actuales que trabajan para estructurar y clasificar el léxico del español.

En definitiva, ante la ausencia de un inventario léxico hispánico con marcas geolectales y distribuido en niveles de aprendizaje, nuestra tesis constituye una respuesta, razonada y documentada, para integrar en el *PCIC* una selección de voces representativas que cumplan con criterios de rentabilidad y extensión léxica. El capítulo 6 de este trabajo es la culminación de nuestra investigación y la consecución del principal objetivo: realizar una propuesta de

ampliación de las «Nociones específicas» que sea representativa de la diversidad geolectal del español e incluya, por un lado, voces representativas de toda la geografía hispánica (*panhispanismos*); por otro, voces americanas con alta rentabilidad y extensión (*americanismos*); y, por último, una marca en aquellas nociones que son exclusivas del español de España (*españolismos*).

Nuestra aportación a este campo de estudio no habría sido posible sin los datos e informaciones recogidas por el grupo de Varilex, que es la única investigación exhaustiva, sistemática y continuada durante más de dos décadas, que ha publicado sus resultados sobre léxico en las grandes ciudades hispánicas.

En este sentido, es fundamental invitar a este y a otros grupos de investigación a continuar recopilando datos sobre la variación léxica, ampliando el número de encuestas y conceptos estudiados, para que los datos analizados en trabajos como el nuestro puedan considerarse más significativos. Asimismo es incuestionable que los medios actuales para la recolección de la información y las enormes posibilidades que permiten las técnicas estadísticas para el análisis de los datos aportan un cambio sustancial en el estudio del vocabulario.

Como en cualquier investigación, algunos temas no han podido tratarse con la profundidad oportuna. Por el momento, no existe un estudio específico que se ocupe de la validación de los datos de Varilex, depurando los errores en los registros informatizados o las diferencias connotativas y denotativas respecto a las respuestas aportadas por los informantes. La presente investigación no puede suplir tal ausencia y esto supone que los resultados demolingüísticos deben tratarse con cautela antes de plantear cualquier aplicación de los mismos. El estudio minucioso sobre la distribución léxica de cada concepto estudiado por Varilex queda pendiente, por tanto, para futuras investigaciones.

Otra cuestión que queda aplazada, y de la que nos habría gustado ocuparnos con mayor profundidad, es el tema de los españolismos. La falta de investigaciones al respecto y la ausencia de esta marca en los diccionarios (Werner, 2002) complica enormemente el estudio de este fenómeno.

El problema de la distribución de las variantes geolectales en los diferentes niveles de dominio tampoco queda resuelto, y merece también un estudio exhaustivo aparte sobre la secuenciación de estas voces en el aprendizaje de español. Tanto el proceso de adquisición de este tipo de nociones como las técnicas y actividades que facilitan el tratamiento de las variantes de un mismo concepto, en función de la región geográfica, son aspectos poco desarrollados en el ámbito de la didáctica de español como segunda lengua y lengua extranjera.

Esperamos que nuestra tesis sirva de base para afrontar estas cuestiones fundamentales relacionadas con el tratamiento de la variación léxica en las clases de español y que surjan de ella trabajos posteriores que arrojen nuevas perspectivas para el estudio de este fenómeno todavía poco presente en la bibliografía actual pero que ofrece múltiples posibilidades para la investigación lingüística.

BIBLIOGRAFÍA

Abad F. (1993). La variación lingüística. *Revista Española de Lingüística*, 23 (I): 73-86.

Almela, R., Cantos, P., Sánchez, A., Sarmiento, R. y Almela, M. (2005). *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Universitas.

Alvar, M. (1969). *Tendencias de la lingüística actual*. Madrid: Universidad Autónoma de Madrid.

Alvar, M. (1980). Hacia una geografía lingüística de América. En J. M. Lope Blanch, *Perspectivas de la investigación lingüística en Hispanoamérica* (pp. 79-92). México: Universidad Nacional Autónoma de México.

Alvar, M. (1982). Atlas lingüísticos y diccionarios. *Lingüística Española Actual*, 4 (2): 253-324.

Alvar, M. (1996). *Manual de dialectología hispánica: El español de América*. Barcelona: Ariel.

Alvar M. y Quilis A. (1984). *Atlas lingüístico de Hispanoamérica: cuestionario*. Madrid: Instituto de Cooperación Iberoamericana.

Andión Herrero, M.^a A. (2001). La lengua en la prensa española e hispanoamericana en Internet: El fantasma de la diferenciación. *Español Actual*, 76: 71-92.

Andión Herrero, M.^a A. (2003). El léxico americano en los manuales españoles de E/LE. Consideraciones. En F. Moreno Fernández, J. A. Samper Padilla, M. Vaquero, M. L. Gutiérrez Araus, C. Hernández Alonso, F. Gimeno Menéndez, *Lengua, variación y contexto. Homenaje a Humberto López Morales* (vol. 1, pp. 105-126). Madrid: Arco/Libros.

Andión Herrero, M.^a A. (2007). Las variedades y su complejidad conceptual en el diseño de un modelo lingüístico para el español L2/ LE. *Estudios de Lingüística* (ELUA), 21: 21-33.

Andión Herrero, M.^a A. (2008). Modelo, estándar y norma..., conceptos imprescindibles en el español L2/LE. *Revista Española de Lingüística Aplicada*, 21: 9-26.

Andión Herrero, M. A. (2008). La diversidad lingüística del español: la compleja relación entre *estándar, norma y variedad*. En A. Moreno Sandoval, *Actas del VIII Congreso de Lingüística General. El valor de la diversidad [meta]lingüística*. [CD-ROM]. Madrid: Universidad Autónoma de Madrid.

Andión Herrero, M.^a A. y Casado Fresnillo, C. (2014). *Variación y variedad del español aplicadas a E-LE/L2*. [Libro electrónico]. Madrid: UNED.

Aoto, S. (2003). Acerca de la variación de las perífrasis verbales de gerundio. *Varilex*, 11: 10-28.

Aoto, S. (2004). Aplicación de LMS en la cartografía electrónica de Varilex. *Varilex*, 12: 6-9.

Aoto, S. (2005). Estudio sobre la variación léxica del español en Paraguay. *Varilex*, 13: 6-27.

Asociación de Academias de la Lengua Española y Real Academia Española (2009). *Nueva Gramática De La Lengua Española*. Madrid: Espasa Libros.

Asociación de Academias de la Lengua Española (2010). *Diccionario de americanismos*. Madrid: Santillana.

Ávila, R. (1990). Las palabras de todos y las de cada uno: un análisis estadístico del español hablado en México. En V. Demonte y B. Garza Cuarón, *Estudios de Lingüística de España y México* (pp. 335-350). México: UNAM-El Colegio de México.

Ávila, R. (1997). Variación léxica: connotación, denotación, autorregulación. *Anuario de Letras: Lingüística y Filología*, 35: 77-102.

Ávila, R. (2001). Los medios de comunicación masiva y el español internacional. En *II Congreso Internacional de la Lengua Española: el español en la Sociedad de la Información*, Valladolid. Recuperado de: http://congresosdelalengua.es/valladolid/ponencias/unidad_diversidad_del_espanol/1_la_norma_hispanica/avila_r.htm

Ávila, R. (2005). Españolismos y mexicanismos: un análisis cuantitativo. *Nueva Revista de Filología Hispánica*, LIII: 413-455.

Baldinger, K. (1964). Semasiologie et onomasiologie. *Revue de Linguistique Romane*, 28: 250-272.

Baldinger, K. (1970). *Teoría semántica. Hacia una semántica moderna*. Madrid: Ediciones Alcalá.

Beacco, J. C. et al. (2004). *Niveau B2 pour le français: un référentiel*. Paris: Didier.

Beaven, T. y Garrido, C. (2000). El español tuyo, el mío, el de aquel... ¿cuál para nuestros estudiantes? ¿Qué español enseñar? En M. A. Martín Zorraquino y C. Díez Pelegrín, *Actas del XI Congreso Internacional ASELE: Norma y Variación Lingüísticas en la enseñanza del Español a Extranjeros* (pp. 13-16). Zaragoza: Universidad de Zaragoza.

Béjoint, H. (1979). The use of informants in dictionary-making. *ITL International Journal of Applied Linguistics*, 45/46: 25-29.

Bentivoglio, P. (2001). La variación sociosintáctica en español. En *II Congreso Internacional de la Lengua Española: El Español en la sociedad de la información*, Valladolid. Recuperado de: http://congresosdelalengua.es/valladolid/ponencias/unidad_diversidad_del_espanol/1_la_norma_hispanica/bentivoglio_p.htm

Berruto, G. T. (1979). *La semántica*. México: Nueva Imagen.

Blanco, C. (2000). El dominio del concepto de norma como presupuesto del profesor de ELE. En M.^a A. Martín Zorraquino y C. Díez Pelegrín, *Actas del XI Congreso Internacional ASELE: ¿Qué español enseñar? Norma y variación lingüísticas en la enseñanza del español a extranjeros* (pp. 209-216). Zaragoza: Universidad de Zaragoza.

Bravo, E. (2008). *El español internacional*. Madrid: Arco/Libros.

Bréal M. (2001). *Ensayo de semántica: Ciencia de las significaciones*. Pamplona: Analecta.

Bustos Gisbert, E. (1995). Dialectología, sociolingüística y español en América. *Dicenda: Cuadernos de filología hispánica*, 13: 39-64.

Cahuzac, P. (1980). La división del español de América en zonas dialectales: solución etnolingüística o semántico-dialectal. *Lingüística Española Actual*, 2: 385-461.

Canfield, D. (1981). *Spanish pronunciation in the Americas*. Chicago: University Chicago Press.

Castillo Peña, C. (1990). Del atlas lingüístico al diccionario: Experiencias lexicográficas. En *Actas del Congreso de La Sociedad Española de Lingüística* (vol. 1, pp. 363-371). Madrid: Gredos.

Chacón García, C. (2014). El léxico del español: campos léxicos y variedades. En M.^a A. Andión Herrero y C. Casado Fresnillo, *Variación y variedad del español aplicadas a E-LE/L2*. [Libro electrónico]. Madrid: Editorial UNED.

Chacón García, C. (2016). *Grado de Representatividad léxica del español*. [Aplicación virtual dinámica]. Disponible en: <https://carmenchacon.shinyapps.io/PCIC/>

Chambers, J. K. y Trudgill, P. (1994). *La dialectología*, Madrid: Visor Libros.

Chaudenson, R. (1993). Francophonie, “Français zero” et français regional. En D. Robillard y M. Beniamino (dir.) *Le Français dans l’espace francophone*. París: Champion.

Clyne, M. (1992). *Pluricentric languages. Differing norms in different nations*. Berlin/New York: De Gruyter.

Consejo de Europa (2002). *Marco Común Europeo de Referencia para las lenguas*. Estrasburgo: Consejo de Europa, Ministerio de Educación, Cultura y Deporte/Instituto Cervantes.

Colón Doménech, G. (2002). *Para la historia del léxico español*. Madrid: Arco/Libros.

Coseriu, E. (1955). *La geografía lingüística*. Montevideo: Universidad de la República. Facultad de Humanidades y Ciencias.

- Coseriu, E. (1967). Sistema, norma y habla. *Teoría del lenguaje y Lingüística general*. Madrid: Gredos.
- Coseriu, E. (1981). *Lecciones de Lingüística General*. Madrid: Gredos.
- Coseriu, E. (1992). Lingüística histórica e historia de las lenguas. *Boletín de Filología*, 33 (1): 27-33.
- Coseriu, E. y Azáceta, J. M. (1981). *Lecciones de lingüística general*. Madrid: Gredos.
- Criado de Val, M. (1972). *Fisionomía del español y de las lenguas modernas*. Madrid: Saeta.
- Crystal, D. (1994). Which English –or English which? En M. Hayhoe y S. Parker (eds.), *Who owns English?* (pp. 108-114). Oxford: Oxford University Press.
- Crystal D. (1997). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- De Sterck, G. (2000). Gramática: registros y áreas geográficas en lingüística: valores y usos de las formas verbales en –ra, –se, –ría y –re. En J. De Kock (ed.), *Gramática española: enseñanza e investigación*, vol. II (7). Salamanca: Universidad de Salamanca.
- Demonte, V. (2001). El español estándar (ab)suelto. Algunos ejemplos del léxico y la gramática. En *II Congreso Internacional de la Lengua Española: El español en la Sociedad de la Información*, Valladolid. Recuperado de: http://congresosdelalengua.es/valladolid/ponencias/unidad_diversidad_del_espanol/1_la_norma_hispanica/demonte_v.htm
- Demonte, V. (2003). Lengua estándar, norma y normas en la difusión actual de la lengua española. *Circunstancia: Revista de Ciencias Sociales del Instituto Universitario de Investigación Ortega y Gasset*, 1. Recuperado de: <http://www.ortegaygasset.edu/publicaciones/circunstancia/ano-i---numero-1---abril-2003/estados-de-la-cuestion/lengua-estandar--norma-y-normas-en-la-difusion-actual-de-la-lengua-espanola>
- Dent, B.D. (1999). *Cartography. Thematic Map Design*. Nueva York: WCB McGraw-Hill.
- Domene Verdú, J. F. (2010). *Lingüística y matemáticas: axiomatización de la teoría gramatical y su aplicación a la tipología lingüística*. San Vicente del Raspeig (Alicante): Publicaciones Universidad de Alicante.

Elías-Olivares, L., y Valdés, G. (1979). Language diversity in chicano speech communities: Implications for language teaching. *Working Papers in Sociolinguistics*, 54. Recuperado de: <http://eric.ed.gov/?id=ED168345>

Entwistle, W. J. (1973). *Las lenguas de España: castellano, catalán, vasco y gallego-portugués*. Madrid: Istmo.

Eugenio, C. (1967). *Teoría del lenguaje y lingüística general*. Madrid: Gredos.

Fajardo Aguirre, A. (2010). La lexicografía del español de América: Evolución, tipología y metodologías. En M. Aleza Izquierdo y J. M.^a Enguita Utrilla, *La lengua española en América: normas y usos actuales* (pp. 317-354). Valencia: Universitat de València.

Fernández, F. J. (1994-2004). Comentarios. En *Comentarios críticos sobre Varilex* (1-9). *Variación léxica del español del mundo*. Tokio: Universidad de Tokio.

Figueras, N. (2008). El MCER, más allá de la polémica. En F. Puig, *Monográficos MarcoELE*, 7: 26-35. Recuperado de: <http://marcoele.com/el-mcer-mas-alla-de-la-polemica/>

Finke, P. (2002). Die Wildnis der Sprache. Ein kurzer Galopp durch die Ökologuistik. *Sprachtheorie und germanistische Linguistik*, 12 (1): 43-80.

Flórez Márquez, Ó. A. 2000. ¿Qué español enseñar? o ¿Cómo y cuándo "enseñar" los diversos registros o hablas del castellano? En M.^a A. Martín Zorraquino y C. Díez Pelegrín, *Actas del XI Congreso Internacional ASELE: Norma y variación lingüísticas en la enseñanza del español a extranjeros* (pp. 311-316). Zaragoza: Universidad de Zaragoza.

Forastieri Braschi, E., Cardona, J., López Morales, H. y Morales de Walters, A. (1999). *Estudios de lingüística hispánica: homenaje a María Vaquero*. San Juan: Universidad de Puerto Rico.

García de Diego, V. (1950). El castellano como complejo dialectal y sus dialectos internos. *Revista de Filología Española*, 34: 107-124.

García Fernández, E. (2010). *El tratamiento de las variedades de español en los manuales de EL2/LE*. [Memoria de máster]. Universidad Nacional de Educación a Distancia. Recuperado de http://www.mecd.gob.es/dctm/redele/Material-RedEle/Biblioteca/2010_BV_11/2010_BV_11_2_semestre/2010_BV_11_08GarciaFernandez.pdf?documentId=0901e72b80e1f5b7

García Izquierdo, I. (2009). El español neutro en los discursos de especialidad: ¿Mito, utopía o realidad? *Íkala: Revista de Lenguaje y Cultura*, 14 (23): 15-39.

García Mouton, P. (1990). El estudio del léxico en los mapas lingüísticos. En F. Moreno Fernández (ed.), *Estudios sobre variación lingüística* (pp. 27-75). Alcalá de Henares: Universidad de Alcalá de Henares.

García Mouton, P. (1991). Dialectometría y léxico en Huesca. En T. Buesa Oliver y J. M. Enguita Utrilla, *I Curso de Geografía Lingüística de Aragón* (pp. 311-326). Zaragoza: Institución Fernando el Católico.

García Mouton, P. (1996). Dialectología y geografía lingüística. En M. Alvar, *Manual de dialectología hispánica. El español de España* (pp. 63-77). Barcelona: Ariel Lingüística.

García Mouton, P. (2003). *El Atlas Linguistique Roman (ALiR): Historia y características de un proyecto europeo*. Madrid: Arco/Libros.

García Mouton, P. (2006). Los Atlas lingüísticos y las variedades del español de América. *Boletín Hispánico Helvético*, 8: 111-122.

García Mouton, P. (2007). Las disciplinas tradicionales (II): Dialectología y geolingüística. En J. E. Gargallo y M. Reina, *Manual de lingüística románica* (pp. 319-349). Barcelona: Ariel.

Gimeno Menéndez, F. (1990). *Dialectología y sociolingüística españolas*. Alicante: Universidad de Alicante.

Giráldez Ceballos-Escalera, J. (2009). *Las colocaciones léxicas en el lenguaje jurídico del derecho civil francés*. [Tesis doctoral]. Recuperado de: <http://eprints.ucm.es/8061/>

Gisbert, E. B. (1995). Dialectología, sociolingüística y español de América. *Dicenda. Cuadernos De Filología Hispánica*, 13: 39-64.

Goebel, H. (1981). Eléments d'analyse dialectométrique (avec application à l' AIS). *Revue de linguistique romane*, 45: 349-420.

Gómez Font, A. (2003). El español global en los medios de comunicación hispanos de los EE. UU. y en los grandes medios transnacionales. [Publicación digital]. *Centro Virtual Cervantes*. Recuperado de: http://commonweb.unifr.ch/artsdean/pub/gestens/f/as/files/4740/21370_094832.pdf

Gutiérrez Ordóñez S. (1996). *Introducción a la semántica funcional*. Madrid: Síntesis.

Haensch G. (1982). *La lexicografía: de la lingüística teórica a la lexicografía práctica*. Madrid: Gredos.

Haensch, G. (1986). La situación actual de la lexicografía del Español de América. *Revista De Filología Románica*, 4: 281-293.

Haensch, G. (2001). Español de América y español de Europa (1.ª parte). *Panace@*, 2 (6): 63-72.

Haensch, G. (2002). Español de América y español de Europa (2.ª parte). *Panace@*, 3 (7): 37-64.

Hasselgreen, A. (2003). *Bergen «Can Do» Project*. Strasbourg: Council of Europe Publishing.

Henríquez Ureña, P. (1921). Observaciones sobre el español de América. *Revista de Filología Española*, 8: 357-390.

Huhta, A., Luoma, S., Oscarson, M., Sajavaara, K., Takala, K. y Teasdale, A. (2002). DIALANG, A diagnostic language assessment system for adult learners. En J. C. Alderson, *Common European Framework of Reference for Languages: learning, teaching, assessment. Case studies*. Strasbourg: Council of Europe Publishing.

Iglesias, L. y Sopova, J. (2012). Jorge Edwards: «Uno que habla la lengua de Cervantes no está en inferioridad de condiciones con alguien que habla la lengua de Shakespeare». *El Correo De La UNESCO*. Recuperado de: <http://www.unesco.org/new/es/media-services/single-view-tv-release/news/jorge-edwards-uno-que-habla-la-lengua-de-cervantes-no-esta-en-inferioridad-de-condiciones-con-alguien-que-habla-la-lengua-de-shakespeare/>

Instituto Cervantes. (2006). *Plan Curricular del Instituto Cervantes: Niveles de referencia para el español*. Madrid: Instituto Cervantes.

Instituto Cervantes. (2008) *Enciclopedia del español en los Estados Unidos: Anuario del Instituto Cervantes 2008*. [Recurso electrónico]. Recuperado de http://cvc.cervantes.es/Lengua/anuario/anuario_08/default.htm

Jaberg K. (1955). Grossräumige und kleinräumige Sprachatlanten. *Vox romanica*, 14: 1-61.

Jenkins, J. (2006). Current perspectives on teaching world englishes and English as a *lingua franca*. *Tesol Quarterly*, 40 (1): 157-181.

- Jud J. (1917). Probleme der altromanischen Wortgeographie. *Zeitschrift für romanische Philologie*, 38: 1-75.
- Kirkpatrick, A. (2006). Which model of english: Native-speaker, nativised or lingua franca? En R. Rubdy y M. Saraceni (eds.), *English in the World: Global rules, global roles* (pp. 71-83). London: Continuum.
- Koike, K. (2001). Variación fraseológica del español. *Varilex*, 9: 77-92.
- Labov W. (1983). *Modelos sociolingüísticos*. Madrid: Cátedra.
- Lapesa, R. (1992). *Historia de la lengua española*. Madrid: Gredos.
- Lavandera, B. (1978). Where does the sociolinguistic variable stop? *Language in Society*, 7: 171-182.
- Lewandowski, T. (1982). *Diccionario de lingüística*. Madrid: Cátedra.
- Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*. London and New York: Routledge.
- Llorente Maldonado, A. (1986). *Aspectos lingüísticos de la tierra de Béjar*. Salamanca: Cuerpo de Hombre.
- Llorente Pinto, M.^a del Rosario. (2013). *El español neutro existe*. [Recurso electrónico]. Recuperado de http://gredos.usal.es/jspui/bitstream/10366/121971/3/DLE_LlorentePinto_El_espanol_neutro_existe.pdf
- Lope Blanch, J. M. (1995). El problema de la lengua española en América. *Nueva Revista de Filología Hispánica*, 43 (1): 17-36.
- Lope Blanch, J. M. (1995). Los estudios sobre la fonética del español americano y las lenguas amerindias. *Thesaurus: Boletín del Instituto Caro y Cuervo*, 50 (1): 129-147.
- Lope Blanch, J. M. (1997). *La lengua española y sus problemas*. México: Universidad Nacional Autónoma de México.
- Lope Blanch, J. M. (2000). Diversidad léxica y uniformidad gramatical en torno al porvenir de la lengua española. *Revista de Filología Española*, 80: 201-214.
- Lope Blanch, J. M. y Bartoš, L. (1971). *El presente y el porvenir del español en América*. Brno: Universita J. E. Purkyně.

López Chávez, J. (1993). Léxico fundamental panhispánico: Realidad o utopía. En A. Matus Oliver *et al.* (eds.), *El español de América. Actas Del IV Congreso Internacional sobre "El Español de América"* (vol. 2, pp. 1006-1014). Santiago de Chile: Pontificia Universidad Católica de Chile.

López González, A. M. (2002). La lengua internacional de los medios de comunicación: Una convergencia de modelos lingüísticos. En M. Pérez Gutiérrez, *XIII Congreso Internacional De ASELE: El español, lengua del mestizaje y la interculturalidad*. Murcia: Centro Virtual Cervantes. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/asele_xiii.htm

López Morales, H. (1983). Lingüística estadística. En H. López Morales, *Introducción a la Lingüística actual* (pp. 209-225). Madrid: Playor.

López Morales, H. (1992). *El español del Caribe*. Madrid: Mapfre.

López Morales, H. (1995). Disponibilidad léxica en Puerto Rico: Primera aproximación. En L. Rodríguez-Fonseca e I. Vázquez, *III Seminario Internacional sobre «Aportes de la lingüística a la enseñanza de la lengua materna»*. Carolina: Universidad de Puerto Rico.

López Morales, H. (2001). Tendencias actuales del léxico hispanoamericano. *Revista de Occidente*, 240: 5-24.

López Morales, H. (2010). *La andadura del español por el mundo*. Madrid: Taurus.

Lüdtke, J. y Schmitt, C. (2004). *Historia del léxico español: enfoques y aplicaciones: Homenaje a Bodo Müller*. Madrid: Iberoamericana.

Mangado Martínez, J. J. (2007). Norma idiomática y lengua oral. En E. Balsameda Maestu, *Las destrezas orales en la enseñanza del español L2-LE: XVII Congreso Internacional de ASELE* (pp. 39-64). La Rioja: Universidad de La Rioja. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/17/17_0039.pdf

Manning C. y Schütze H. *Foundations of Statistical Natural Language Processing*. Cambridge: MIT Press.

Martín Camacho, J. C. (2008). La variación formal y conceptual en la terminología de la lingüística. Reflexiones sobre sus consecuencias positivas y negativas. En A. Moreno Sandoval, *Actas del VIII Congreso de Lingüística General*, (pp. 67-67) [Recurso electrónico]. Recuperado de <http://www.lllf.uam.es/clg8/actas/ActasCLG8.pdf>

Martín Peris, E. (2001). Textos, variedades lingüísticas y modelos de lengua en la enseñanza del español como lengua extranjera. *Carabela. Modelos de uso de la lengua española*, 50: 103-137.

Martín Zorraquino, M.^a A. y Díez Pelegrín, C. (2001). *Actas del XI Congreso Internacional ASELE: Norma y variación lingüísticas en la enseñanza del Español a Extranjeros*. Zaragoza: Universidad de Zaragoza.

Mendizábal, N. (1997). ¿Qué lengua enseñar? La norma estándar en la enseñanza del español como L2. *Frecuencia*, L(5) 50-52.

Messner, D. (2006). Sobre estudios históricos del léxico español. *Alfa: Revista de Lingüística*, 2: 69.

Moreno de Alba, J. (1998). Sobre españolismos léxicos. [Recurso electrónico]. Recuperado de http://archivo.estepais.com/inicio/historicos/87/19_cultura_lexico_alba.pdf

Moreno de Alba, J. (2001). *El español en América*. México: Fondo de Cultura Económica.

Moreno Fernández, F. (1990). *Metodología sociolingüística*. Madrid: Gredos.

Moreno Fernández, F. (1994). Sociolingüística, estadística e informática. *Lingüística*, (6): 95-154.

Moreno Fernández, F. (1997). ¿Qué español hay que enseñar? Modelos lingüísticos en la enseñanza del español/LE. *Cuadernos Cervantes de la lengua española*, 3(14), 7-15.

Moreno Fernández, F. (1997). Geografía lingüística y variacionismo. En J. Dorta Luis y M. V. Almeida Suárez, *Contribuciones al estudio de la lingüística hispánica: Homenaje al profesor Ramón Trujillo*, vol. 2 (347-356). Barcelona: Montesinos Editor.

Moreno Fernández, F. (1998). *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel.

Moreno Fernández, F. (2000). *Qué español enseñar*. Madrid: Arco/Libros.

Moreno Fernández, F. (2005). Geografía lingüística de Hispanoamérica. En J. M.^a Enguita, T. Buesa y M.^a A. Martín Zorraquino, *Jornadas Internacionales en memoria de Manuel Alvar* (89-108). Zaragoza: Institución Fernando el Católico.

Moreno Fernández, F. (2009). *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel.

Moreno Fernández, F. y Otero Roth, J. (2008). *Atlas de la lengua española en el mundo*. Barcelona: Ariel-Fundación Telefónica.

Moreno Fernández, F., Moreno Fernández, J. E. y García de las Heras, A. J. (1995). Cálculo de disponibilidad léxica. El programa LexiDisp. *Lingüística*, 7: 243-250.

Mühlhäusler, P. (2000): Humboldt, Whorf and the roots of ecolinguistics. En M. Pütz y M. H. Verspoor, *Explorations in Linguistic Relativity* (pp. 89-99). Amsterdam / Filadelfia: John Benjamins.

Nakajima, S. (1999). Comentarios críticos sobre Varilex 6. *Varilex*, 7: 43-45.

Navarro Carrasco A. I. (1993). Breve panorama de la dialectología. *Anuario de estudios filológicos*, 16: 309-330.

Navarro Carrasco, A. I. (1993). Geografía lingüística y diccionarios. *ELUA. Estudios de Lingüística*, 9: 73-96.

Navarro Gala, R. (2010). *La lengua española en América: normas y usos actuales*. Valencia: Universitat de València.

Navarro Giner, R. M., y Navarro Serrano, P. (2008). Cómo aplicar los niveles de referencia a la elaboración de materiales didácticos: Estudio sobre pasaporte A1, *MarcoELE: Revista de didáctica*, 6 Recuperado de: http://marcoele.com/descargas/6/navarro-navarro-equivalencias_niveles.pdf

Nazar, R. (2008). Seminario Infolex: Aplicaciones de Lingüística Cuantitativa. [Recurso electrónico]. Recuperado de <https://www.iula.upf.edu/materials/080307nazar.pdf>

Nishimura, K. 2002. *Diminutivos españoles. Morfología, unciones semánticas y análisis diatópicas y diastráticas*. [Tesis doctoral]. Universidad de Estudios Extranjeros de Osaka.

Penadés Martínez I. (1999). *La enseñanza de las unidades fraseológicas*. Madrid: Arco/Libros.

Pennycook, A. (2008). English as a language always in translation. *European Journal of English Studies*, 12 (1): 33-47.

Pérez, F. J. (2000). *Diccionarios, discursos etnográficos, universos léxicos : propuestas teóricas para la comprensión cultural de los diccionarios*. Caracas: Universidad Católica Andrés Bello.

Pérez, F. J. (2005). *Pensar y hacer el diccionario: nociones de lexicografía: definiciones, géneros, crítica e historia. Criterios para elaborar diccionarios, métodos de investigación y bibliografía lexicográfica*. Caracas: El Nacional.

Pérez Gómez, Á. I. (2002). Un aprendizaje diverso y relevante. *Cuadernos de pedagogía*, 311: 66-70.

Pérez Hernández, C. (2002). Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. *Estudios de Lingüística del Español*, 18. Recuperado de <http://elies.rediris.es/elies18/>

Pop, S. (1950). *La Dialectologie*. Louvain.

Prodromou, L. (2008). *English as a Lingua Franca: A Corpus-Based Analysis*. London: Continuum.

Quilis, A. (2001). Les travaux de l'Atlas Lingüístico de Hispanoamérica. *Géolinguistique*, 187-195.

Rabanales, A. (1953). *Introducción al estudio del español de Chile. Determinación del concepto de chilenismo*. Santiago de Chile: Editorial Universitaria.

Ramírez, A. G. (1992). *El español de los Estados Unidos: el lenguaje de los hispanos*. Madrid: Mapfre.

Real Academia Española (2005). *Diccionario del estudiante*. Madrid: Santillana.

Real Academia Española (2009). *Nueva Gramática de la lengua española*. Madrid: Santillana.

Real Academia Española (2010). *Ortografía de la lengua española*. Madrid: Espasa.

Real Academia Española (2014). *Diccionario de Lengua Española*. Madrid: Espasa.

Real Academia Española. (2005). *Diccionario panhispánico de dudas*. Madrid: Santillana.

Real Academia Española. Banco de datos (CORPES XXI)[en línea]. *Corpus del español del siglo XXI (CORPES)*. Recuperado de: <http://www.rae.es/recursos/banco-de-datos/corpes-xxi>

Real Academia Española y Asociación de Academias de la Lengua Española (2013). *El buen uso del español*. Madrid: Espasa.

Resnick, M. (1976). *Phonological variants and dialect identification in Latin American Spanish*. Mouton: La Haya.

Reyes Heróles, F. (2005). Español internacional e internacionalización del español. *Revista de la Universidad de México*: 19-22.

Rivarola, J. L. (2006). El español en el siglo XXI: Los desafíos del pluricentrismo. *Boletín Hispánico Helvético*, 8: 97-109.

Rona, J. P. (1969). ¿Qué es un americanismo? En *El Simposio de México. Enero de 1968. Actas, informes y comunicaciones* (pp. 135-148). México: UNAM.

Rona, J. P. (1993). El problema de la división del español americano en zonas dialectales. En F. Moreno Fernández (coord.). *La división dialectal del Español de América* (pp. 63-76). Madrid: Universidad de Alcalá.

Rosenblat, A. (1971). *La lengua del "Quijote"*. Madrid: Gredos.

Rotaetxe, K. (1990). *Sociolingüística*. Madrid: Síntesis.

Rubdy, R., y Saraceni, M. (2006). *English in the World: Global Rules, Global Roles*. London / New York: Continuum.

Rubio, C. y Ueda, H. (1992). *Diccionario Crown japonés—español*. Tokio: Kenkyusha.

Ruiz Tinoco, A. (1999). El Proyecto Varilex en Internet. Base de datos compartida de variación léxica. *Varilex*, 7: 50-60

Ruiz Tinoco, A. (2000). Manual del lenguaje SQL aplicado a la base de datos de Varilex. *Varilex*, 8: 48-70.

Ruiz Tinoco, A. (2001). Cartografía automática en Internet. *Bulletin of the Faculty of Foreign Studies*, 36: 77-93.

Ruiz Tinoco, A. (2004). Léxico variable en los países hispanohablantes según el sexo. En J. Prado Aragonés y V. Galloso Camacho, *Diccionario, léxico y cultura* (pp. 189-198). Huelva: Universidad de Huelva.

-
- Ruiz Tinoco, A. (2008). Los informantes en las encuestas de variación léxica. *Dialectología: Revista Electrónica*, 1: 83-105.
- Salvador, G. (1980). Lexicografía y geografía lingüística. *Revista Española de Lingüística*, X (1): 49-57.
- Salvador, G. (1985). *Semántica y lexicología del español*. Madrid: Paraninfo.
- Sánchez, A., Sarmiento, R., Cantos, P. y Simón, J. (1995). *Cumbre: Corpus lingüístico del español contemporáneo: Fundamentos, metodología y aplicaciones*. Madrid: SGEL.
- Sánchez, R. (1987). *Chicano discourse: Socio-historic perspectives*. Houston: Arte Publico Press.
- Santa-Cecilia, A. G. (2007). Plan Curricular del Instituto Cervantes: Niveles de Referencia para el español. *MarcoELE: Revista de didáctica de Español Lengua Extranjera* (5), 1. Recuperado de <http://marcoele.com/plan-curricular-del-instituto-cervantes-niveles-de-referencia-para-el-espanol/>
- Sapon, S. M. (1962). Contribución metodológica a la encuesta dialectal. En *IX Congreso Internacional de Lingüística Románica* (pp. 69-77). Lisboa: Centro de Estudos Filológicos.
- Serrano Montesinos, M. J. (1999). *Estudios de variación sintáctica*. Madrid: Iberoamericana.
- Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5 (1): 3-55.
- Sierra Bravo, R. (1992). *Técnicas de investigación social*. Madrid: Paraninfo.
- Silva-Corvalán, C. (1980). La función pragmática de la duplicación de pronombres clíticos. *Boletín de Filología de la Universidad de Chile*, XXXI: 561-570.
- Silva-Corvalán, C. (1986). On the problem of meaning in sociolinguistic studies of syntactic variation. *Linguistics across historical and geographical boundaries*, 1, 111-123.
- Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español*. Washington, D.C.: Georgetown University Press.
- Sinclair J. M. (2004). *How to use Corpora in Language Teaching*. Amsterdam: John Benjamins.

Slagter, P. J. (1979). *Un nivel umbral*. Strasbourg: Consejo de Europa.

Slocum, T.A., McMaster, R.B., Kessler, F.C. y Howard, H.H. (2010): *Thematic Cartography and Geovisualization*. New Jersey: Pearson Education International.

Sousa Fernández, X. S. (2010). Entre el atlas lingüístico y el diccionario. Un diccionario de léxico tradicional a partir de los materiales del ALPI. En I. Ahumada, *Metalexigrafía variacional: Diccionarios de regionalismos y diccionarios de especialidad* (pp. 237-256). Málaga: Universidad de Málaga.

Stork, Y. (2006). La situación lexicográfica del español hispanoamericano. Consideraciones desde la perspectiva de la ecolingüística. [Edición digital a partir de *El diccionario como puente entre las lenguas y culturas del mundo: Actas del II Congreso Internacional de Lexicografía Hispánica*]. Recuperado de: <http://www.cervantesvirtual.com/obra/la-situacin-lexicografica-del-espaol-hispanoamericano-consideraciones-desde-la-perspectiva-de-la-ecolingstica-0/>

Takagaki, T. (1993). Hacia la descripción del español contemporáneo de las grandes ciudades del mundo hispánico. *Lingüística Hispánica*, 16: 65-86.

Takagaki, T. (ed.) (2003). *Diccionario de bolsillo español—japonés, japonés—español*. Tokio: Shogakkan.

Tappolet, E. (1895). *Die romanischen Verwandtschaftsnamen*. Estrasburgo.

Terrádez Gurrea, M. (2001). *Frecuencias léxicas del español coloquial: Análisis cuantitativo y cualitativo*. Valencia: Universitat de Valencia.

Thompson, R. W. (1992). Spanish as a pluricentric language. En M. Clyne, *Pluricentric languages. Differing norms in different nations* (pp. 45-70). Berlin/New York: De Gruyter.

Thun, H. (1998). La geolingüística como lingüística variacional general (con ejemplos del *Atlas lingüístico Diatópico y Diastrático del Uruguay*). En G. Ruffino (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza. Sezione 5: Dialettologie, geolingüística, sociolingüística* (pp. 701-729). Tübingen: Niemeyer.

Torres Torres, A. (2001). Bases para el estudio del léxico hispánico de flora, fauna y minería recogido en el *Diccionario de voces americanas* (siglo XVIII) de Manuel José de Ayala. En J. Brumme (ed.), *La historia de los lenguajes iberorrománicos de especialidad: la divulgación de la ciencia* (pp. 195-205). Madrid: Iberoamericana.

Torres Torres, A. (2013). Del castellano de «un pequeño rincón» al español internacional. *Normas: revista de estudios lingüísticos hispánicos*, 3: 205-224.

Tovar, A. (1968). *Lo que sabemos de la lucha de lenguas en la Península Ibérica*. Madrid: Gregorio del Toro.

Ueda, H. (1994). Aproximación translingual a la lengua española. *Revista de Estudios de Adquisición de la Lengua Española*, 1: 157-172.

Ueda, H. (1995a). Variación fonológica en el léxico urbano español. *Lingüística Hispánica*, 18: 105-125.

Ueda, H. (1995b). Zonificación del español. Palabras y cosas de la vida urbana. *Lingüística*, 7: 43-86.

Ueda, H. (1996a). Variación léxica del español urbano. Distribución: Vestuario y equipo. *The Proceedings of the Foreign Language Sections. Graduate School of Arts and Sciences*, vol. I: 41-83. Tokio: Universidad de Tokio. Recuperado de http://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/chiri/vestua2.htm#N_8

Ueda, H. (1996b). Accounting for morpho-syntactic change in Spanish: the Present Perfect Case. *Pennsylvania Working Papers in Linguistics*, 3: 51-62.

Ueda, H. (1996c). Análisis demolingüístico del léxico variable español. *Lingüística Hispánica*, 19: 63-98.

Ueda, H. (1999). Distribución de las palabras variables en España y en América: Léxico del transporte (pp. 637-655). *Estudios de Lingüística Hispánica: Homenaje a María Vaquero*.

Ueda, H. (2003). Tipos de léxico y enseñanza de español como segundo idioma. *Cuadernos Hispanoamericanos*, 631: 27-33.

Ueda, H. (2007a). Análisis dialectométrico del léxico variable español: Interpretación taxonómica de resultados. En *El español de América: Actas del VI Congreso Internacional de "El español de América"* (pp. 813-824). Valladolid: Diputación de Valladolid.

Ueda, H. (2007b). Un estudio comparativo del léxico de Guinea Ecuatorial: Resultado de encuestas y clasificación. En G. Nistal Rosique y G. Pié Jahn, *La situación actual del español en África* (pp. 479-495). Madrid: Sial / Casa de África.

Ueda, H. (2009). Resultados y proyectos en las investigaciones sobre variación léxica del español. *Dialectología: Revista Electrónica*, 2: 51-80. Recuperado de <http://www.raco.cat/index.php/Dialectologia/article/view/198816>

Ueda, H. (2011). Nuevo método para la recogida de datos de variación léxica. Encuestas en web en el proyecto de Varilex. En A. M. Cestero, I. Molina Martos y F. Paredes García, *La Lengua, lugar de encuentro: Actas del XVI Congreso Internacional de ALFAL* (pp. 3733-3740). Alcalá De Henares: Universidad de Alcalá de Henares.

Ueda, H. (2013). Dialectometría en el proyecto «Varilex». Historia y geografía de las voces variantes. *Estudios de Lingüística Galega*, 5: 127-143.

Ueda, H., y Ruiz Tinoco, A. (2003). Varilex. Variación léxica del español en el mundo: Proyecto internacional de investigación léxica. En R. Ávila, J.A. Samper, H. Ueda *et al.*: *Pautas y pistas en el análisis del léxico hispano(americano)* (pp. 141-278). Madrid: Iberoamericana.

Ueda, H., y Takagaki, T. (1993). *Varilex. Variación léxica del español del mundo: Mapas y estadísticas*. Proyecto Internacional de investigación: Español del mundo. Tokio: Universidad de Tokio.

Van Ek, J. A. (1975). *The Threshold Level in a European Unit/Credit System for Modern Language Learning by Adults*. Estrasburgo: Publicaciones del Consejo de Europa.

Vázquez, G. (2008). «¿Qué español enseñar? Preguntas frecuentes». Ponencia presentada en las *III Jornadas de Español como Lengua Extranjera/ I Congreso Internacional de Enseñanza e investigación en ELSE*. Universidad Nacional de Córdoba.

Referencia electrónica:

http://www.lenguas.unc.edu.ar/elsecongreso/teleconferencia_vazquez.pdf

Vila Pujol, M. R. (2009). Dialectos, niveles, estilos y registros en la enseñanza del español como lengua extranjera. *Monográficos MarcoELE*, 8. Recuperado de http://marcoele.com/descargas/expolingua1994_vila.pdf

Villa, D. J. (1996). Choosing a “standard” variety of Spanish for the instruction. *Foreign Language Annals*, 29: 191-200.

Werner, R. (2001). Áreas pendientes en la lexicografía del español americano. En *II Congreso Internacional de la Lengua Española: El Español en la sociedad de la información*, Valladolid. Recuperado de: http://cvc.cervantes.es/obref/congresos/valladolid/ponencias/unidad_diversidad_del_espanol/2_el_espanol_de_america/werner_r.htm

Werner, R. (2002). Identidad nacional y regional de las palabras en los diccionarios españoles. En M.^a T. Fuentes Morán y R. Werner (eds.) *Diccionarios: textos con pasado y futuro* (pp. 75-88). Madrid: Vervuert.

Wolf, L., Ettinger, S., Werner, R. y Haensch, G. (1982). *La lexicografía: de la lingüística teórica a la lexicografía práctica*. Madrid: Gredos.

Wood, G. R. (1989). Using a printed vocabulary checklist. *Journal of English Linguistics*, 22 (1): 1-5.

Zamora, J. y Guitart, J. (1988). *Dialectología hispanoamericana*. Salamanca: Almar.

Zauner, A. (1903). Die romanischen Namen der Körperteile. *Romanische Forschungen*, 14 (1): 339-530.

Zimmerman, L. J. (2003). *Indios americanos. Las primeras naciones: Vida, mitología y arte de los indios norteamericanos*. Madrid: Jaguar.

Zimmermann, K. (2003). El fin de los diccionarios de mexicanismos, colombianismos, argentinismos, cubanismos etc. La situación de la lexicografía del español de América después de la publicación de los Diccionarios contrastivos del español de América: Español de América – Español de España, dirigidos por Günther Haensch y Reinhold Werner Madrid: Gredos 2000. *Revista Internacional de Lingüística Iberoamericana*, I (1): 71-83.

ANEXOS

ANEXO I

Fuentes de consulta sobre la variación léxica americana

OBRAS LEXICOGRÁFICAS REGIONALES SOBRE EL ESPAÑOL AMERICANO	
País	Obras lexicográficas destacadas
Argentina	<p>Abad de Santillán, D. 1991. <i>Diccionario de argentinismos de ayer y de hoy</i>. Buenos Aires: Tipográfica Editora Argentina.</p> <p>Haensch, G. y Werner, R. (eds.). 2000. <i>Diccionario el español de Argentina. Español de Argentina / Español de España</i> (coord. C. Chuchuy). Madrid: Gredos.</p> <p>Tejera, M. J., y de Letras, A. A. 2004. <i>Diccionario del habla de los argentinos</i>. Buenos Aires: Espasa.</p> <p>Verdevoye, P. (coord. científico) y Colla, H. F (coord. técnico). 1992. <i>Léxico argentino-español-francés/Lexique argentin-espagnol-français</i>. Madrid/Paris: Colección Archivos, n.º 47.</p>
Chile	<p>Brennan, J. y Taboada, A. <i>How to Survive the Chilean Jungle</i>. Santiago: CL: Dolmen Ediciones.</p> <p>Carrillo Herrera, Gastón. 1981. "Diccionario del Habla Chilena", <i>Romanische Forschungen</i> 93 1/2: 216-222.</p> <p>Morales Pettorino, F., Quiroz Mejías, O., Peña Álvarez. <i>Diccionario ejemplificado de chilenismos y de otros usos diferenciales del español de Chile</i>. Valparaíso: Academia Superior de Ciencias Pedagógicas de Valparaíso.</p> <p>Subercaseaux, M. 1986. <i>Diccionario de chilenismos</i>. Santiago de Chile: Editorial Juvenil.</p>
Colombia	<p>Academia Colombiana, Comisión de lexicografía. 1975. <i>Breve diccionario de colombianismos</i>. Bogotá.</p> <p>Di Filippo, M. A. 1964. <i>Lexicón de colombianismos</i>. Cartagena: Editora Bolívar.</p>
Costa Rica	<p>Quesada Pacheco, M. A. 1995. <i>Diccionario histórico del español de Costa Rica</i>. San José: Euned.</p> <p>Quesada Pacheco, M. A. 2007. <i>Nuevo diccionario de costarrriqueñismos</i>. Cartago: Editorial Tecnológica de CR.</p>
Cuba	<p>Cárdenas Molina, G., Tristán Pérez, A. M.^a, Werner, R. (coords.). 2000. <i>Diccionario del Español de Cuba. español de Cuba –Español de España</i>. Madrid: Gredos.</p>

	<p>Ortiz, F. 1985. <i>Nuevo catauro de cubanismos</i>. La Habana: Editorial de Ciencias Sociales.</p> <p>Pichardo, E. 1861. <i>Diccionario provincial de voces cubanas</i>. Imprenta del Gobierno.</p> <p>Santiesteban, A. 1982. <i>El habla popular cubana de hoy</i>. La Habana: Editorial Ciencias Sociales.</p>
España	<p>Corrales, C., y Corbella, D. 2010. <i>Tesoro léxico canario-americano</i>. Las Palmas de Gran Canaria: Departamento de Ediciones de la Consejería de Cultura y Patrimonio Histórico y Cultural.</p>
Guatemala	<p>Arriola, J. L. 1954. <i>Pequeño diccionario etimológico de voces guatemaltecas</i>. Guatemala: Ministerio de Educación Pública.</p>
Honduras	<p>Membreño, A. 1982. <i>Hondureñismos Vocabulario de los Provincialismos de Honduras con "Hondureñismos" de Jeremías Cisneros</i>. Guaymuras (Tegucigalpa): Lámpara crítica y cultura.</p>
México	<p>Cobos, R. 1983. <i>A dictionary of New Mexico and southern Colorado Spanish</i>. Santa Fe: Museum of New Mexico Press.</p> <p>Lara, L. F. 1986. <i>Diccionario básico del español de México</i>. México: El Colegio de México.</p> <p>Lara, L. F., Bracho, C., Colegio de México y Fondo de Cultura Económica. 1982. <i>Diccionario fundamental del español de México</i>. México: Colegio de México.</p> <p>Lara, L. F. 2000. <i>Diccionario del español usual en México</i>. Alicante: Biblioteca Virtual Miguel de Cervantes.</p> <p>Santamaría, F. J. 1959. <i>Diccionario de mejicanismos</i>. Madrid: Porrúa.</p>
Nicaragua	<p>Academia Nicaragüense de la Lengua. 2001. <i>Diccionario de uso del español nicaragüense</i>. Managua: Academia Nicaragüense de la Lengua.</p> <p>Arellano Oviedo, F. 2007. <i>Diccionario del español de Nicaragua</i>. Managua: Pavsa.</p> <p>Gulden, C. M. Van der, 1995. <i>Vocabulario nicaragüense</i>. Managua: UCA.</p>
Panamá	<p>Aguilera Patiño, L. 1951. "Diccionario de panameñismos". <i>Boletín de la Academia Argentina de la Lengua</i> 20: 405-506.</p> <p>Isaza Calderón, B., y Alfaro, R. J. 1968. <i>Panameñismos</i> Ciudad de Panamá: Impresora Panamá. SA,</p>
Perú	<p>Álvarez Vita, J. 1990. <i>Diccionario de peruanismos</i>. Lima: Studium.</p> <p>Arona, J. de: <i>Diccionario de peruanismos</i>. 1938. París: Desclée de Brouwer.</p>

	<p>Foley Gambetta, E. 1983. <i>Léxico del Perú. Diccionario de peruanismos, replana criolla, jerga del hampa, regionalismos y provincialismos del Perú</i>. Lima: Jahnos.</p> <p>Hildebrandt, M. 1969. <i>Peruanismos</i>. Moncloa-Campodónico.</p>
Puerto Rico	<p>Malaret, A. 1999. <i>Vocabulario de Puerto Rico</i>. Madrid: Arco/Libros.</p> <p>Vaquero de Ramírez, M. T., y Morales de Walters, A. 2005. <i>Tesoro lexicográfico del español de Puerto Rico</i>. San Juan: Academia Puertorriqueña de la Lengua Española.</p>
República Dominicana	<p>Deive, C. E. 2002. <i>Diccionario de dominicanismos</i>. Santo Domingo: Editora Manati.</p> <p>Demorizi, E. R. 1983. <i>Del vocabulario dominicano</i>. Santo Domingo: Editora Taller.</p>
Uruguay	<p>Kühl de Mones, U. K. 1993. <i>Nuevo diccionario de uruguayismos</i>. Santa Fé de Bogotá: Instituto Caro y Cuervo.</p>
Venezuela	<p>Núñez, R., y Pérez, F. J. 1994. <i>Diccionario del habla actual de Venezuela: Venezolanismos, voces indígenas, nuevas acepciones</i>. Caracas: Universidad Católica Andrés Bello, Centro de Investigación Lingüística y Literaria.</p> <p>Tejera, M. J. 1983. <i>Diccionario de venezolanismos</i>. Caracas: Academia Venezolana de la Lengua: Universidad Central de Venezuela.</p>

DICCIONARIOS GENERALES DEL ESPAÑOL AMERICANO

- Arias de la Cruz, M. A.. 1980. *Diccionario temático, americanismos*. León: Everest.
- Asociación de Academias de la Lengua Española. 2010. *Diccionario de americanismos*. Lima: Santillana Ediciones Generales.
- Ayala, M. J. 1995. *Diccionario de voces americanas*. Madrid: Arco/Libros.
- Malaret, A. 1946. *Diccionario de americanismos*. Buenos Aires: Mayagüez.
- Morínigo, M. A. 1996. *Diccionario del español de América*. Buenos Aires: Anaya.
- Neves, A. N. 1973. *Diccionario de americanismos*. Buenos Aires: Editorial Sopena Argentina.
- Renaud, R. 1997. *Diccionario de hispanoamericanismos: no recogidos por la Real Academia. Formas homónimas, polisémicas y otras derivaciones*. Madrid: Cátedra.
- Santamaría, F. J. 1942. *Diccionario general de americanismos*. México: Pedro Robredo.
- Sopena, R. 1983. *Diccionario ilustrado de americanismos*. Barcelona: Muchnik.
- Steel, B. 1990. *Diccionario de americanismos: ABC of Latin American Spanish*. Madrid: SGEL.
- Steel, B. 1999. *Breve diccionario ejemplificado de americanismos*. Madrid: Arco/Libros.

ANEXO II

Conceptos de la base de datos de Varilex estudiados en esta investigación

Concepto	Definición
1. ABUNDANT	De gran cantidad. Ej. Este país es *** en petróleo.
2. AMUSEMENT	Cosa que divierte. Ej. Aquéllos eran lindos tiempos. No había tantas (tantos) ***(e)s como ahora.
3. ANGER	Enfado grande y violento. Ej. En verdad me da *** que haya quienes piensan que invirtiendo en quince días, se puede exportar y ganar divisas.
4. ANGRY	Con enfado o disgustado. Ej. El jefe está *** porque ustedes llegaron diez minutos tarde.
5. ANGRY	Que tiene un disgusto muy fuerte. Ej. El profesor está muy ***, porque unos estudiantes han faltado a su clase.
6. ANNOYING	Que causa disgusto. Ej. ¡Qué *** es la vida... no hay justicia!
7. APERITIF	Copa de bebida alcohólica que se toma antes de las comidas. Ej. Oye, ¿vamos a tomar un (***) antes de volver a casa?
8. APRON	Prenda de vestir que se coloca de cintura para abajo en la parte delantera del cuerpo, con el fin de proteger la ropa de manchas o suciedad. Se usa en la cocina, en cierto tipo de trabajos, etc.
9. ARROGANCE	Orgullo excesivo o sentimiento de superioridad frente a los demás. Ej. Su *** le impide aceptar las observaciones de los demás sobre su trabajo.
10. ARROGANT	Que siente un orgullo excesivo o sentimiento de superioridad. Ej. Es muy *** con los inferiores.
11. ATTRACTIVE (MAN)	Un hombre muy atractivo. Ej. El actor de la película es ***.
12. ATTRACTIVE (MAN)	Muy atractivo: El actor de la película está ***.
13. AUTOMATIC CASHIER	Máquina automática donde se hace la cuenta bancaria y se puede sacar cierta cantidad de dinero.
14. AVARICE	Deseo excesivo de dinero, poder o riquezas. Ej. El (La) *** no le dejaba ser feliz.

15. BABY	¿Cómo llama al niño recién nacido? Ej. La madre tomó a su *** en brazos.
16. BALL-POINT PEN	Utensilio para escribir con una bolita por donde fluye la tinta para escribir.
17. BANANA	Fruto comestible, alargado, cubierto por una corteza lisa y amarilla. Ej. José resbaló en una piel de ***.
18. BAR	Barra donde se sirve café, cerveza, etc.
19. BAR	Establecimiento donde se sirven café y otras bebidas y comidas ligeras. Ej. Hoy no puedo ir a casa y comeré algo en cualquier ***.
20. BAR	Establecimiento de carácter popular en el que se venden y consumen bebidas alcohólicas. Ej. Todas las noches, el marido volvía borracho del (la) ***.
21. BAR	Local en el que se sirven cerveza y comidas ligeras. Ej. Vamos a entrar a este(a) (***) a tomarnos una cerveza.
22. BATHTUB	Pila de baño.
23. BE BORN LUCKY.	Tener suerte: A María todo le sale bien. Además es inteligente y guapa. (***)).
24. BE FALLING ASLEEP.	Tener mucho sueño: Llevo cerca de dos días sin dormir y ahora (***)).
25. BE INSEPARABLE.	Entenderse o avenirse muy bien con alguien: Agustín y José María se conocen desde que eran niños. (***)).
26. BEAUTIFUL	Bello y agradable a la vista. Ej. Desde aquí se ve un paisaje muy ***.
27. BEDROOM	Habitación de una casa, no de un hotel, en la que se duerme.
28. BEER	¿Qué otro nombre tiene la cerveza? Ej. No tengo tanta prisa y vamos a tomarnos un (una) ***.
29. BITTER	De sabor fuerte y desagradable al paladar, como la hiel. Ej. Ana prueba a sorbos pequeños el café y un gusto *** se queda en la garganta.
30. BLACKBOARD	Planta que se emplea para escribir.
31. BLOCK	Cuaderno con espiral de alambre.

32. BLOCK	Espacio cuadrado limitado por calles y destinado a la construcción de edificios. Ej. Vivían en la segunda (el segundo) *** de la calle C.
33. BORED	Que está cansado de no hacer nada. Ej. Está lloviendo y no ha salido en todo el día. Está en su casa ***.
34. BOSS	El superior en el trabajo. Ej. El *** ha dicho que hay que terminar el trabajo antes de las 5.
35. BRASSIERE	Prenda interior femenina que se utiliza para sujetar los senos.
36. BROTHER	¿Cómo le llama a su hermano? Ej. Juan está con su *** en el parque.
37. BUS	Vehículo de servicio público que sigue un trayecto y paradas fijas, especialmente en las ciudades.
38. BUTTER	Grasa comestible que se saca de la nata de la leche. Ej. Las tostadas me gustan con *** y mermelada.
39. CALL	Comunicación a través del teléfono. Ej. Esperaba un (una) *** importante.
40. CALL THE ROLL	Comprobar la presencia de los alumnos llamándoles por su nombre. Ej: El profesor *** al comenzar la clase.
41. CAN	Envase de refresco.
42. CAN	Envase en que se guarda por ejemplo, gasolina, querosene, etc.
43. CANNOT REMEMBER	No poder recordar algo. Ej. ¿Dónde puse el diccionario? ***
44. CAR	Vehículo destinado al transporte de personas.
45. CELLOPHANE TAPE	Cinta transparente, con una sustancia adhesiva en una de sus caras, que se emplea, por ejemplo, para fijar papel a otra superficie. Ej. ¿Tienes un poco de (***) para pegar este papel en el tablón?
46. CENTER HALF	En el fútbol, jugador que integra la línea media del equipo y que se sitúa en el centro del campo de juego. Ej. Pedro juega muy bien con las dos piernas, y como es muy rápido, el entrenador lo ha puesto de (***)).

47. CENTER HALF	Jugador que juega en el centro del campo formando con otros la línea central del equipo: El equipo lo que necesita es a alguien que se mueva bien por el centro del campo, un ***.
48. CHAT	Conversación que no tiene un fin determinado, sobre temas normalmente poco importantes. Ej. Tuve un (una) *** larga (largo) con el vecino.
49. CLOSE FRIENDO	Entenderse o avenirse muy bien con alguien: Agustín y José María se conocen desde que eran niños. (***)
50. COARSE	Que tiene malos modos; que es poco delicado en el trato y en el comportamiento. Ej. Cuando se enfada se vuelve un poco ***, pero no es mala persona.
51. COFFIN	Caja de madera donde se pone el cadáver de una persona para enterrarla. Ej. Cuando don Felipe murió, sus amigos llevaron el (la) (***) al cementerio.
52. COLANDER	Utensilio de cocina para coger el objeto dejando pasar los líquidos.
53. COLD	Con un malestar físico que se produce generalmente por cambios bruscos de temperatura. Ej. Estornudo tanto porque tengo un (una) ***.
54. COMPETENT	El que sabe mucho y acierta en todo lo que hace. Ej. José es muy *** y nunca se equivoca.
55. COMPUTER	Máquina para procesar los datos de información.
56. COOK	Colocar un alimento crudo en un líquido en ebullición para hacerlo comestible. Ej. Si no te gustan las salchichas crudas, puedes (***)las antes.
57. COOKER (STOVE)	Aparato donde se cocina.
58. COURAGE	Cualidad de valiente. Ej. Es un combatiente con mucho (mucho) ***.
59. DEGREE	Grado académico que permite ejercer una profesión al terminar los estudios en la universidad. Ej. Consiguió su *** en Medicina en sólo cinco años.
60. DIARY	Agenda con hojas cambiables y tapas gruesas.
61. DIE	¿Qué dice cuando una persona deja de existir? Ej. Dionisio estaba enfermo de corazón y ***.

62. DIE	Fallecer: Doña Ana (***) después de una larga enfermedad.
63. DISHWASHER	Máquina para lavar platos.
64. DISMISS	Quitar a una persona el empleo. Ej. Fueron *** de sus cargos todos los funcionarios comprometidos en el caso.
65. DODGE THE ISSUE.	Eludir o evitar un riesgo o compromiso: Siempre que hay algún problema, Eladio intenta (***)
66. DRIVING LICENCE	Documento oficial que autoriza a su poseedor a conducir un vehículo de motor. Ej. Saqué el (la) (***) a los dieciocho años pero todavía no tengo coche.
67. EARRINGS	Adorno que se ponen las mujeres en las orejas.
68. EAT LIKE A HORSE.	Comer mucho: No sé cómo no revientan de tanto comer. Estos muchachos (***)
69. EAT TOO MUCH	Hartarse de comer: El domingo, después del partido se fueron todos a comer chuletas y ***
70. EXPEL	Echar a una persona. Ej. Le *** del cuerpo de policía porque era un ladrón.
71. FALL IN LOVE	Sentir amor hacia una persona. Ej. Teresa es tan simpática que cualquier chico *** de ella.
72. FAT (PLUMP)	Muy gordo. Ej. Pedro es (un) *** porque come mucho.
73. FATHER (DADDY)	¿Cómo le llama a su padre? Ej. *** está trabajando.
74. FLAT (US: APARTMENT)	Vivienda en un edificio de varias plantas.
75. FOLDER	Cubierta doblada de cartón que sirve para archivar papeles.
76. FOOD	Producto que sirve como alimento. Ej. El (La) *** que más ha subido este último tiempo es el arroz.
77. FOOTBALL	Fútbol: Aquí el *** es el deporte que más se practica.
78. FORWARD	Jugador que forma con otros la línea más adelantada de un equipo: Si corriera un poco más, me gustaría jugar de ***
79. GET ANGRY	Enojarse, enfadarse: Cuando supo la verdad, *** mucho.
80. GIVE A FAILING GRADE	Dar la pésima nota. Ej. Hoy recibimos la nota del último examen y me ***.

81.	GLASSES	Anteojos usados corrientemente que se apoyan en la nariz y en las orejas. (No son de color oscuro para el sol.)
82.	GO TO ENORMOUS EXPENSE.	Gastar demasiado dinero, generalmente para celebrar algo: Don Pedro (***) cuando se casó su hija menor.
83.	GOAL	Sitio donde se mete la bola en fútbol.
84.	GOALKEEPER	Jugador que se coloca en la meta para evitar que entre la pelota: Es el mejor *** que he visto. Ha parado los dos penaltis.
85.	GOOD STUDENT	El estudiante que siempre tiene buenas notas (calificaciones) y sabe mucho. Ej: Pedro es *** y sus padres están muy contentos.
86.	GRADUATE	Persona que ha conseguido el grado académico que permite ejercer una profesión. Ej. Es *** en Derecho y actualmente es abogado.
87.	GRADUATE	Recibir un grado en una universidad. Ej. Mi hija está a punto de *** una universidad norteamericana.
88.	GRIND	El que siempre está estudiando (con el sentido peyorativo). Ej. Javier es *** y no juega mucho con sus compañeros.
89.	HAIR	El cabello. (coloquial). Ej. Mira cómo se me ha puesto el (la, las) *** con el viento.
90.	HANDBAG	Bolsa de cuero, tela o similar que llevan las mujeres con sus cosas personales cuando salen de casa.
91.	HAVE PULL.	Gozar de la protección de alguien: El jefe no se atreve a llamarle la atención, porque (***)
92.	HEAD	La cabeza. (coloquial). Ej. Se ha hecho daño en el (la) ***.
93.	HOSPITAL	Establecimiento, generalmente público, con camas, personas y medios para que los enfermos o heridos reciban atención médica. Ej. Tuvo un accidente de tráfico y lo llevaron al ***.
94.	HUNGER	Gana y necesidad de comer. Ej. Con el aumento de los precios se colocó a los trabajadores al borde del (la) ***.
95.	IRON	Metal duro, de color gris, que abunda en la naturaleza y que sirve para hacer todo tipo de herramientas, estructuras y objetos. Ej. El *** se oxida fácilmente.

96. KICK THE BUCKET.	Morir (coloquial): Ese desgraciado por fin (***)
97. LAZY	Persona a la que no le gusta trabajar. Ej. No debes ser ***, es mejor que hagas las tareas cuanto antes.
98. LAZY STUDENT	El que no estudia mucho. Ej. Alfonso es *** y sus padres están preocupados.
99. LIFT (US: ELEVATOR)	Aparato en forma de cajón para llevar gente de un piso a otro.
100. LIVING ROOM	Sala para estar cómodo, para ver la televisión, etc.
101. LONG DISTANCE CALL	Llamada telefónica que se hace a otra ciudad o a otro país. Ej. No puedo llamar muchas veces a mi novia que vive en el extranjero porque las (***) (s) cuestan muy caras.
102. MAKE FLY OFF THE HANDLE.	Enojarse a alguien: José, con sus impertinencias, siempre (***) a Felisa.
103. MARK (US: GRADE)	Nota que se pone a un ejercicio. Ej. El (La) *** de la preguntas me da para pasar al siguiente examen.
104. MEEK	El que está siempre callado y no trata con nadie. Ej. Pablito es *** y siempre está solo en la escuela.
105. MONEY	Dinero. Ej. José tiene mucho(a) *** en el banco.
106. MOSQUITO	Insecto de cuerpo fino, dos alas transparentes y patas alargadas que se alimenta de sangre. Ej. Tengo el brazo lleno de picaduras de ***s.
107. MOTHER (MUMMY)	¿Cómo le llama a su madre? Ej. *** está siempre ocupada en la cocina.
108. NOT ATTRACTIVE	El que no es nada atractivo. Ej. El pobre hombre es (un) ***, pero tiene muy buen corazón.
109. NOTE (US: BILL)	Papel de tamaño pequeño que sirve de pasaje de un tren.
110. NUMBER PLATE (US: LICENSE PLATE)	Placa visible con el número de registro oficial.
111. OLD BOY	Forma de llamar a un amigo. Ej. Tú me comprenderás, ***.
112. OLD MAN	¿Cómo llama a un anciano de 80 años? Ej. Los ***(s, es) están tomando el sol en el parque.

113. OVERALLS	Traje usado por pintores, mecánicos, etc., que consiste en un pantalón con peto y mangas en una sola pieza.
114. OVERCOAT	Prenda de vestir femenina que se usa en invierno y se pone sobre todas las demás para ir por la calle.
115. PANTIES	Prenda interior femenina que recubre la parte inferior de la cintura.
116. PANTS (US: SHORTS)	Ropa interior masculina.
117. PANTY HOSE	Combinación de panty y medias.
118. PAY	El dinero mensual que recibe por su trabajo. Ej. Su *** no es suficiente para mantener la familia.
119. PEN	Utensilio para escribir con tinta.
120. PHOTOCOPIER	Máquina para hacer fotocopias.
121. PICK A GIRL UP (COL.)	Establecer una relación amorosa sin llegar a un noviazgo formal. Ej. Como no tiene novia, los fines de semana se va a las discotecas con sus amigos a (***)
122. PIG	Animal mamífero doméstico, bajo, grueso, de patas cortas cuya carne aprovecha el hombre. Ej. No le gusta la carne de *** porque tiene mucha grasa.
123. POLICEMAN	Persona que se dedica a vigilar el mantenimiento del orden público y la seguridad de los ciudadanos, así como el cumplimiento de las leyes. (No en forma vulgar). Ej. Dos ***s siguieron al sospechoso hasta su casa.
124. POOR	Sin dinero. Ej. Mi hermano no es un *** como ustedes creen.
125. POOR STUDENT	El estudiante que siempre tiene malas notas. Ej. Germán es *** y los maestros le regañan con frecuencia.
126. POSITION	Posición en un lugar determinado. Ej. El edificio tiene un (una) *** privilegiada en el centro de la ciudad.
127. POSTER	Anuncio escrito de un evento, colocado en lugares públicos.
128. POT	Vasija para guisar.
129. POTATO	Tubérculo comestible, de forma redonda o alargada y de color marrón por fuera y blanco o amarillo por dentro, que se usa como alimento. Ej. Peló los (las) ***s y las (los) lavó antes de cortarlas y freírlas.

130. PROPELLING PENCIL	Utensilio para escribir con minas cambiables.
131. PUFF	Acción de inhalar y exhalar una vez el humo del cigarrillo. Ej. Sólo dio un par de ***s y luego tiró el cigarrillo.
132. PUFF	Absorber el humo del cigarrillo. Ej. *** hasta calentar el cigarrillo y se mareó.
133. PUT ON	Colocar o ajustar en el cuerpo. Ej. Con el frío que hacía convenía *** toda la ropa posible.
134. PUT SOMEONE ON THE CARPET.	Reprender a alguien: Miguel nunca hace caso de lo que le dicen, y le tuve que (***) .
135. RADIANT	Que expresa gozo o alegría. Ej. Yo lo veía avanzar *** para recoger el premio.
136. REFRIGERATOR	Aparato que mantiene alimentos a baja temperatura.
137. RELATION	Unión, correspondencia o dependencia de uno con otro. Ej. Detengamos nuestra atención en aquellos que más *** tienen con la novela.
138. REPEATER	El que ha obtenido muy malas notas y no ha podido pasar al grado superior. Ej: En esta clase hay dos ***(s, es) que no pudieron aprobar el curso.
139. RESIGNATION	Acción de dejar de cumplir un cargo por voluntad propia. Ej. También la Junta Directiva tiene la facultad de proponerle al Presidente el (la) *** del Rector.
140. RICH	Con mucho dinero. Ej. Es *** y puede vivir sin trabajar.
141. RING	Aro con una piedra que se pone en los dedos de la mano como adorno.
142. RING	Aro sin piedra que se pone en los dedos de la mano como adorno.
143. RING	Aro que se pone en los dedos de la mano como adorno.
144. ROAD	Camino amplio y pavimentado destinado a la circulación rodada.
145. SAD	Triste: Con la noticia del accidente de su primo, se quedó ***.
146. SANDWICH	Pan que lleva jamón, queso, etc. adentro.

147. SANDWICH	Conjunto de dos rebanadas cuadradas de pan, entre las cuales se pone un alimento. Ej. El niño se está comiendo un (una) *** de jamón y queso.
148. SCHOLARSHIP HOLDER	Persona que disfruta de una beca. Ej. Soy *** del intercambio cultural.
149. SCHOOL	Centro de enseñanza. Ej. Los niños reciben la enseñanza en el (la) ***.
150. SEARCH	Mirar o examinar con cuidado para encontrar lo que pueda estar oculto. Ej. La policía *** las casas del barrio en busca de los secuestradores.
151. SHAWL	Prenda de vestir femenina bordeada de flecos, que llevan las mujeres sobre los hombros.
152. SHOP (US: STORE))	Establecimiento en el que se venden comestibles, ropas u otros productos de consumo. Ej. Isabel trabaja en un (una) *** de ropa.
153. SHORT (PERSON)	Muy baja. Ej. Andrés es ***, pero juega bien al baloncesto.
154. SKIN	Tejido exterior que cubre el cuerpo del hombre. Ej. Ángela es pelirroja y tiene el (la) *** cubierta de pecas.
155. SKIRT	Prenda de vestir femenina que cubre de la cintura para abajo.
156. SLAP	Golpe dado en la cara con la mano abierta sobre la cara. Ej. Cuando le oyó esa respuesta, le dio un (una) ***; (*) No expreso; (#) No sé.
157. SOCKS	Prenda de punto que recubre el pie.
158. SPANIARD	Persona de origen español. Ej. Es una ciudad llena de ***(e)s, italianos e ingleses.
159. SPONGE CAKE	Pasta de repostería, muy esponjosa de color amarillo y en la parte superior, de color marrón (café).
160. STAMP	Stampita emitida por una administración postal al servicio de correos.
161. STANDING	Levantado, de posición vertical. Ej. El autobús estaba muy lleno y yo tenía que estar ***.
162. STAPLER	Utensilio que sirve para grapar papeles.
163. STEWARDESS	En un avión, mujer que atiende a los pasajeros.

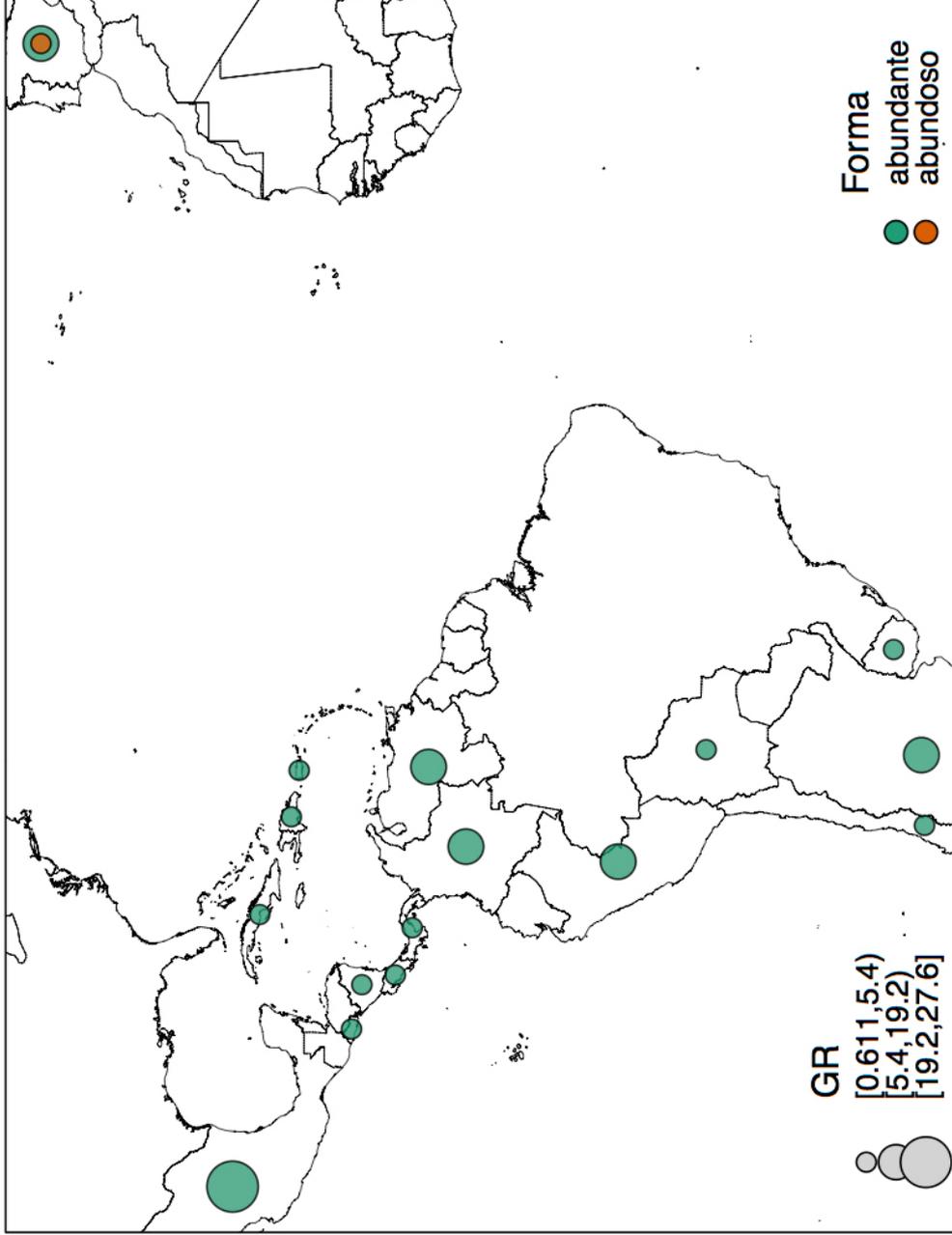
164. STONE (US: PIT)	Parte dura e interior de algunos frutos, como el melocotón, la ciruela, etc.
165. STORM	La lluvia fuerte y con viento. Ej. El domingo no salimos al campo por el (la) ***.
166. STRANGER	La persona que es o viene de otro lugar. Ej. Cuando un *** entró en el bar, todos lo miraron.
167. STRANGER	El que no es de aquí. Ej. Renato es ***, por eso no sabe costumbres de aquí.
168. SUFFER FROM A HANGOVER.	Sentir malestar al día siguiente de haber bebido con exceso: Ayer estuvo bebiendo con sus amigos y hoy (***)
169. SUITCASE	Bolso grande hecho de material duro, con asas para ser transportado a mano, que se usa en los viajes.
170. SWEATER	Prenda de vestir de tejido de punto, con mangas que cubre aproximadamente hasta la cintura.
171. SWIMMING POOL	Sitio donde se practica la natación.
172. SWIMSUIT	Traje de baño femenino.
173. SWITCH	Pieza que se pulsa para accionar un dispositivo.
174. TAKE OFF	Quitarse prendas de vestir. Ej. El marido ayudó a su esposa a *** el abrigo.
175. TAKE OFF	Levantar el vuelo un avión. Ej. La azafata advirtió a los pasajeros que no fumaran antes de (***) el avión.
176. TALKATIVE	Que habla mucho. Ej. Este hombre es muy *** _ y no para de hablar.
177. TALL (PERSON)	Muy alto. Ej. Diego es (un) ***. Es el más alto de la clase.
178. TAP (US: FAUCET)	Dispositivo para abrir o cerrar el paso del agua.
179. TEACHER	Persona que enseña las primeras letras. Ej. Es *** y da clase a niños de cinco años.
180. THIN	De poca grasa y carne. Ej. Tras la enfermedad se ha quedado muy *** y todos los pantalones le están grandes.
181. THIN (SKINNY)	Muy delgado (flaco). Ej. Juanito es ***, como su madre.
182. TIDBITS	Alimento ligero que se sirve como acompañamiento de una bebida. Ej. En ese bar se sirven ricas (ricos) *** todos los días.

183. TIRED	Que se siente sin energía. Ej. Hoy estoy *** y quiero acostarme pronto.
184. TOOTHBRUSH	Cepillo que se utiliza para limpiarse los dientes.
185. TOOTHPASTE	Pasta de dientes.
186. TREMENDOUS	Que sobresale entre lo demás y provoca admiración. Ej. ¡Es un chico ***! Sabe de todo.
187. TROUBLE	Alteración del ánimo, a causa de un enfado. Ej. Entonces, el (la) *** es así: solicitud denegada. No hay dónde ir.
188. UNEMPLOYMENT	Falta de empleo o trabajo. Ej. Es necesario acabar con el (la) ***, si perseguimos la igualdad.
189. VALISE	Estuche para llevar documentos, utilizado por hombres de negocio.
190. WAITER	Persona que se dedica a servir comidas o bebidas en un establecimiento. Ej. Trabaja como *** en un restaurante.
191. WARDROBE	Armario donde se guardan ropas.
192. WASHBASIN	Lugar donde se lavan las manos.
193. WASHER	Aparato eléctrico que sirve para lavar ropa.
194. WHEEL	Unidad rodante completo de un automóvil.
195. WHIP	Instrumento largo y flexible que sirve para castigar o hacer ruido para que los animales se muevan o realicen un trabajo. Ej. Su padre andaba con un *** enorme.
196. WINDBREAKER	Chaqueta fuerte y cómoda que se lleva para hacer actividades fuera de casa.
197. WORKER	Persona que se dedica a hacer un trabajo físico. Ej. Los ***s de la construcción tienen que llevar casco.
198. YOUNG ADULT	Un joven, por ejemplo, de 18 años. Ej. Los ***s no tienen intereses verdaderos.

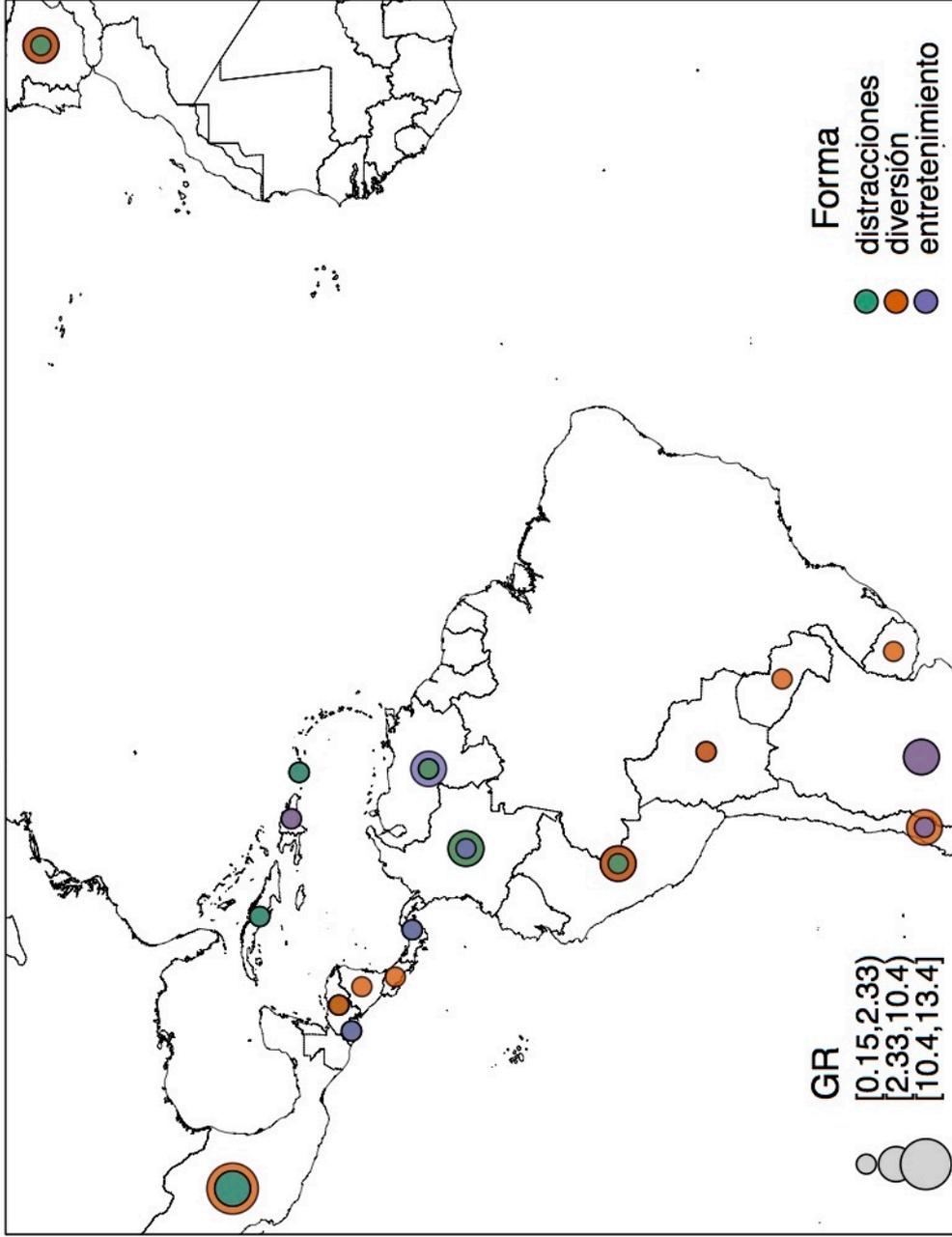
ANEXO III

Mapas estáticos de los conceptos estudiados

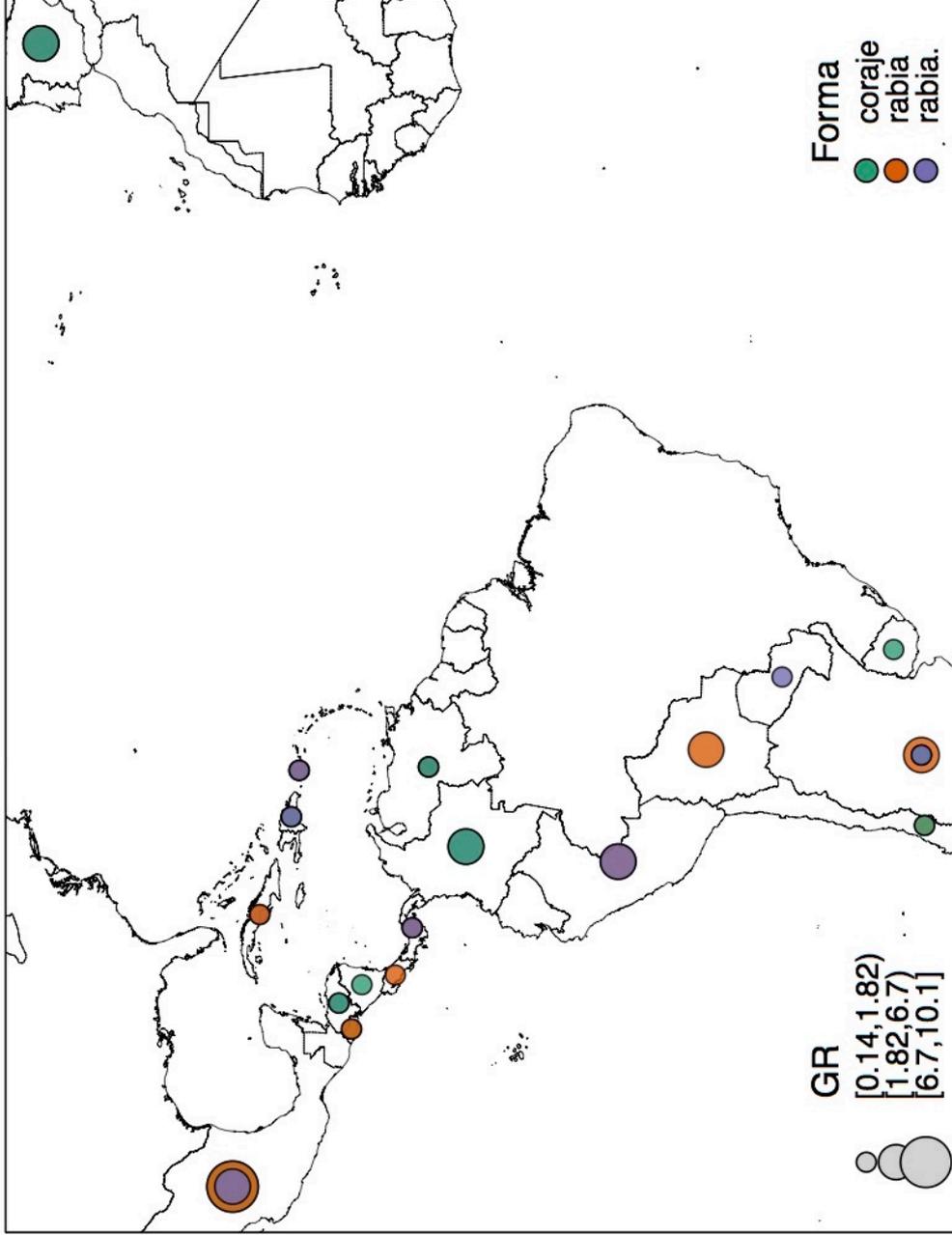
ABUNDANT



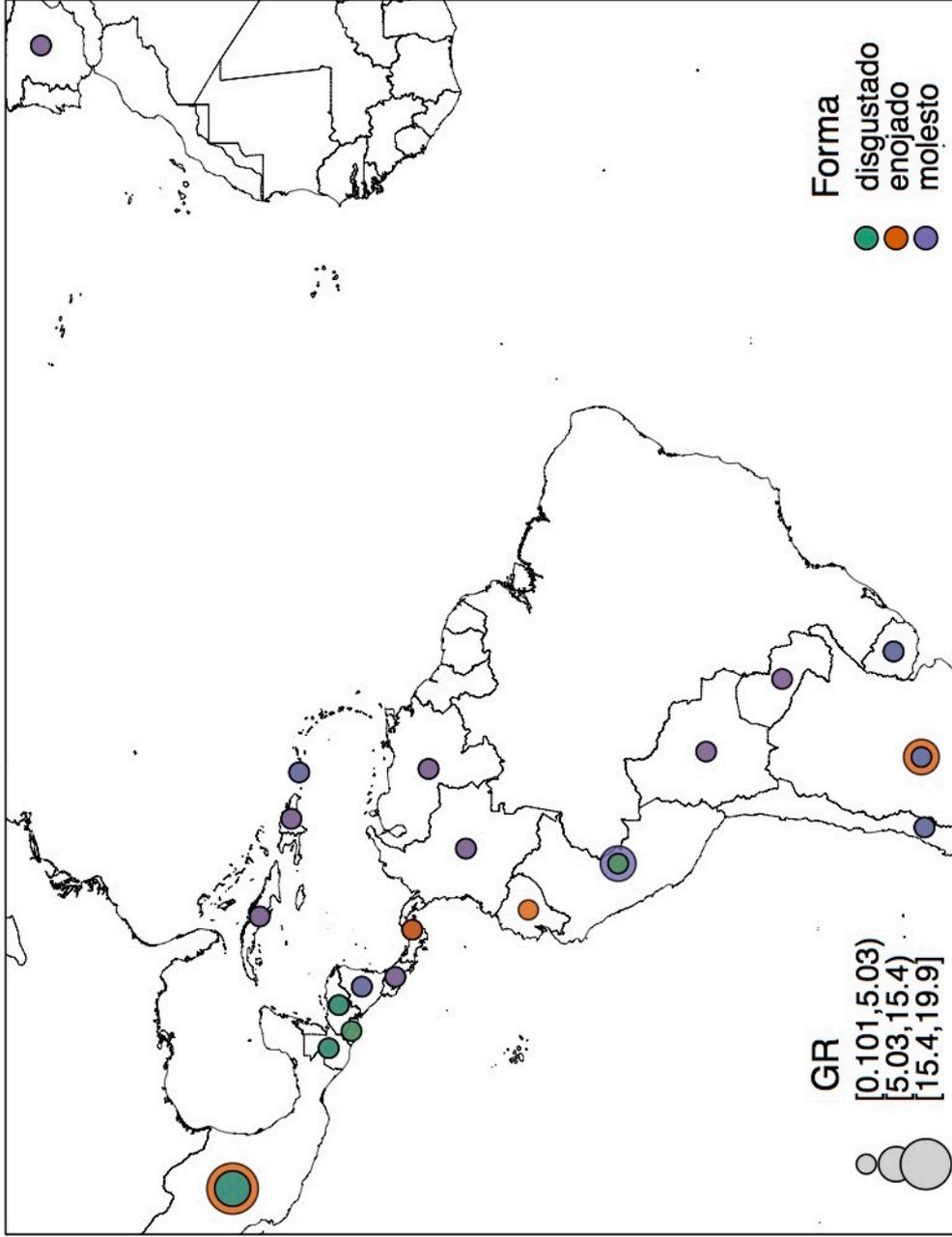
AMUSEMENT



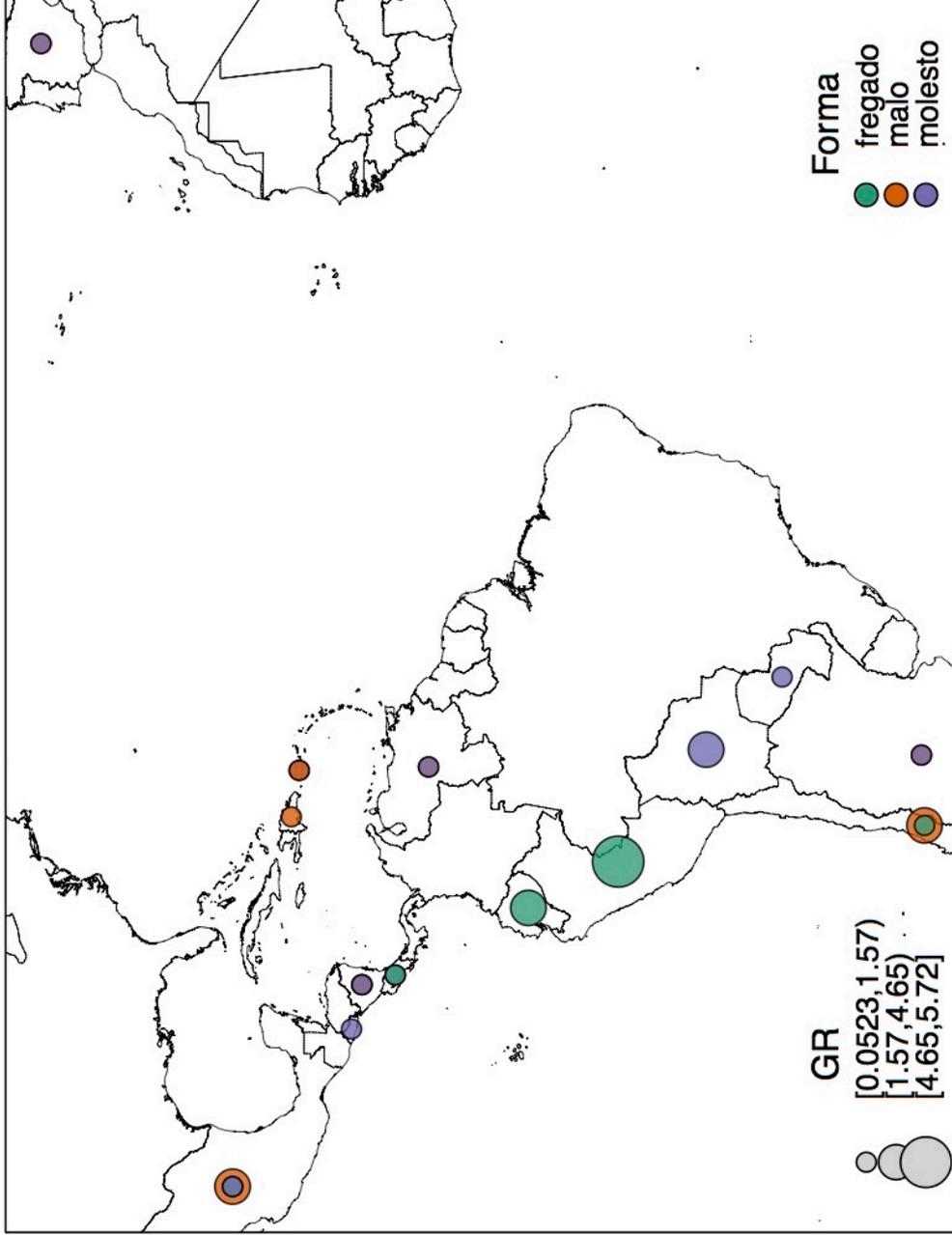
ANGER



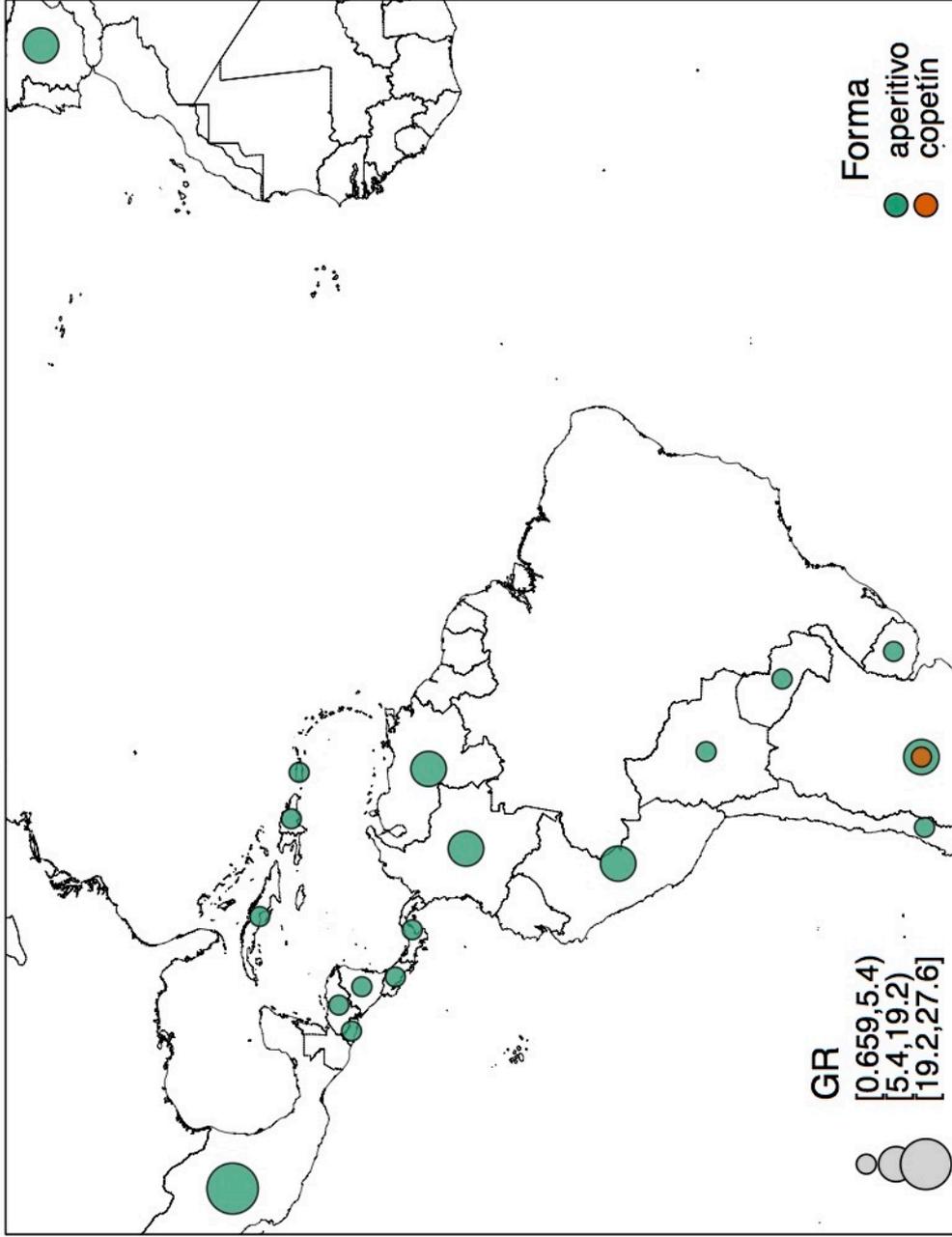
ANGRY



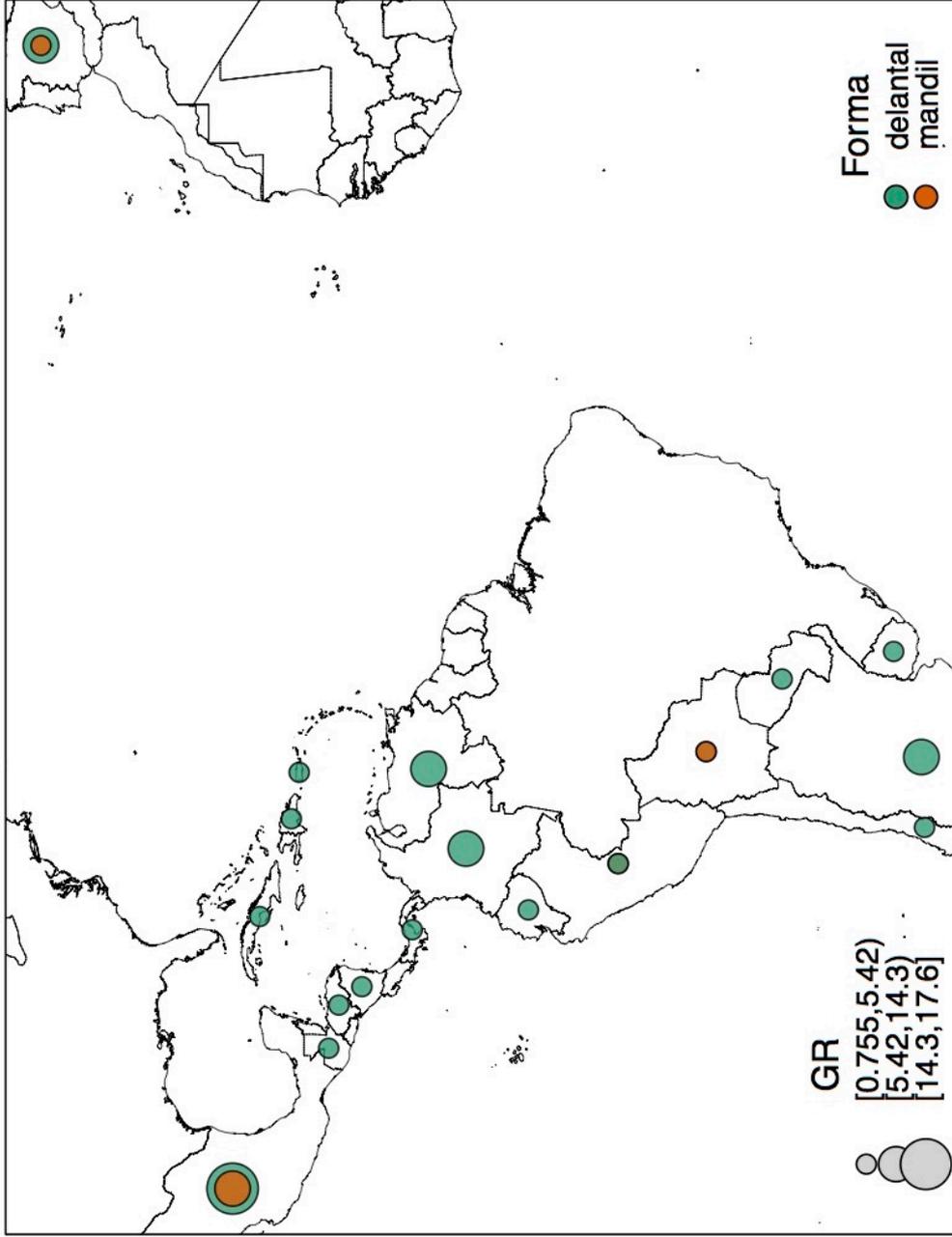
ANNOYING



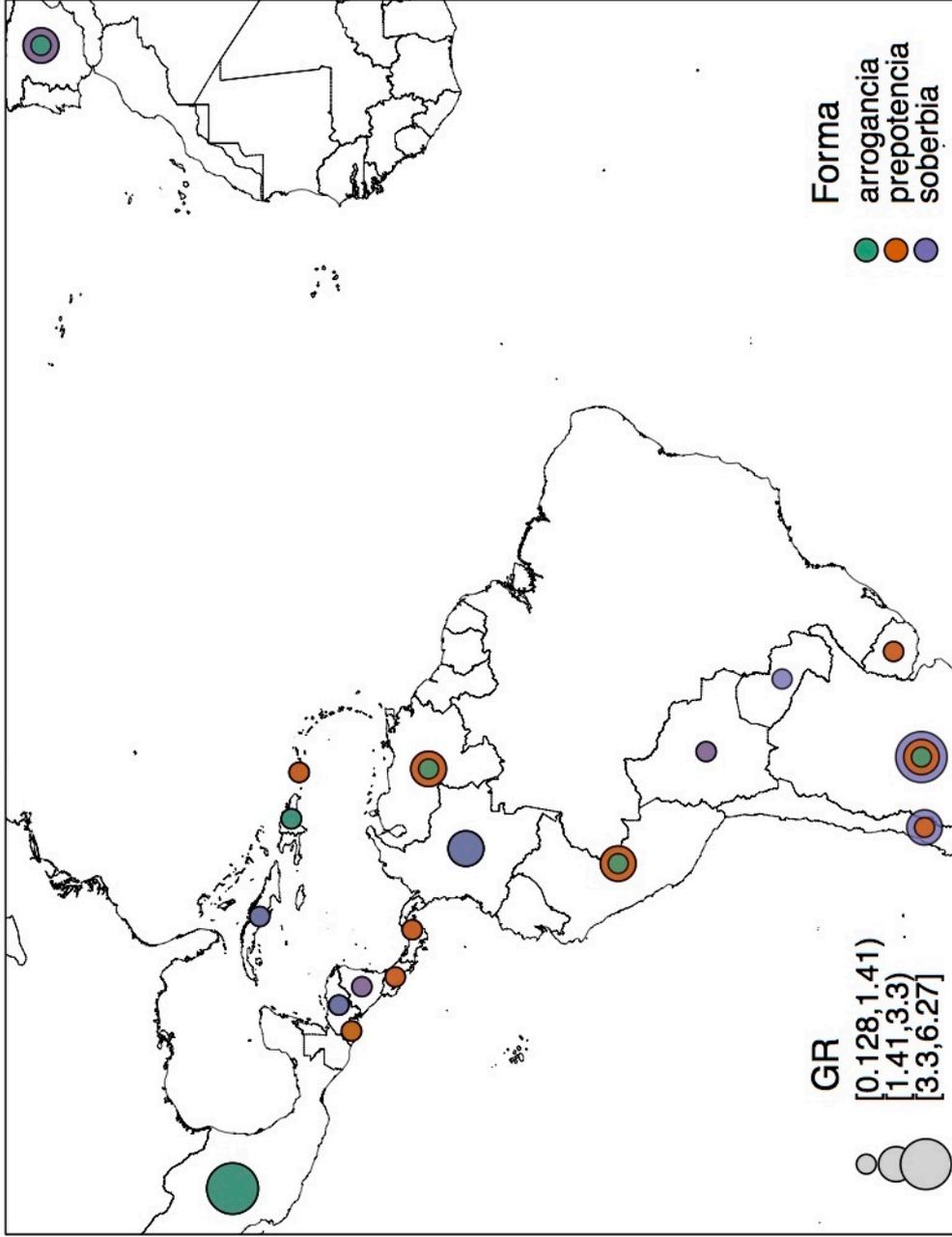
APERITIF



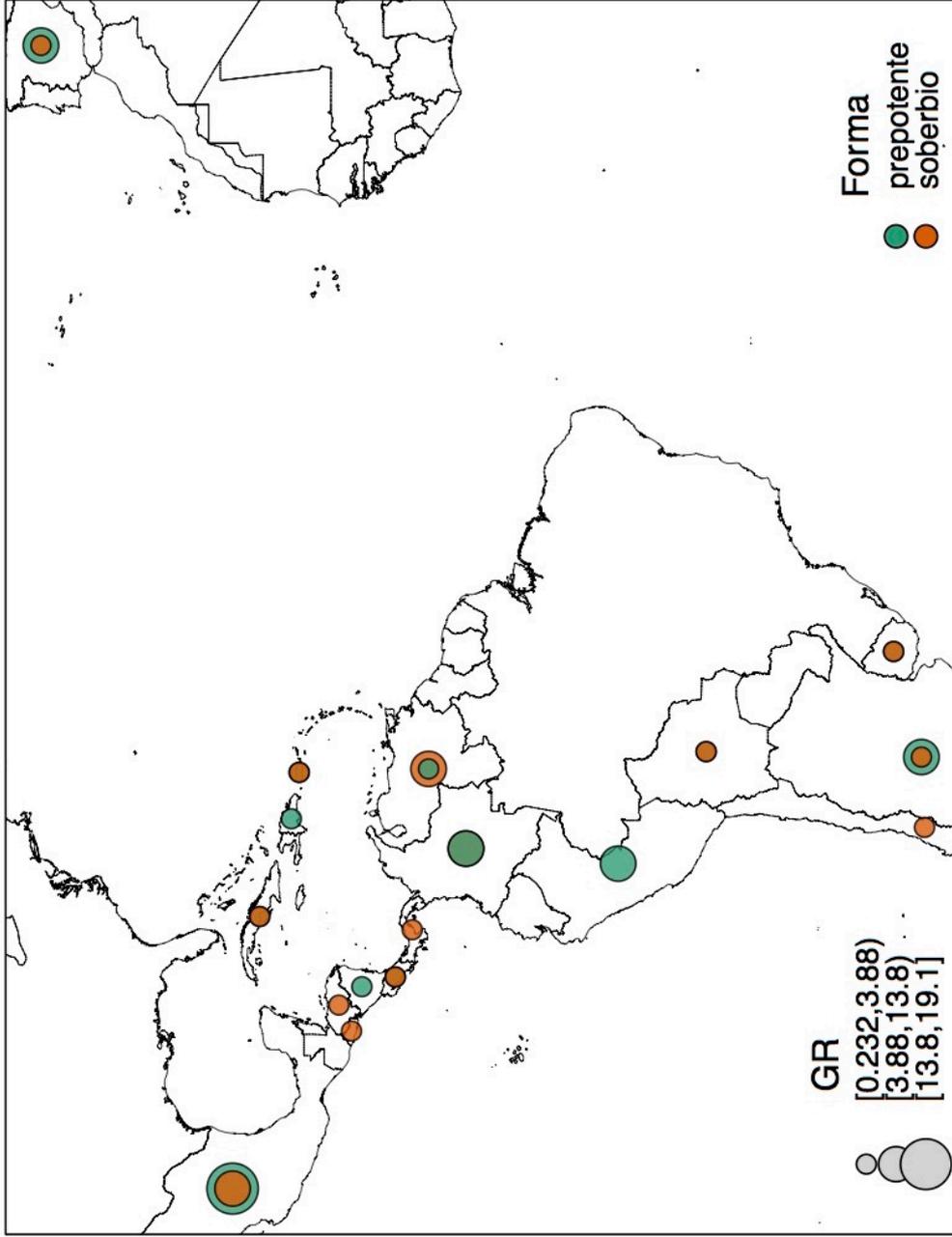
APRON



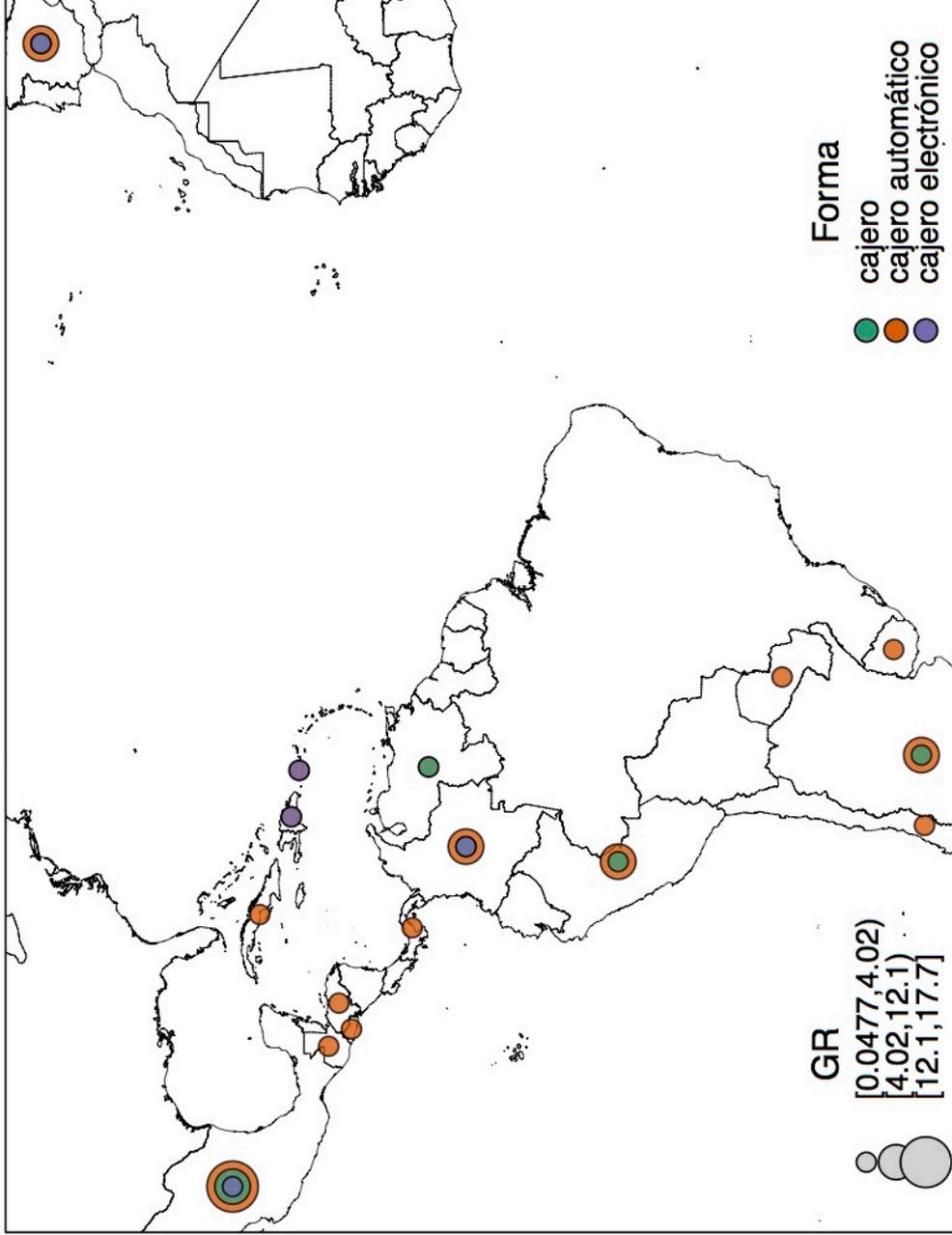
ARROGANCE



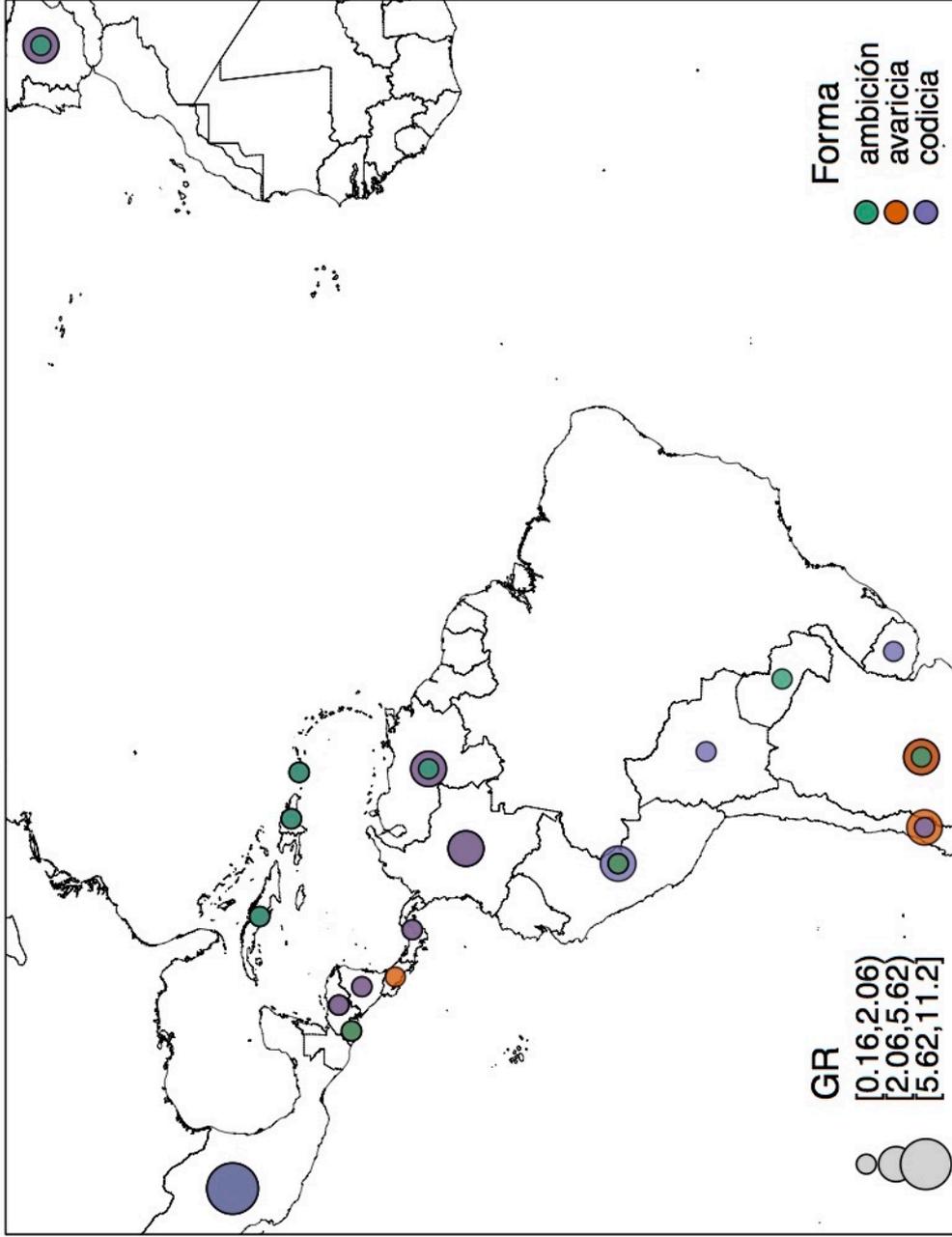
ARROGANT



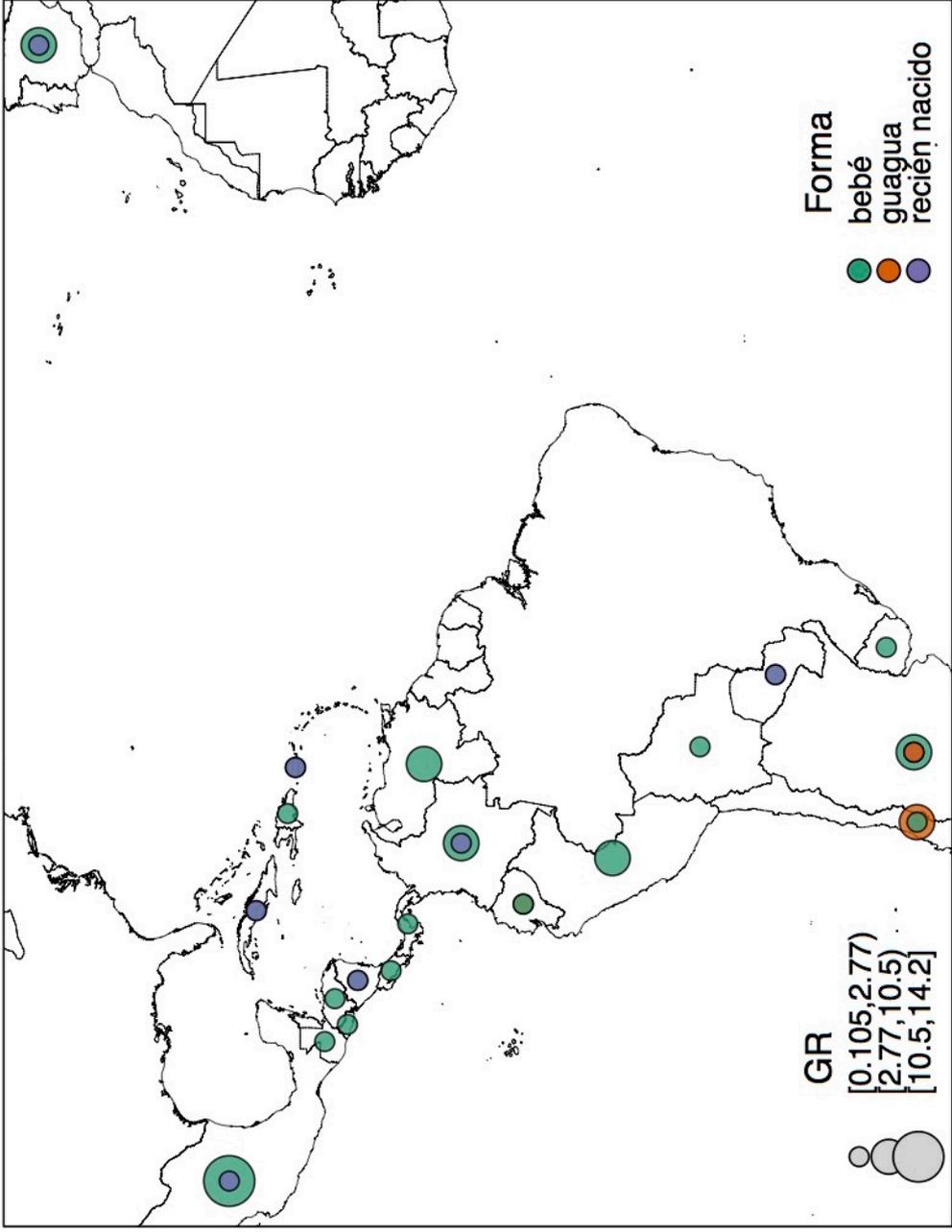
AUTOMATIC CASHIER



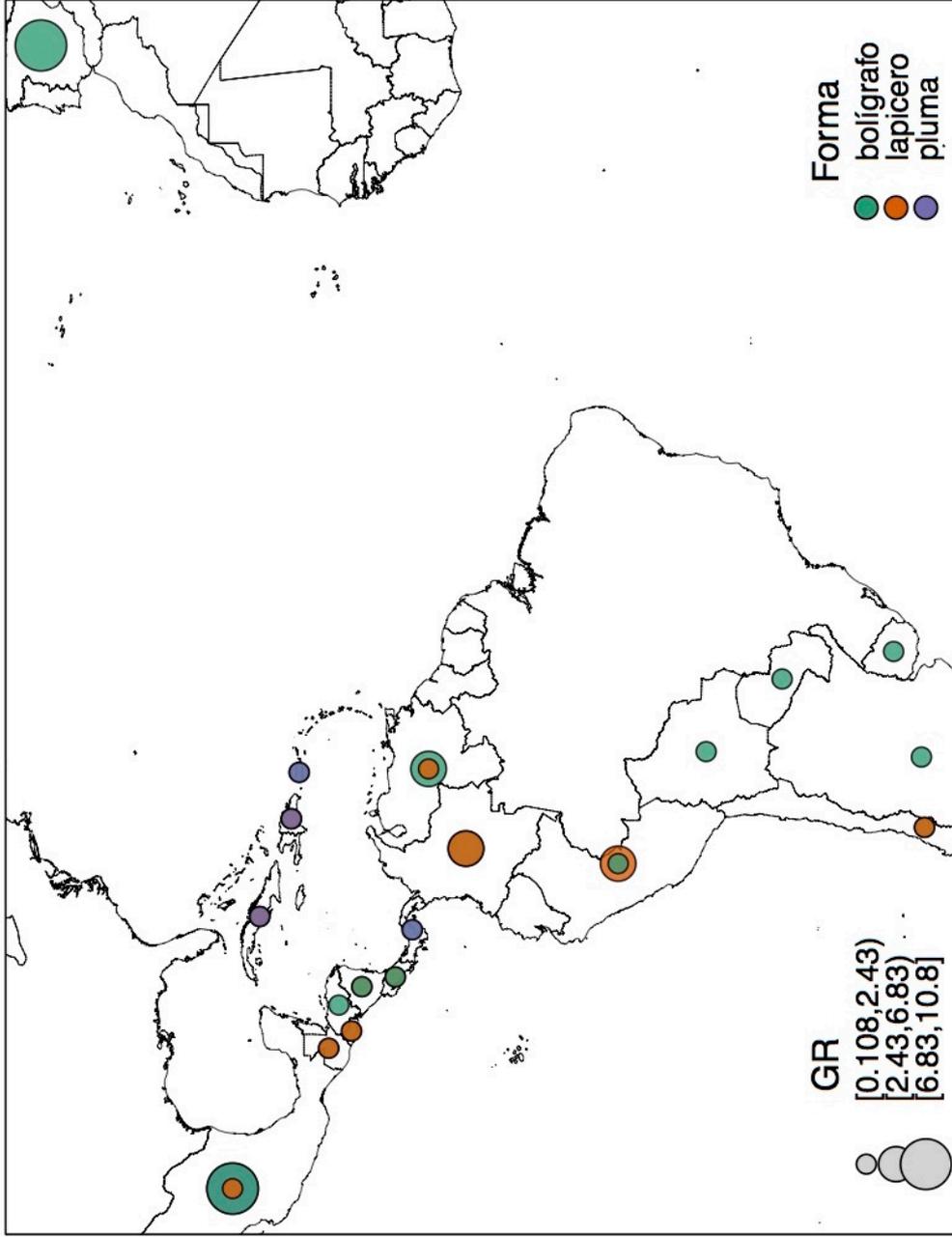
AVARICE



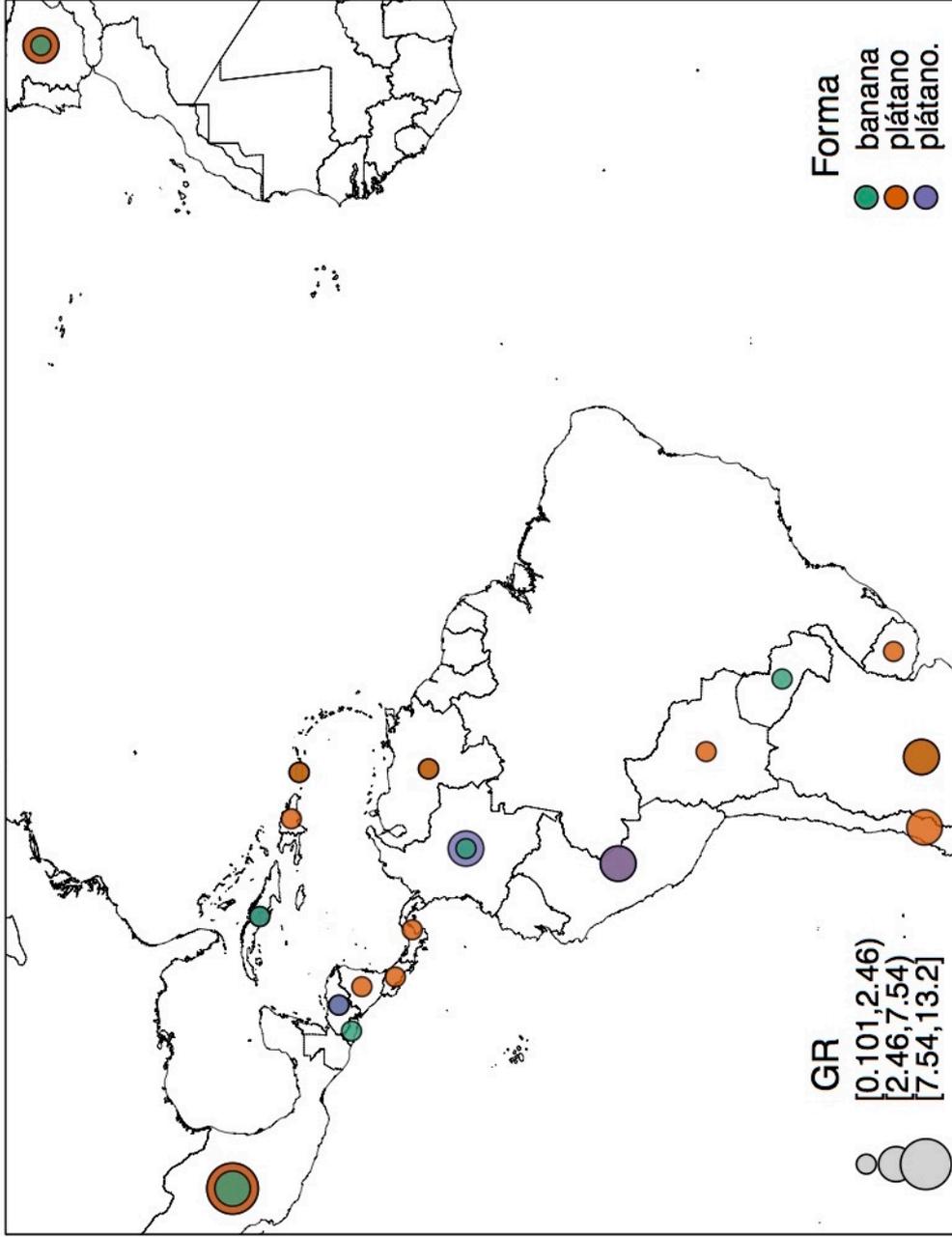
BABY



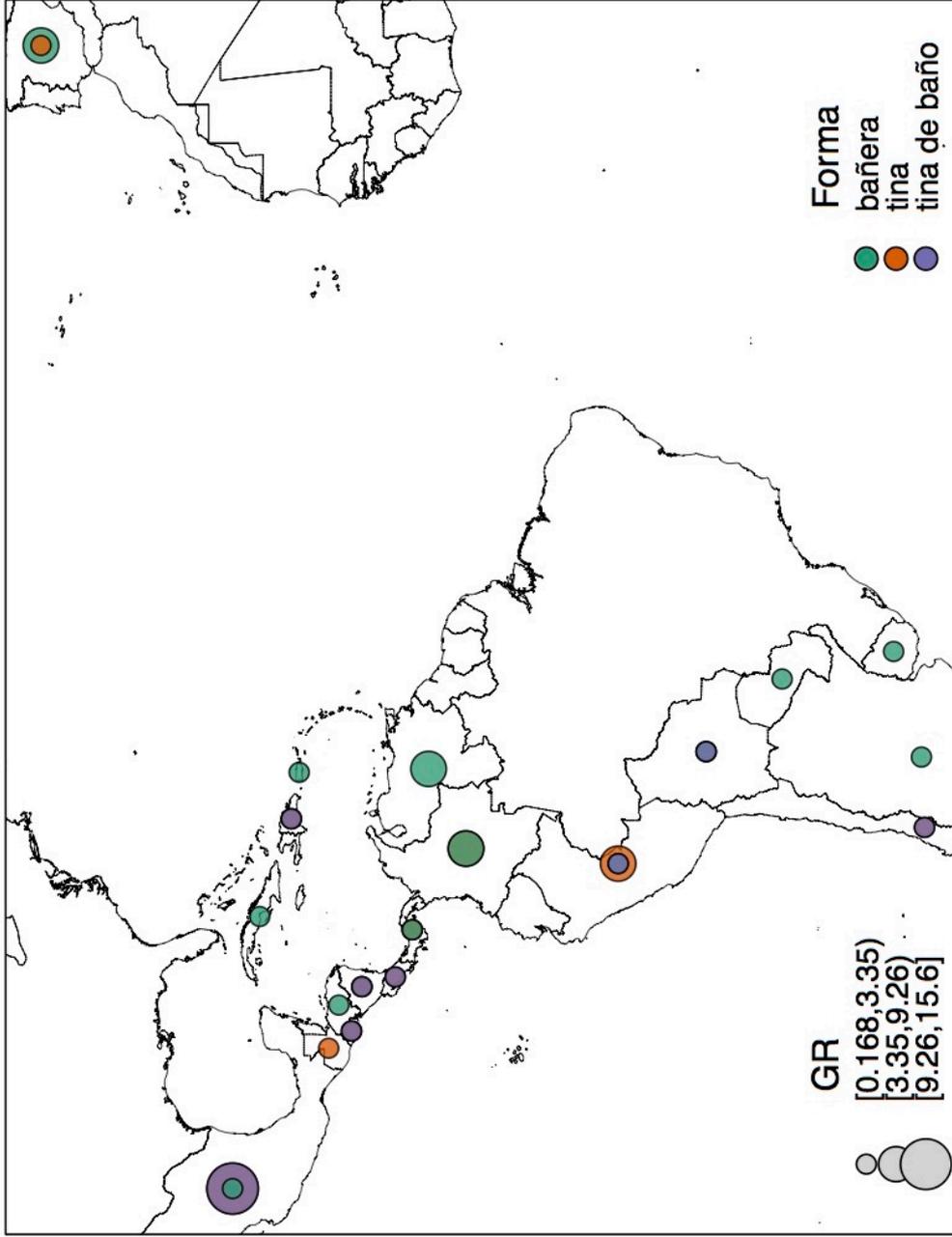
BALL-POINT PEN



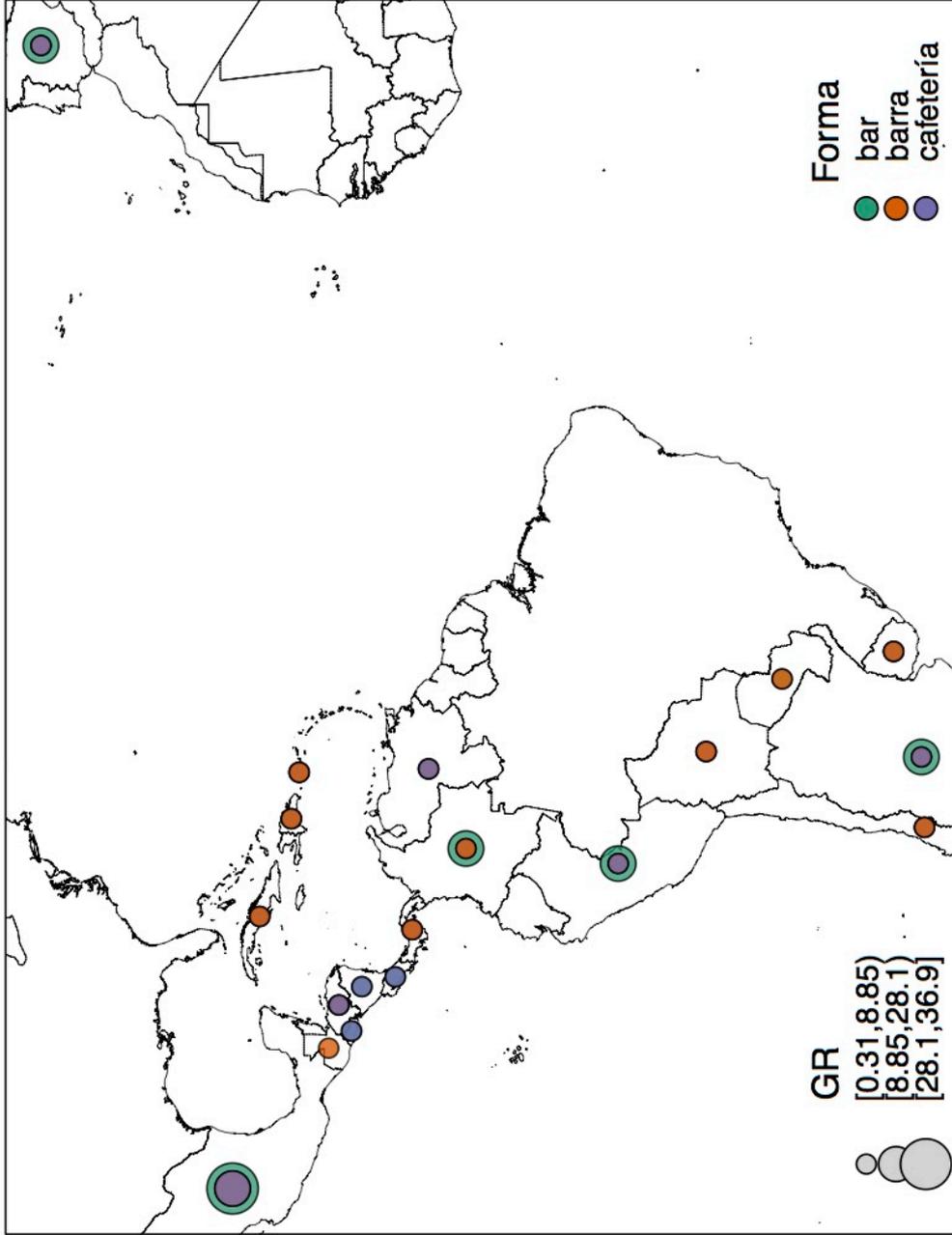
BANANA



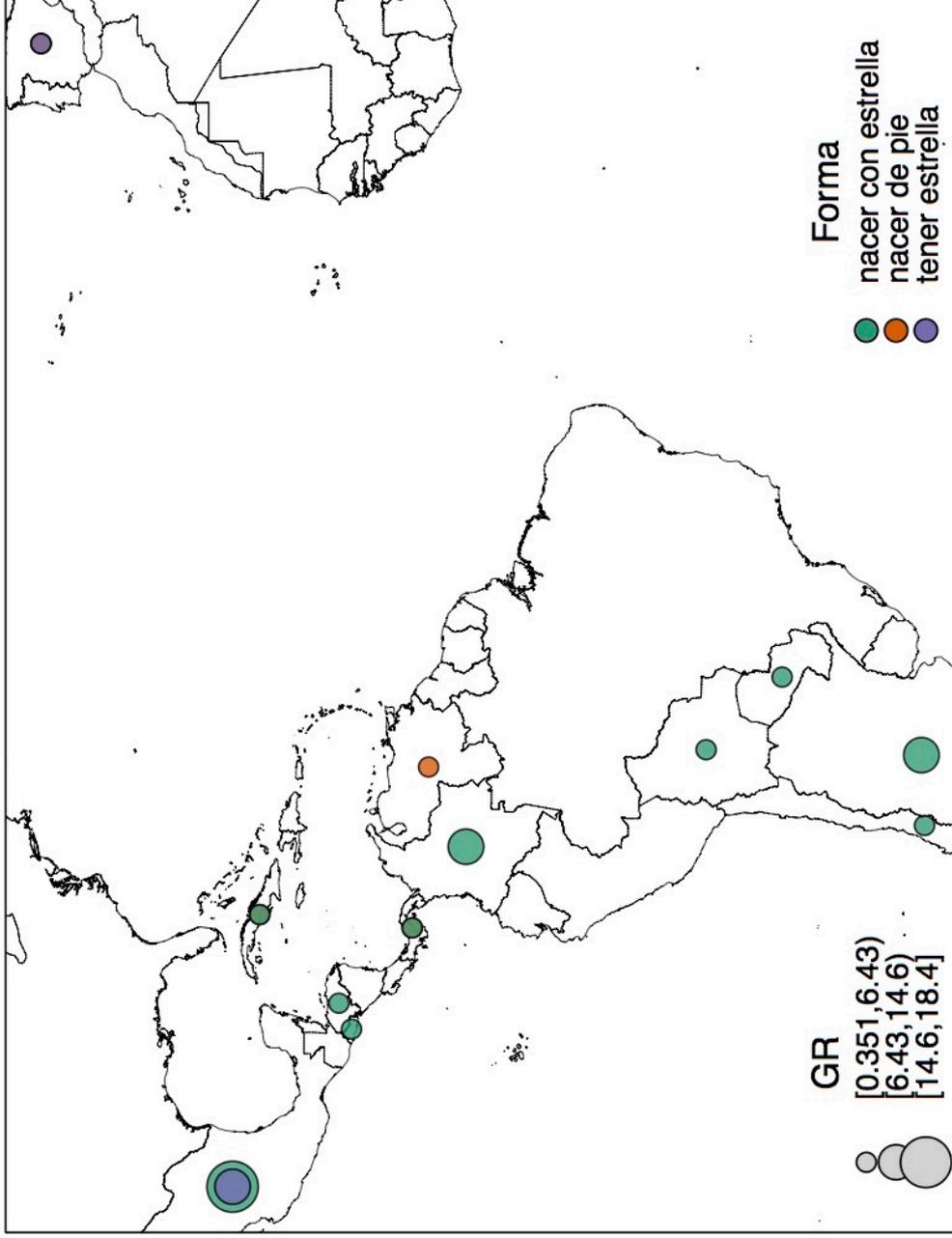
BATHTUB



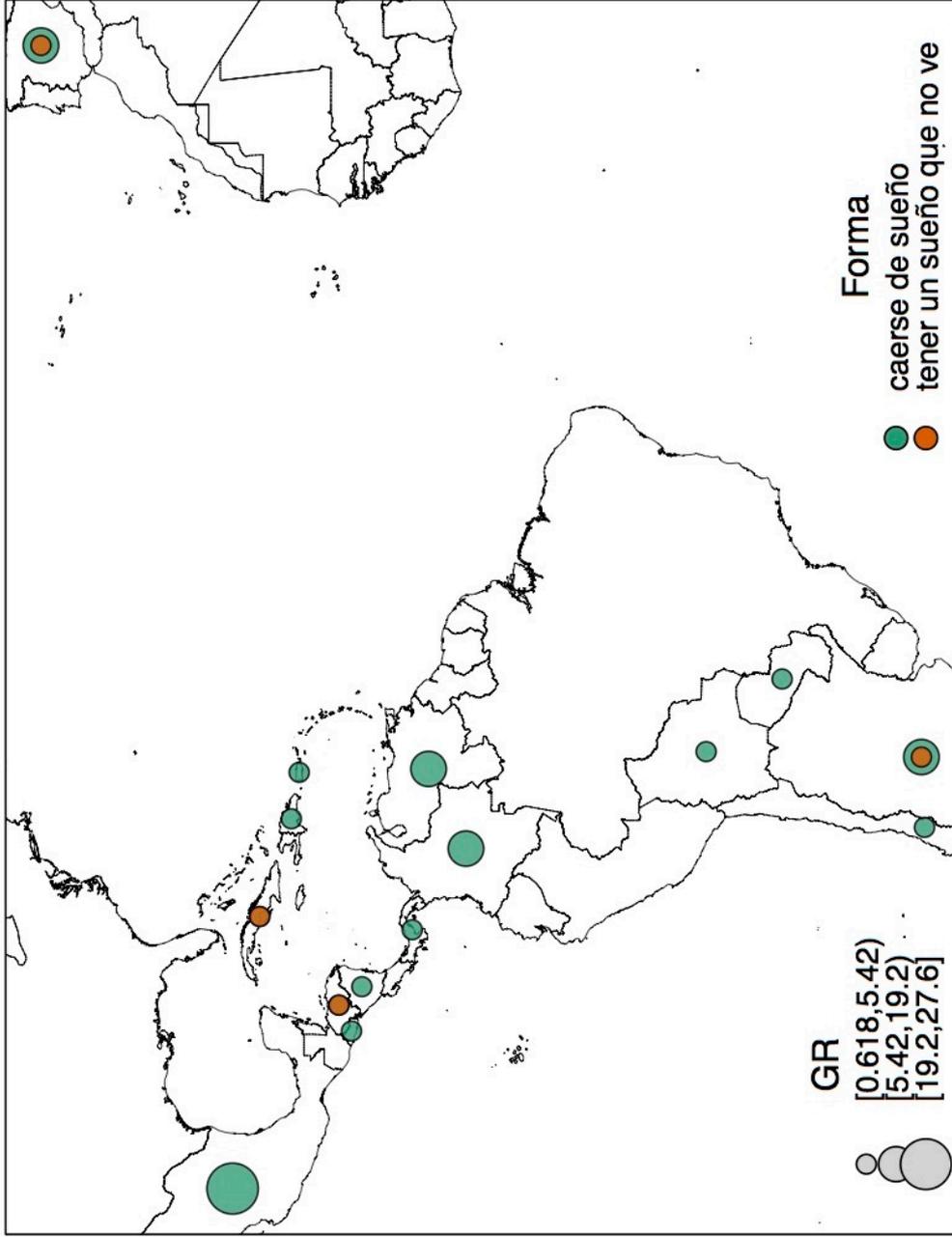
BAR



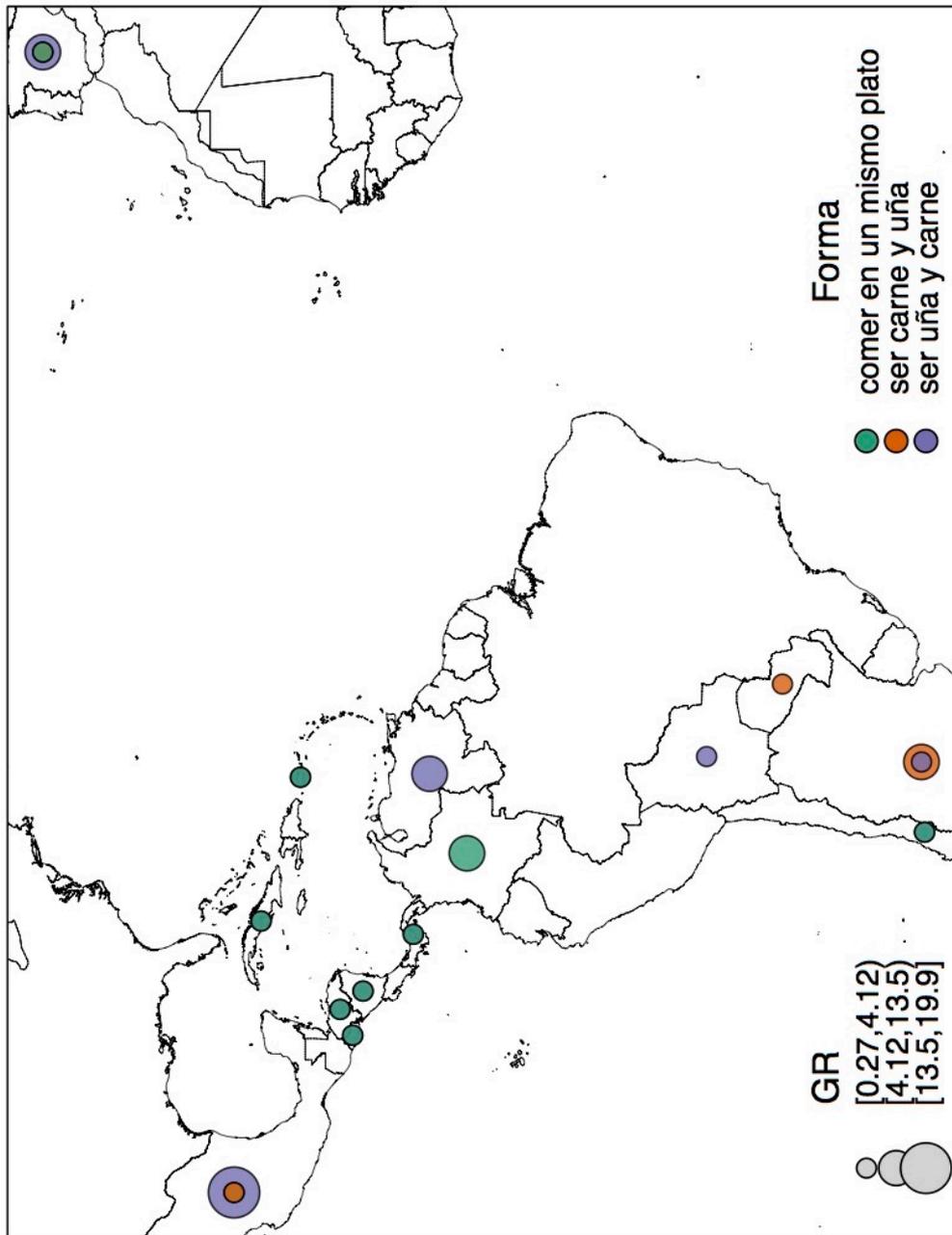
BE BORN LUCKY



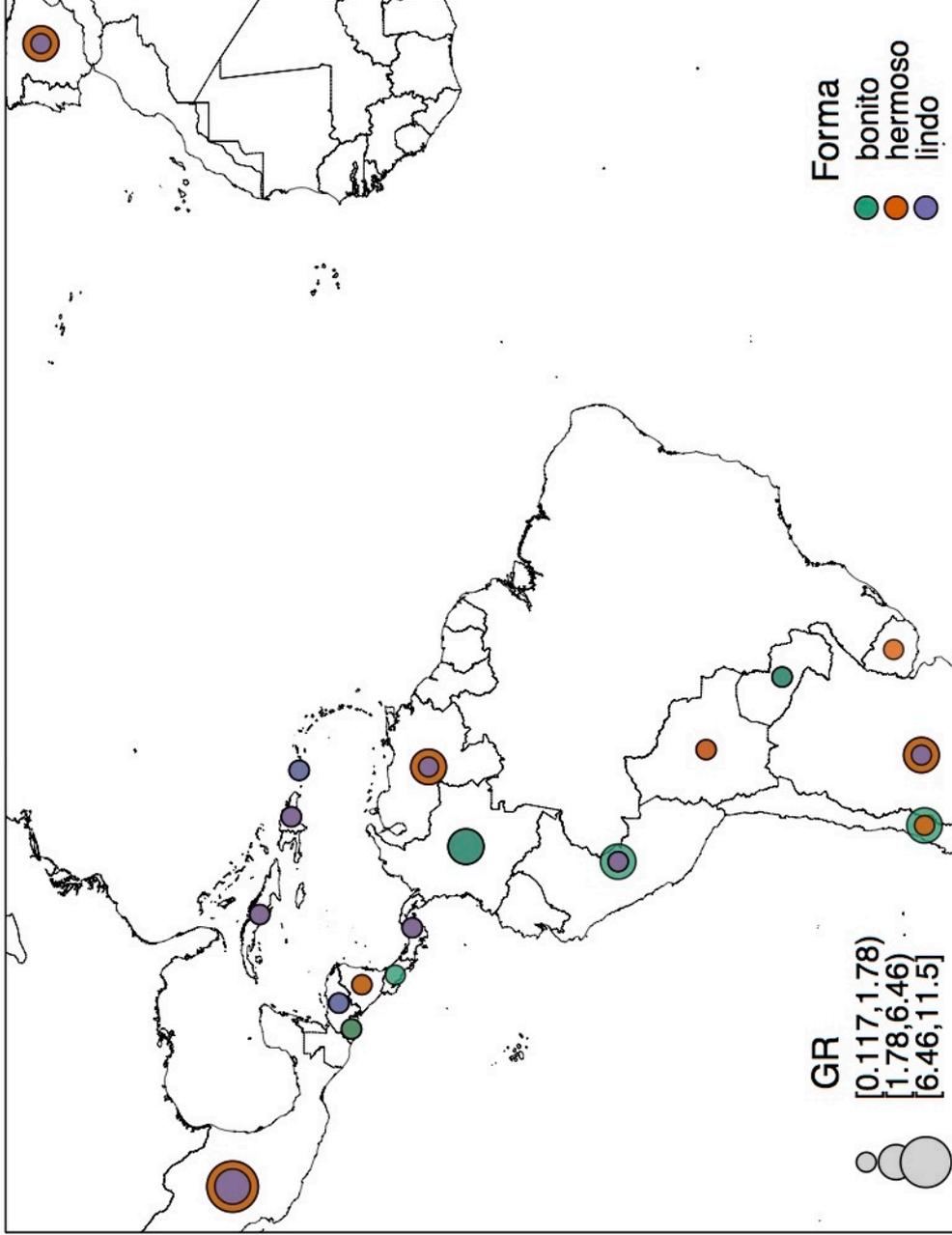
BE FALLING ASLEEP



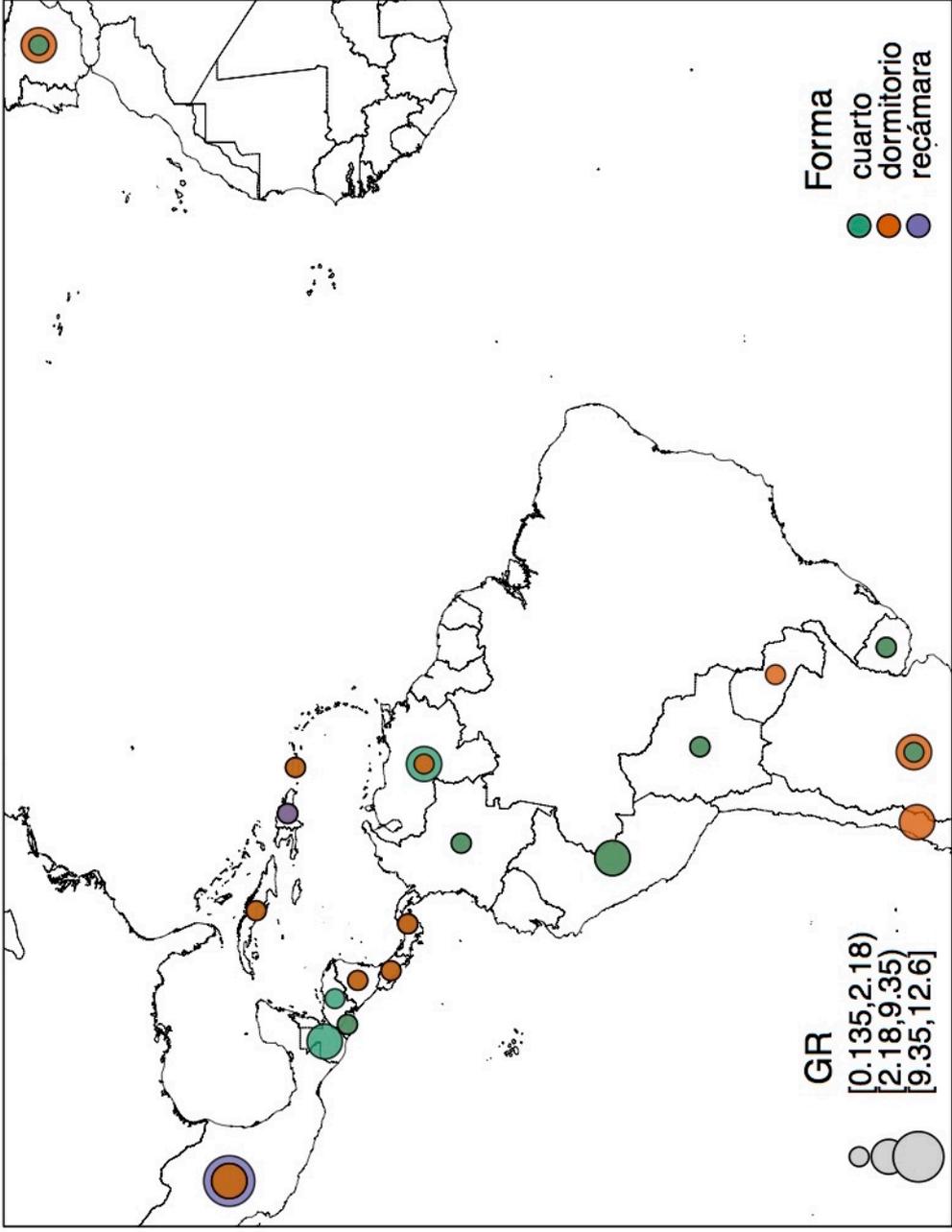
BE INSEPARABLE



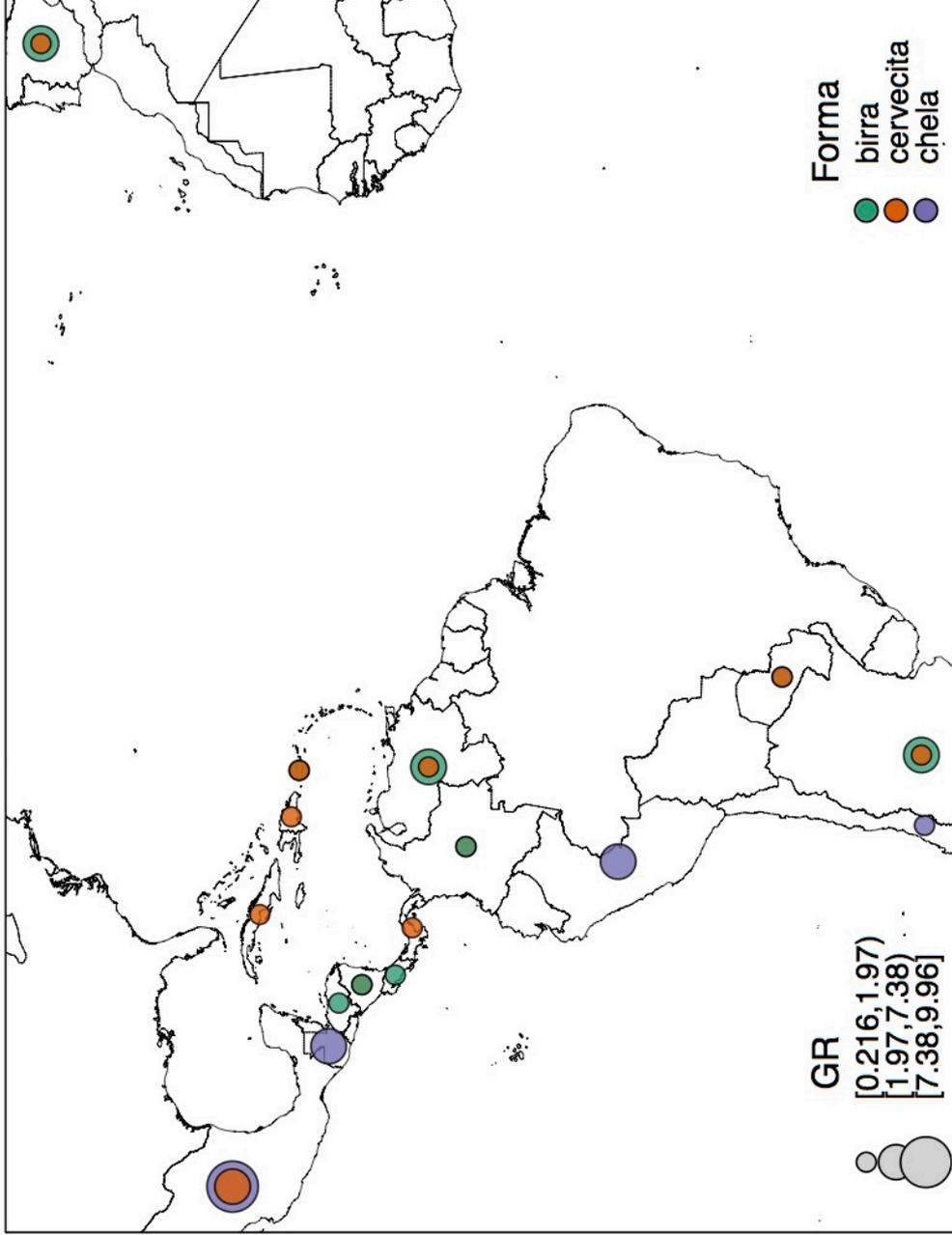
BEAUTIFUL



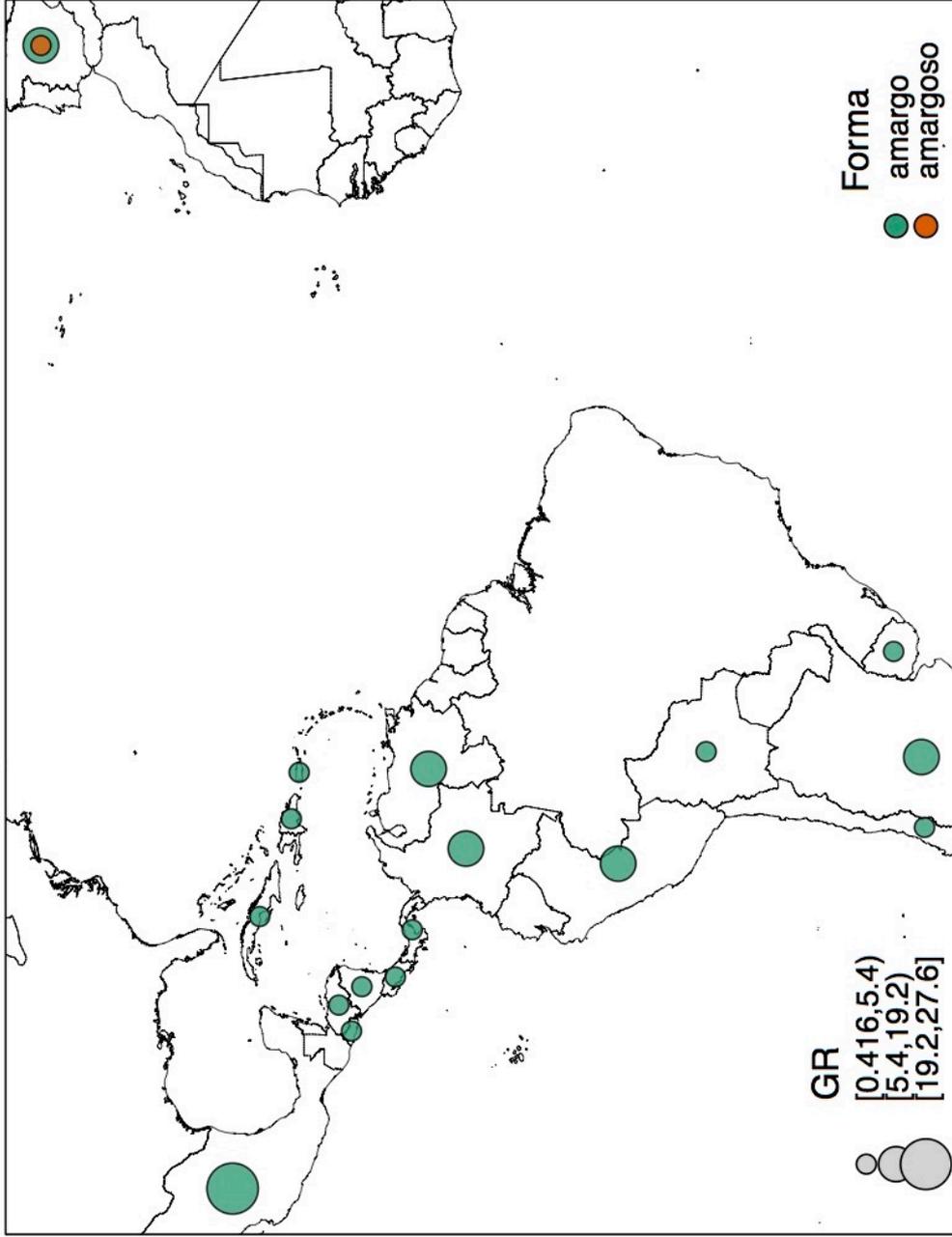
BEDROOM



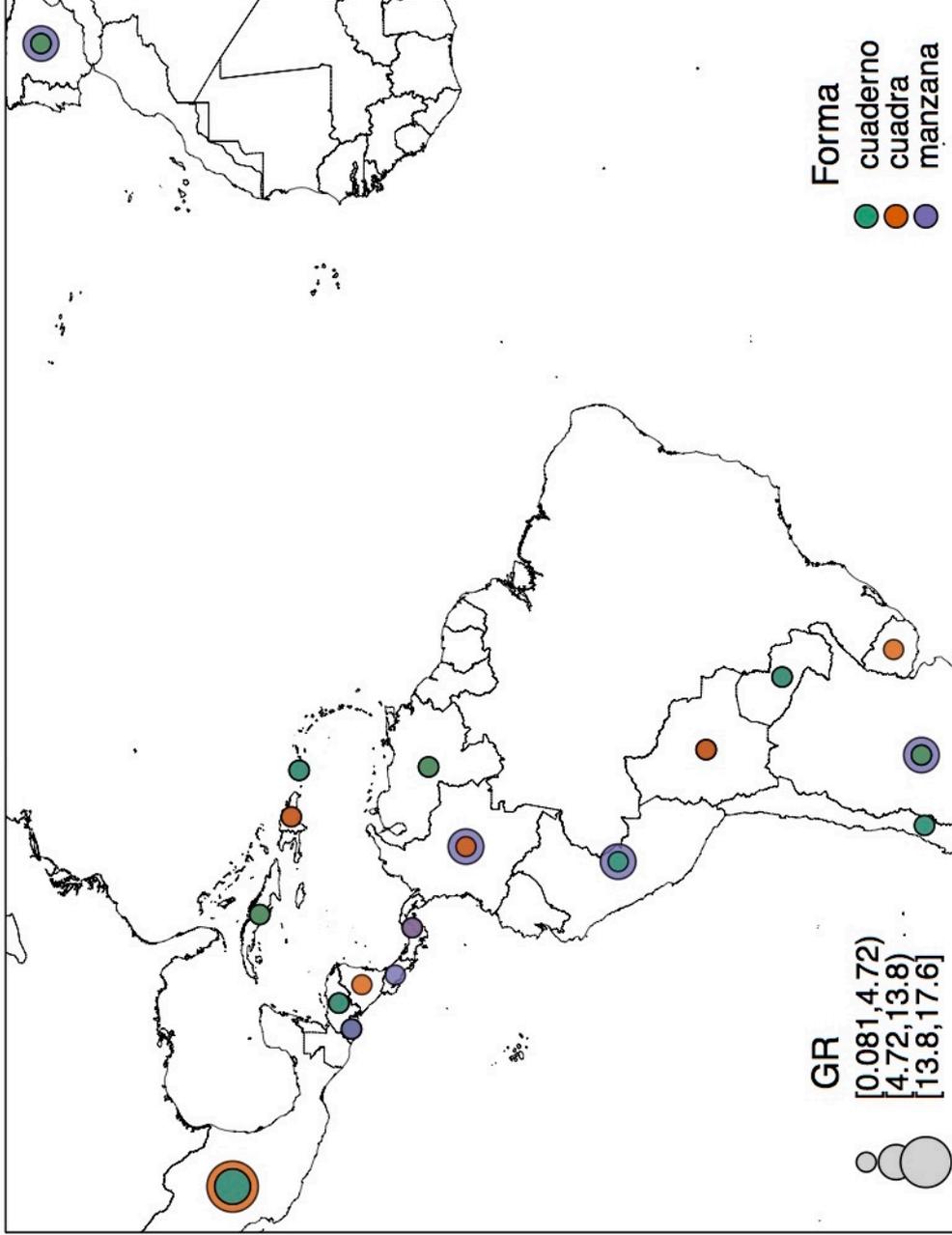
BEER



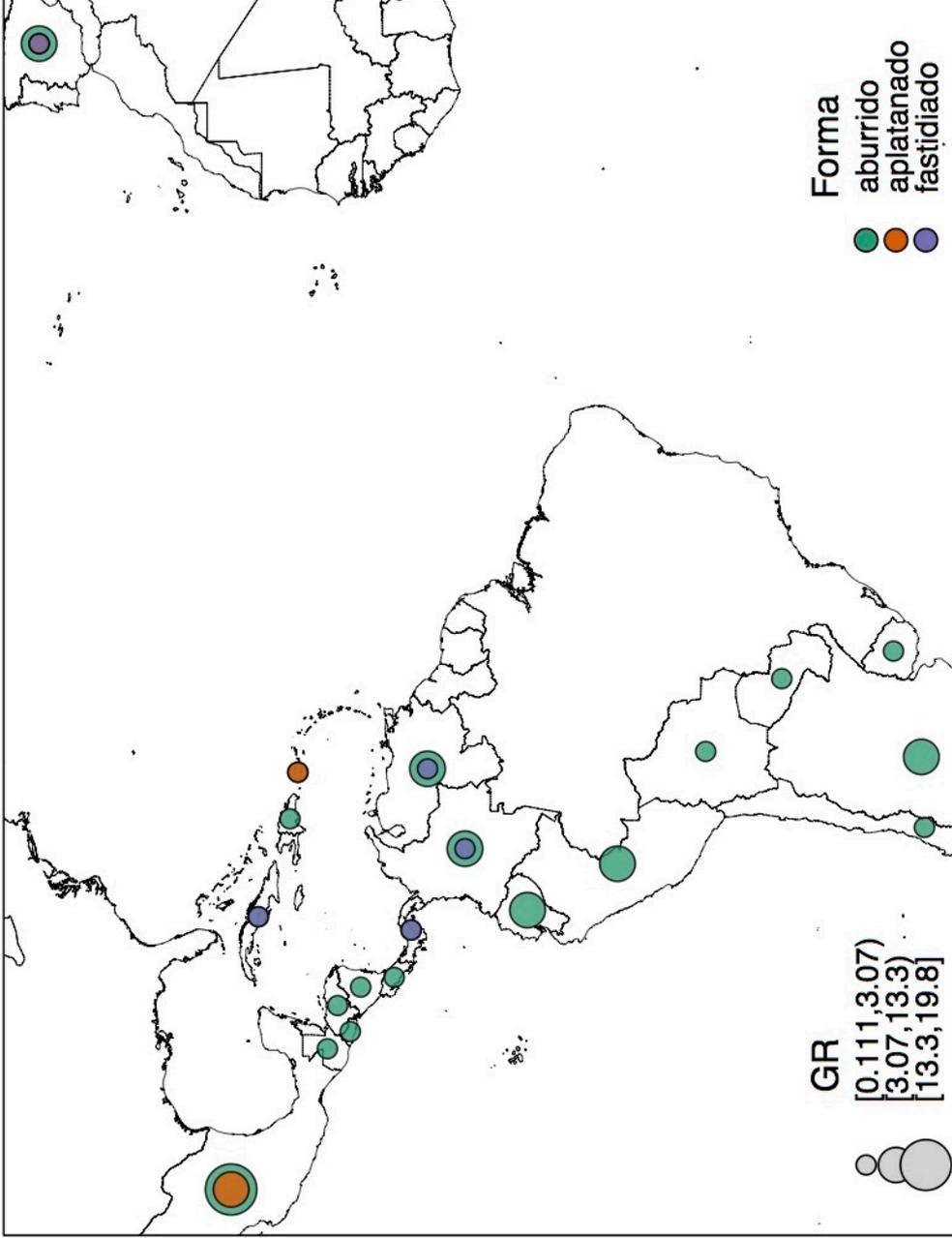
BITTER



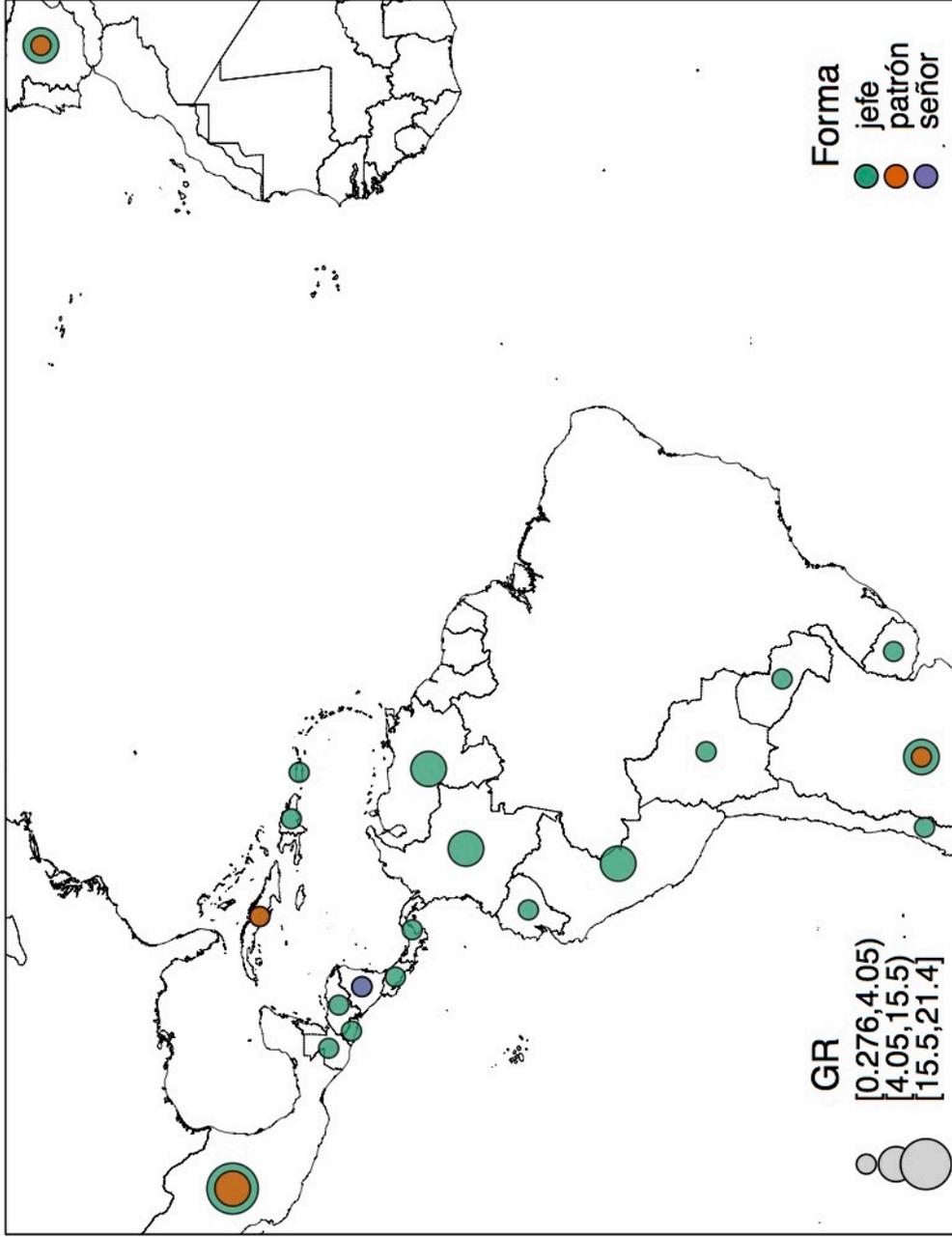
BLOCK



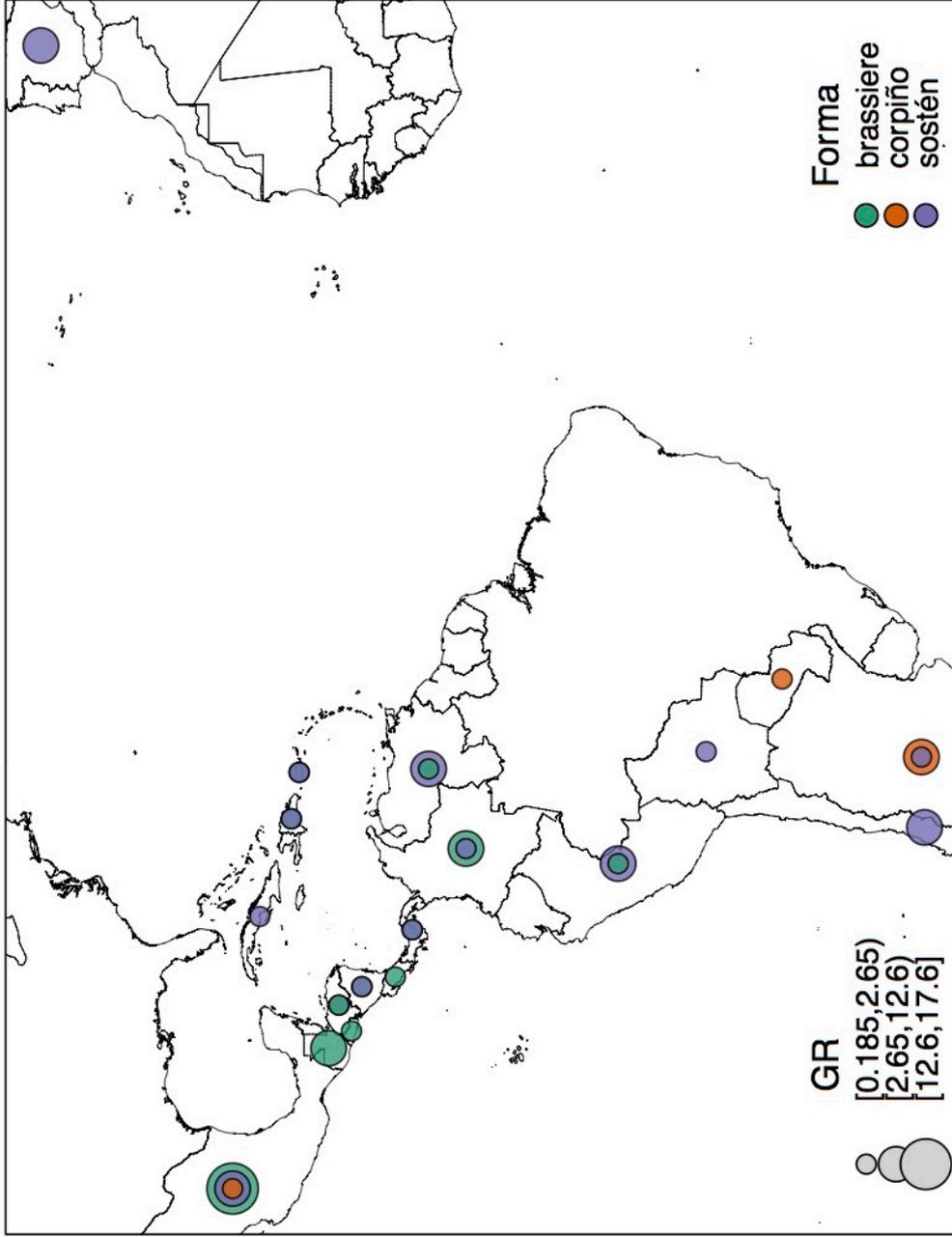
BORED



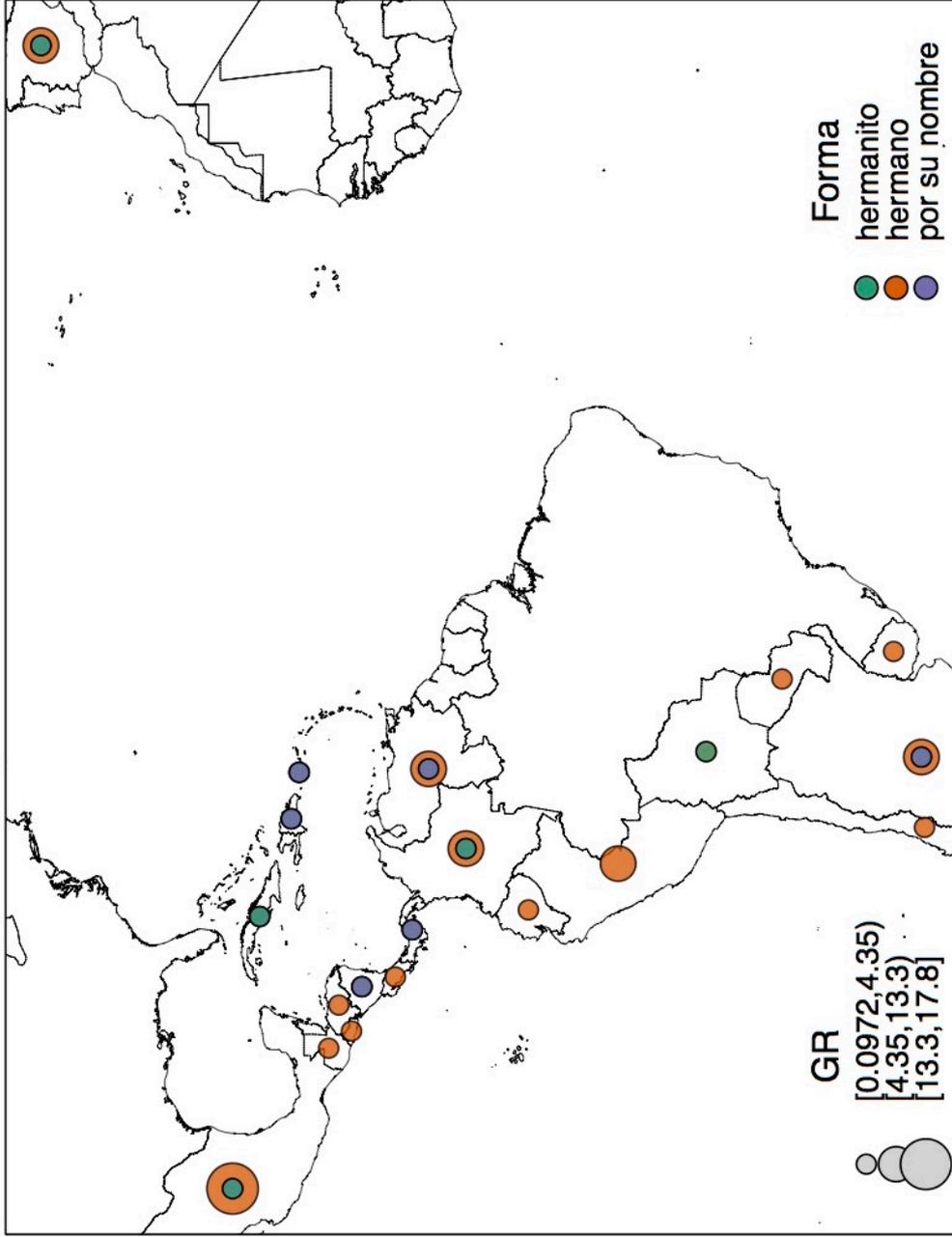
BOSS



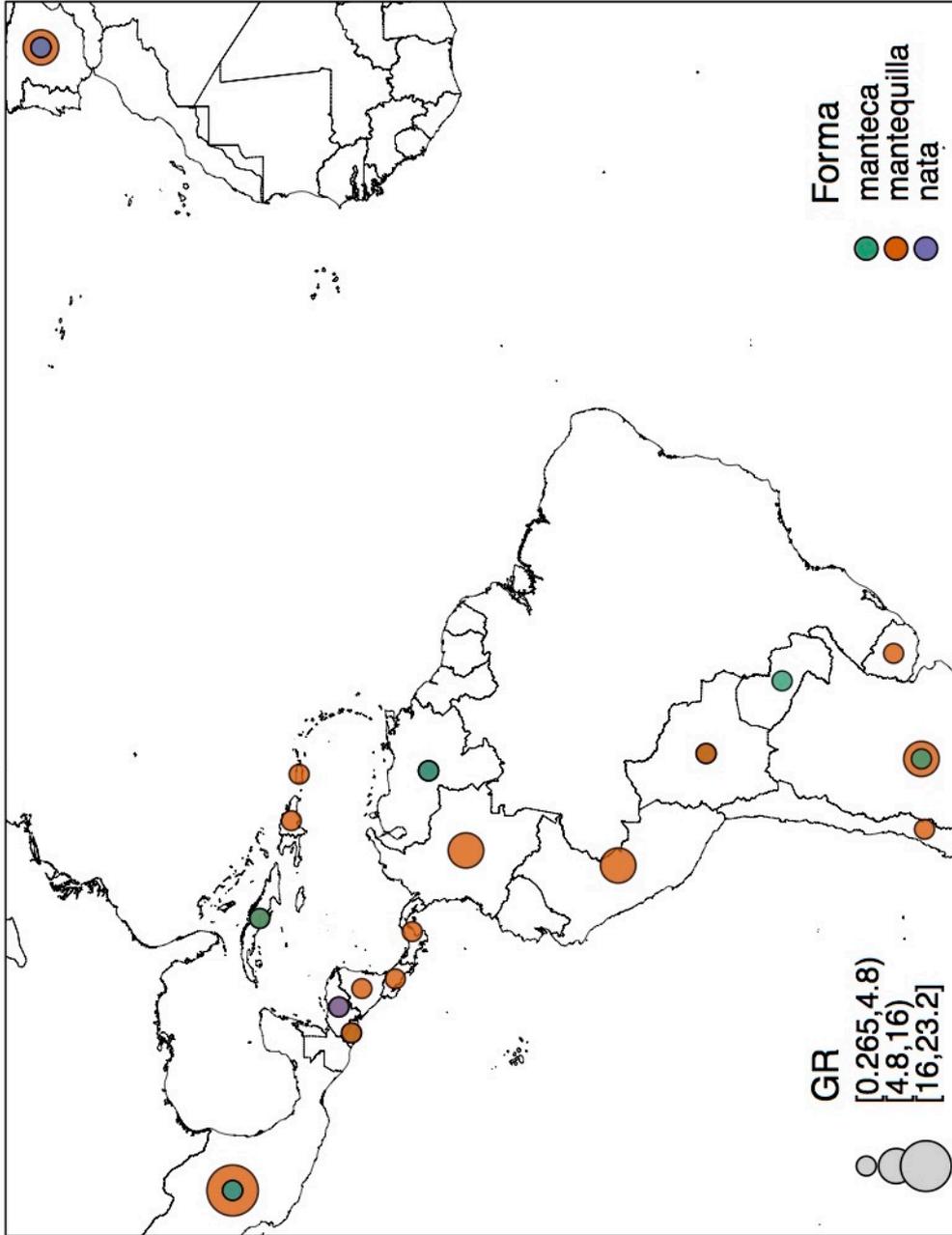
BRASSIERE



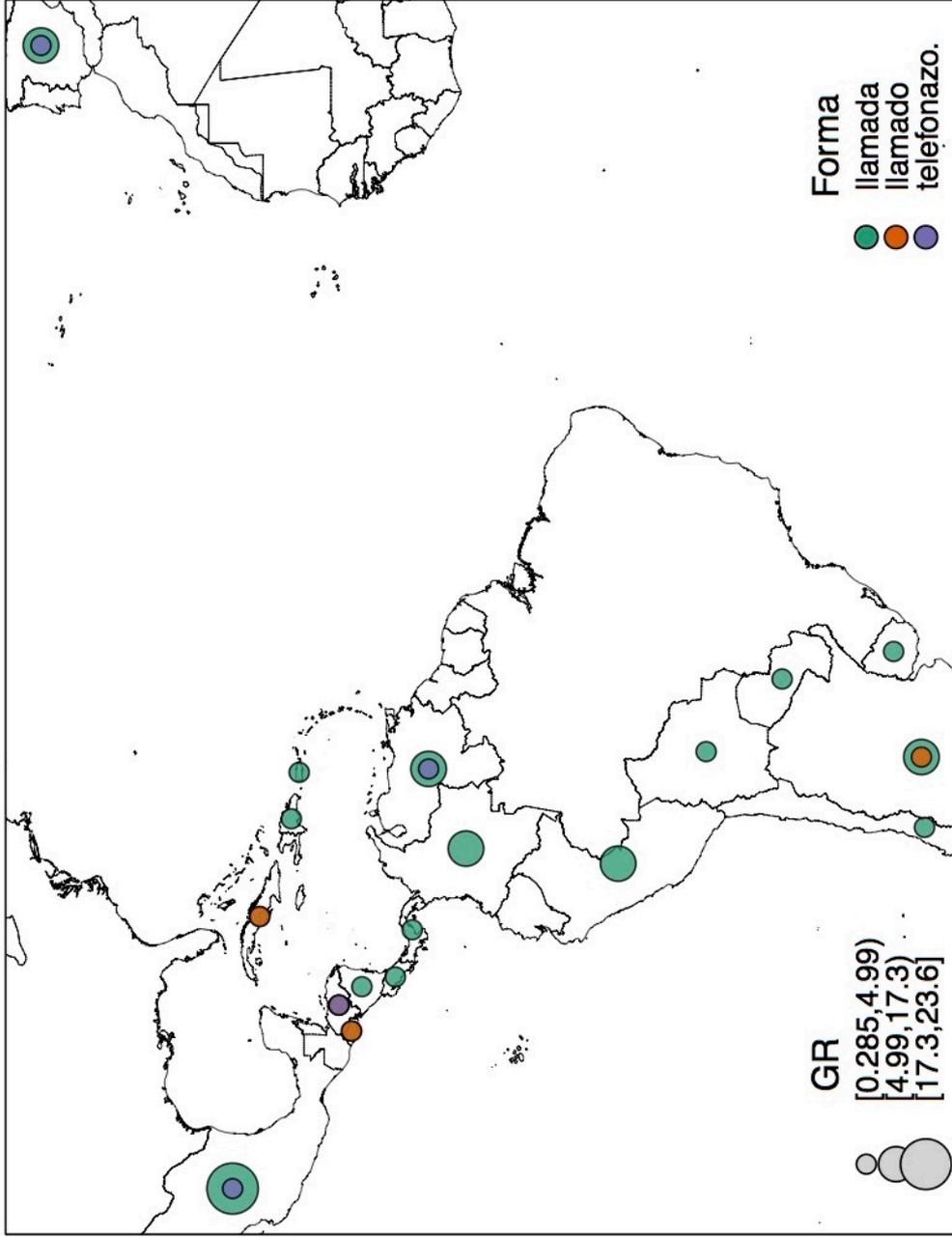
BROTHER



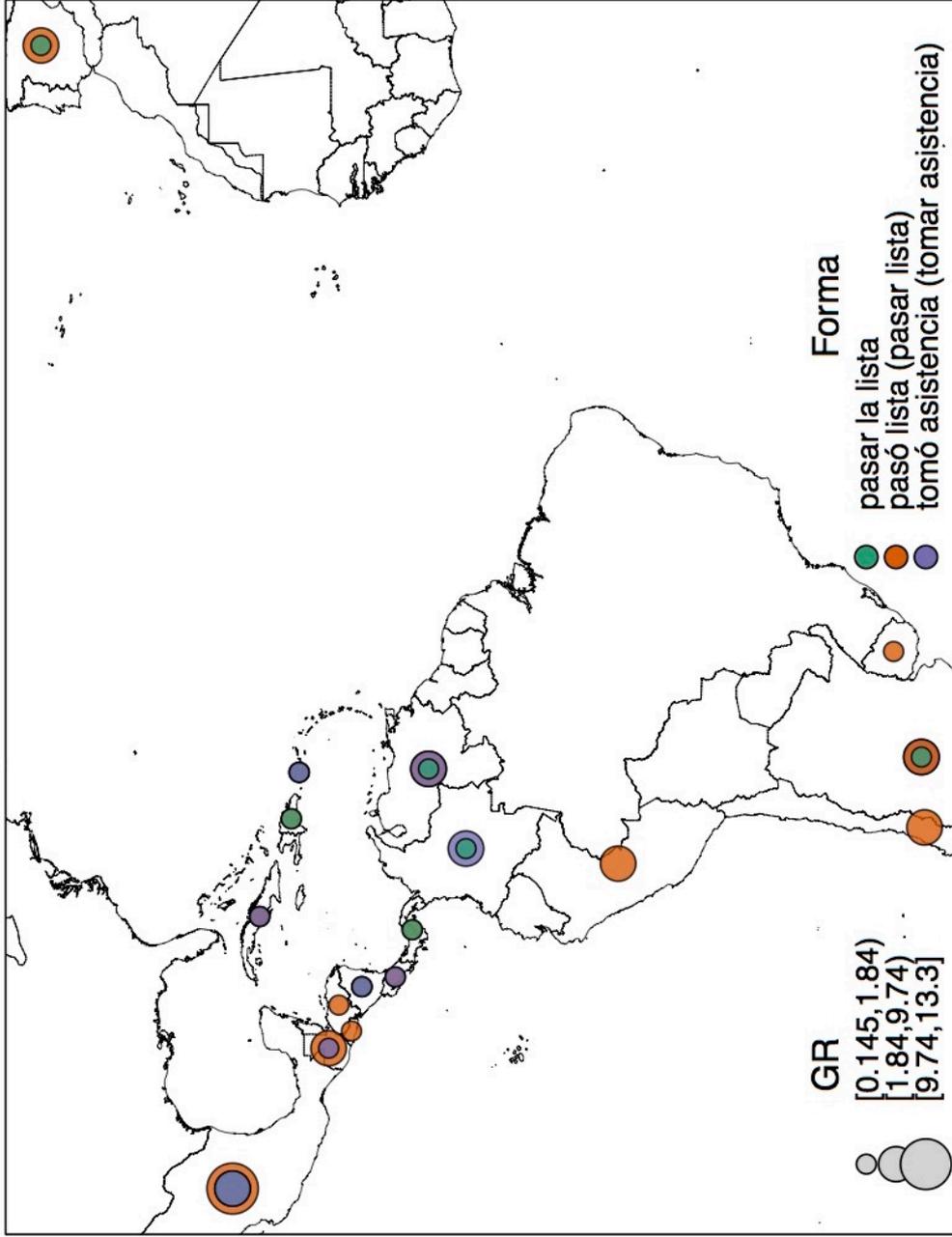
BUTTER



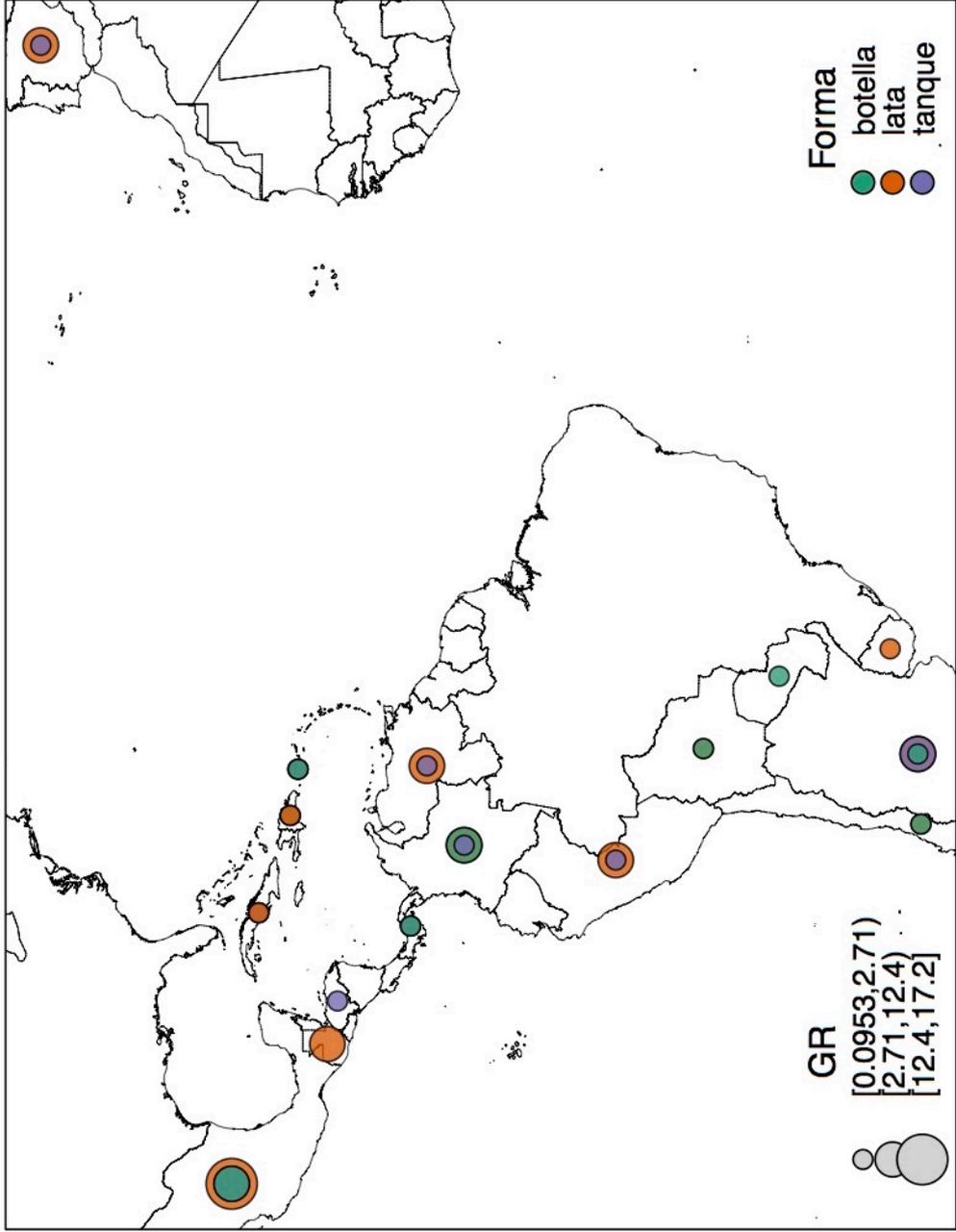
CALL



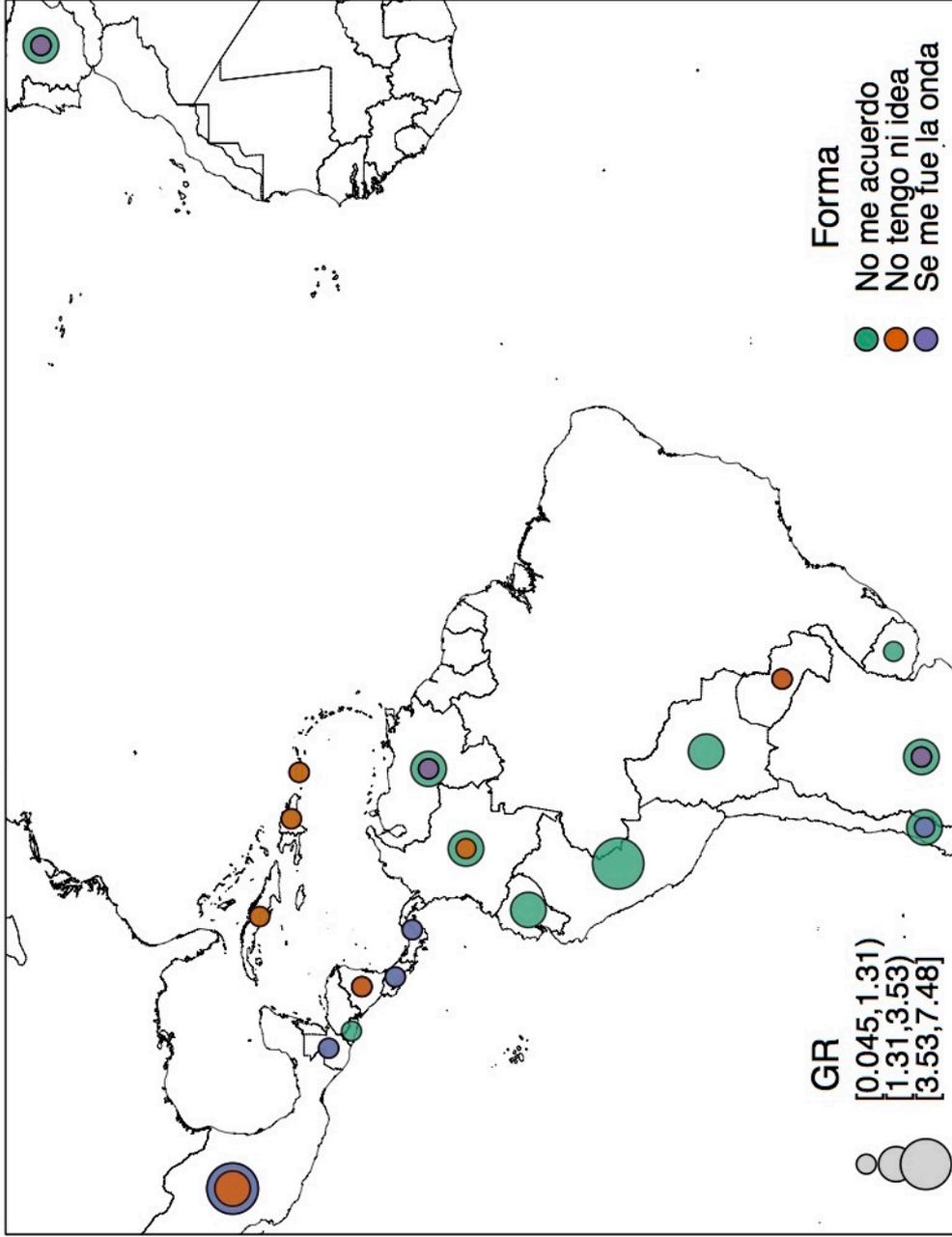
CALL THE ROLL



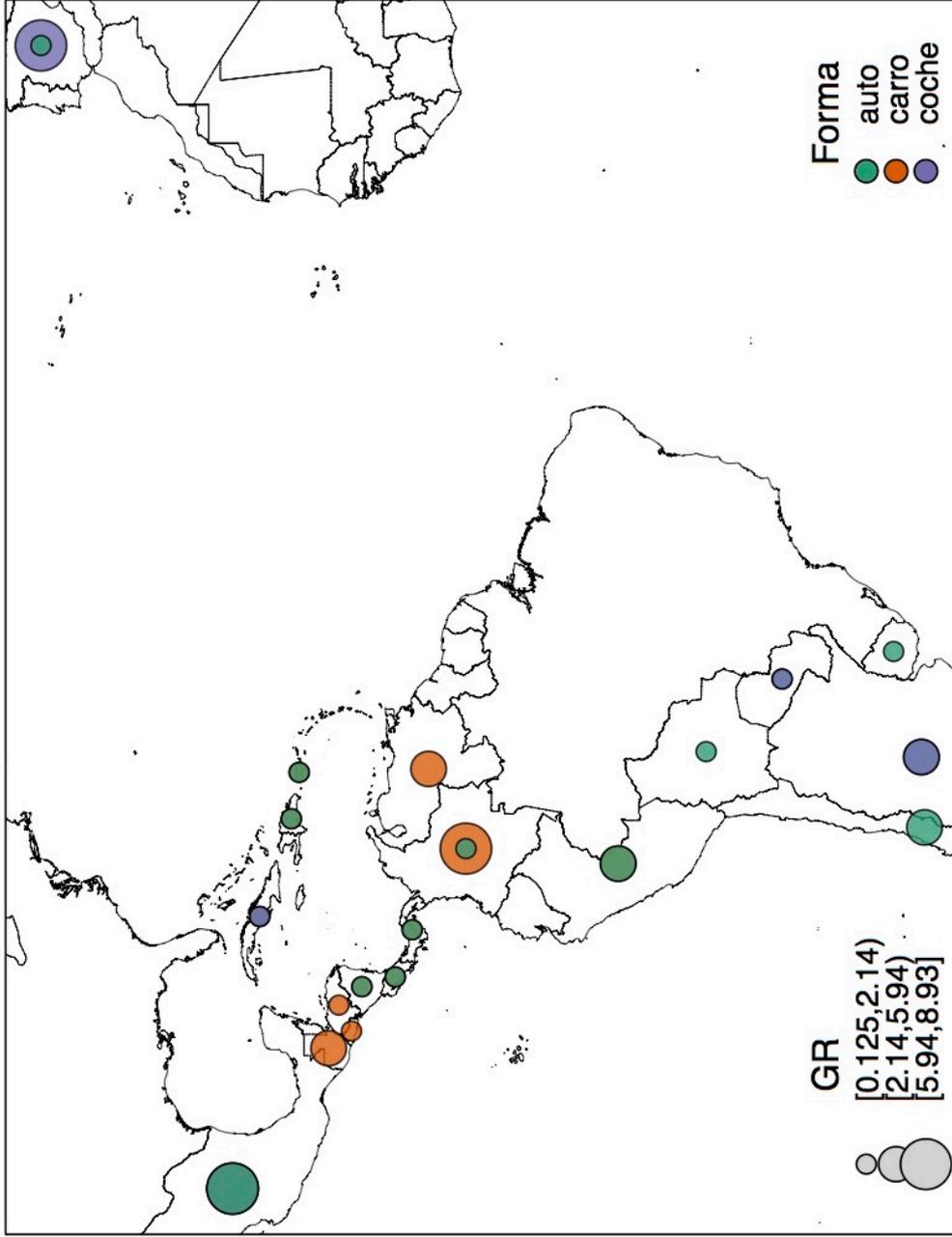
CAN



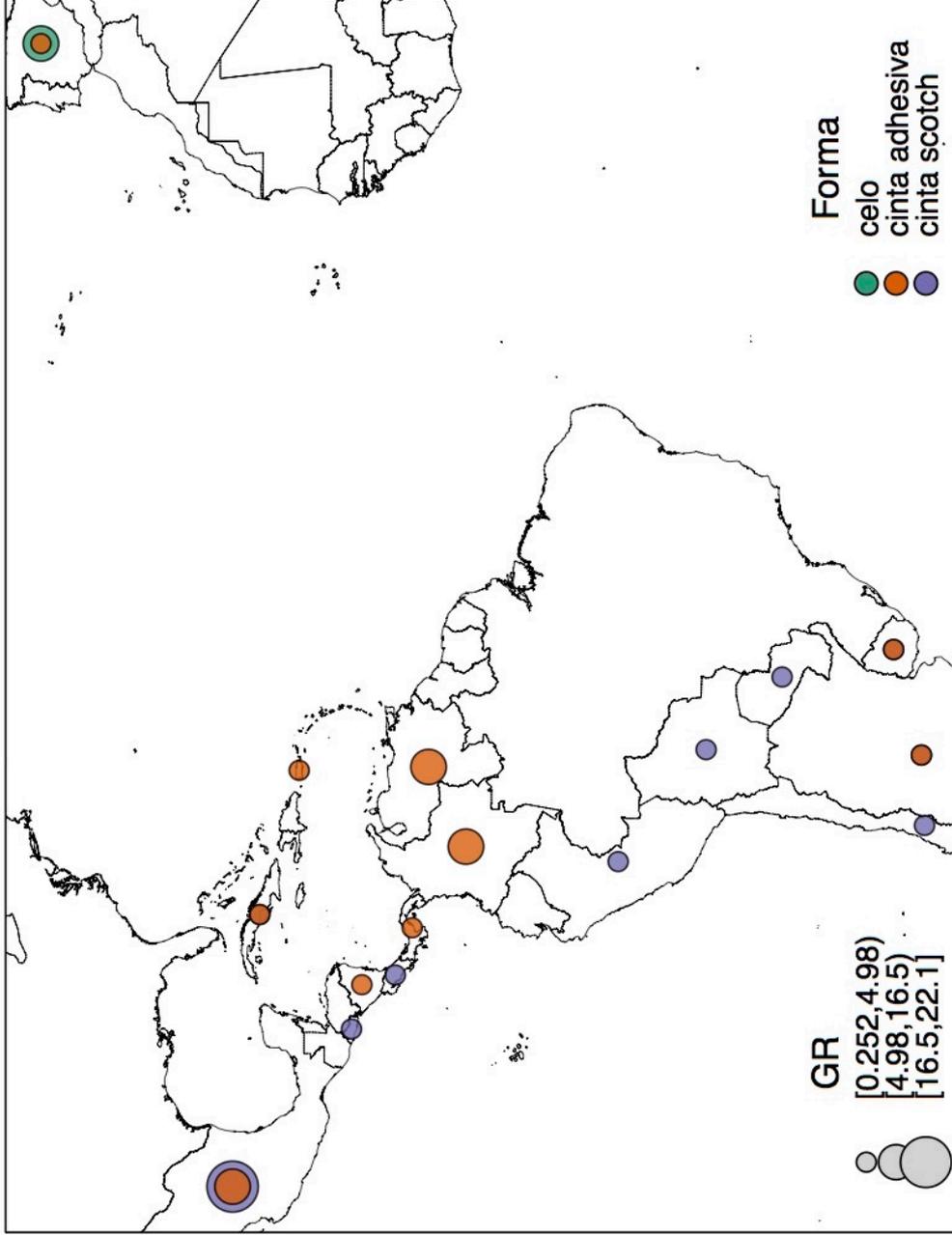
CANNOT REMEMBER



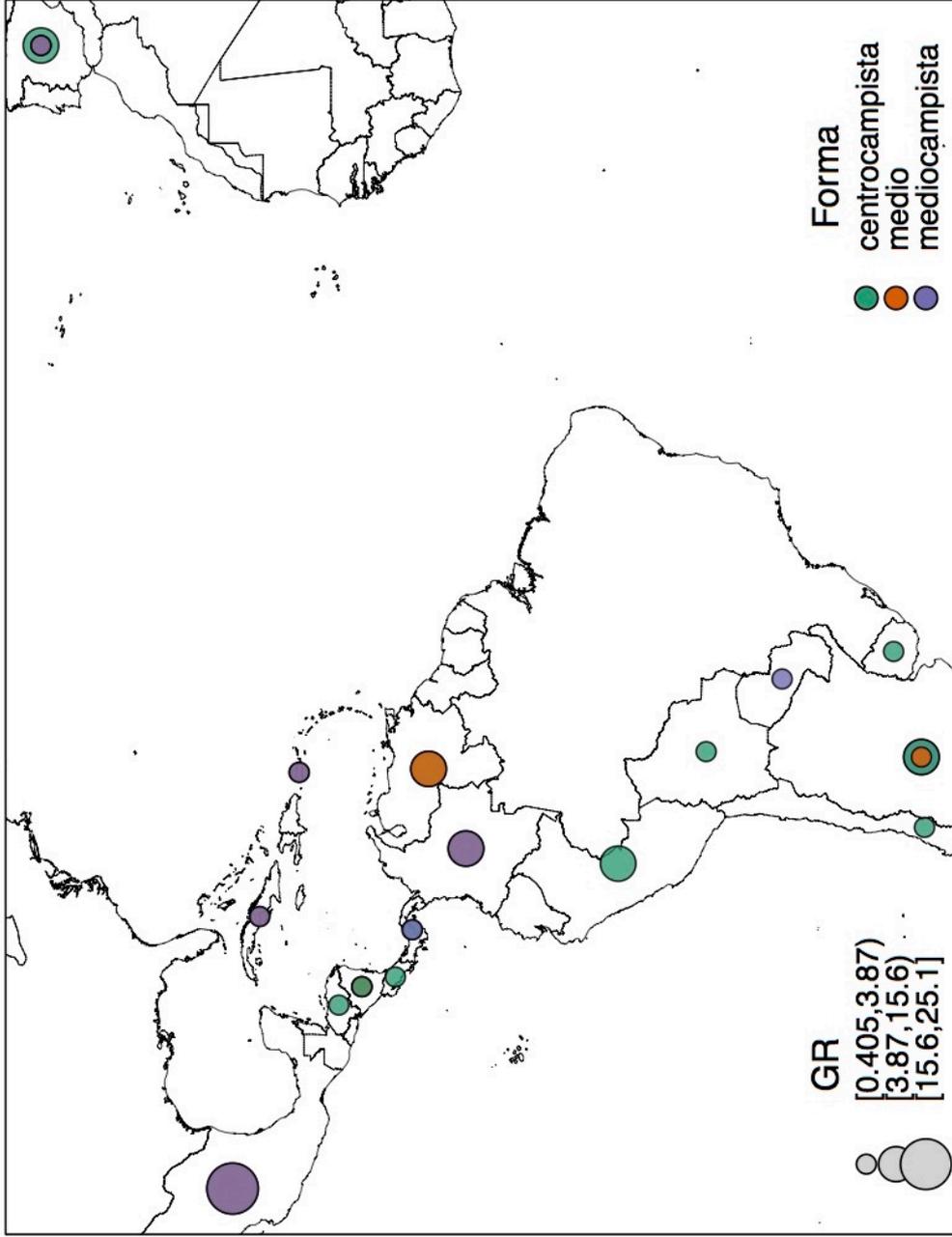
CAR



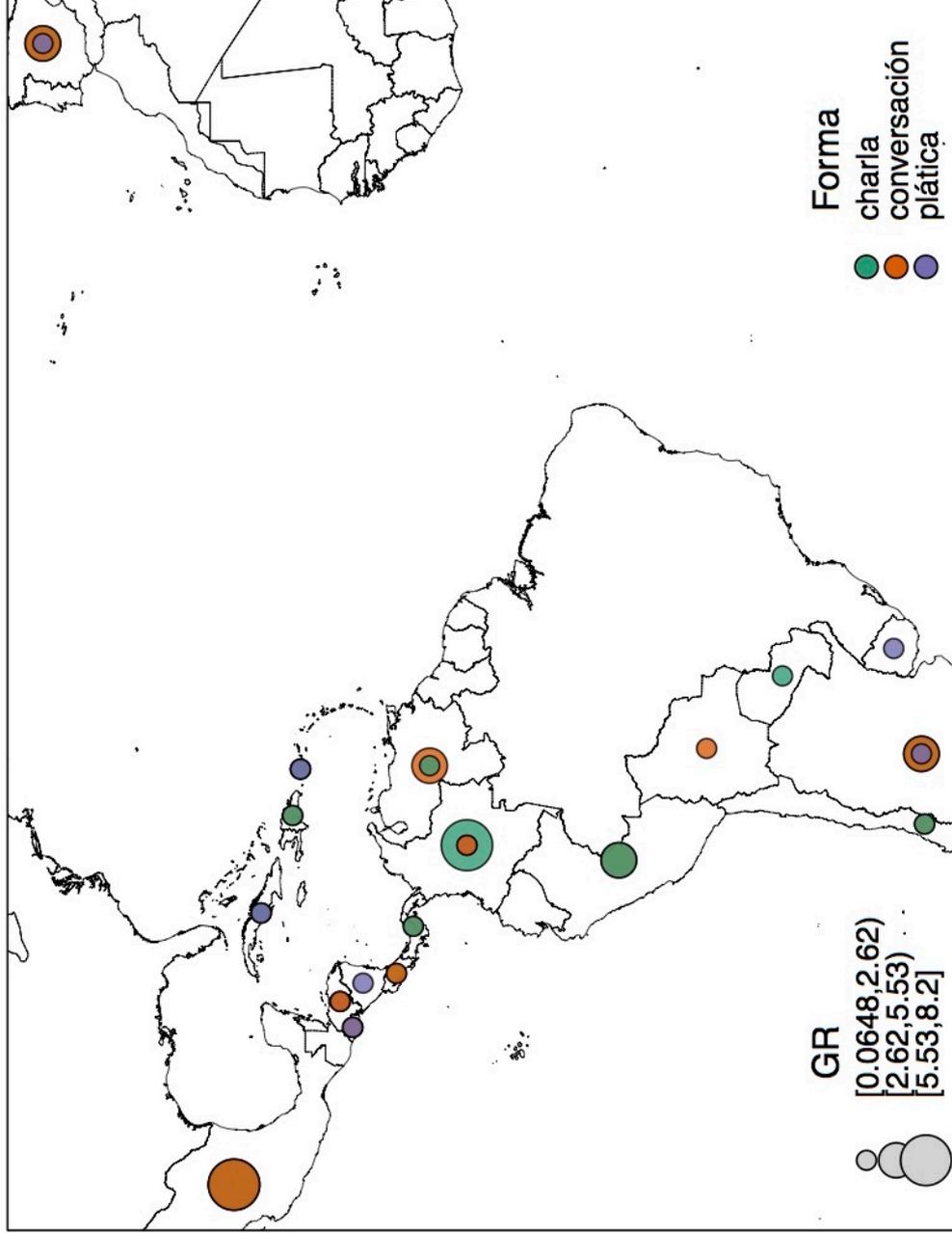
CELLOPHANE TAPE



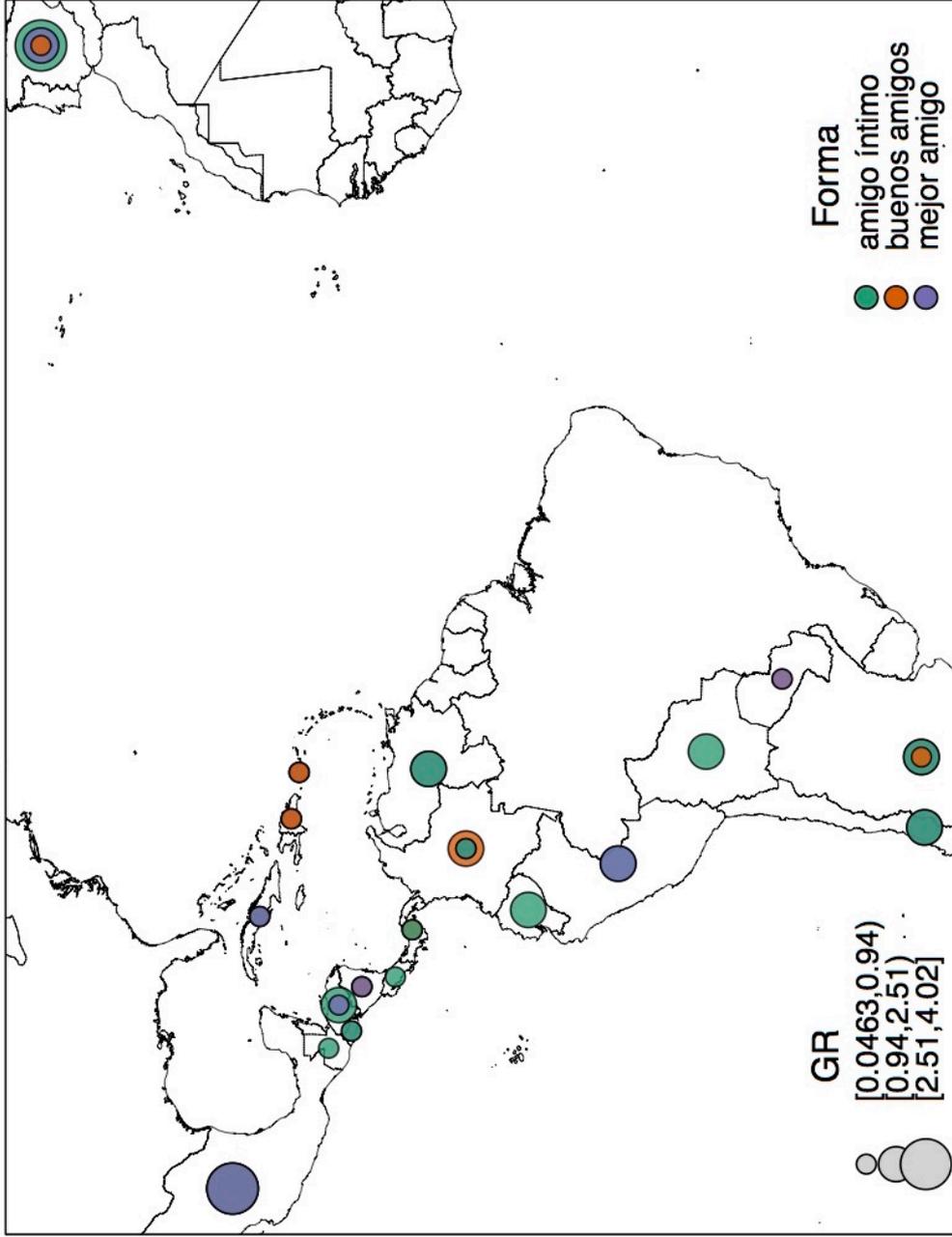
CENTER HALF



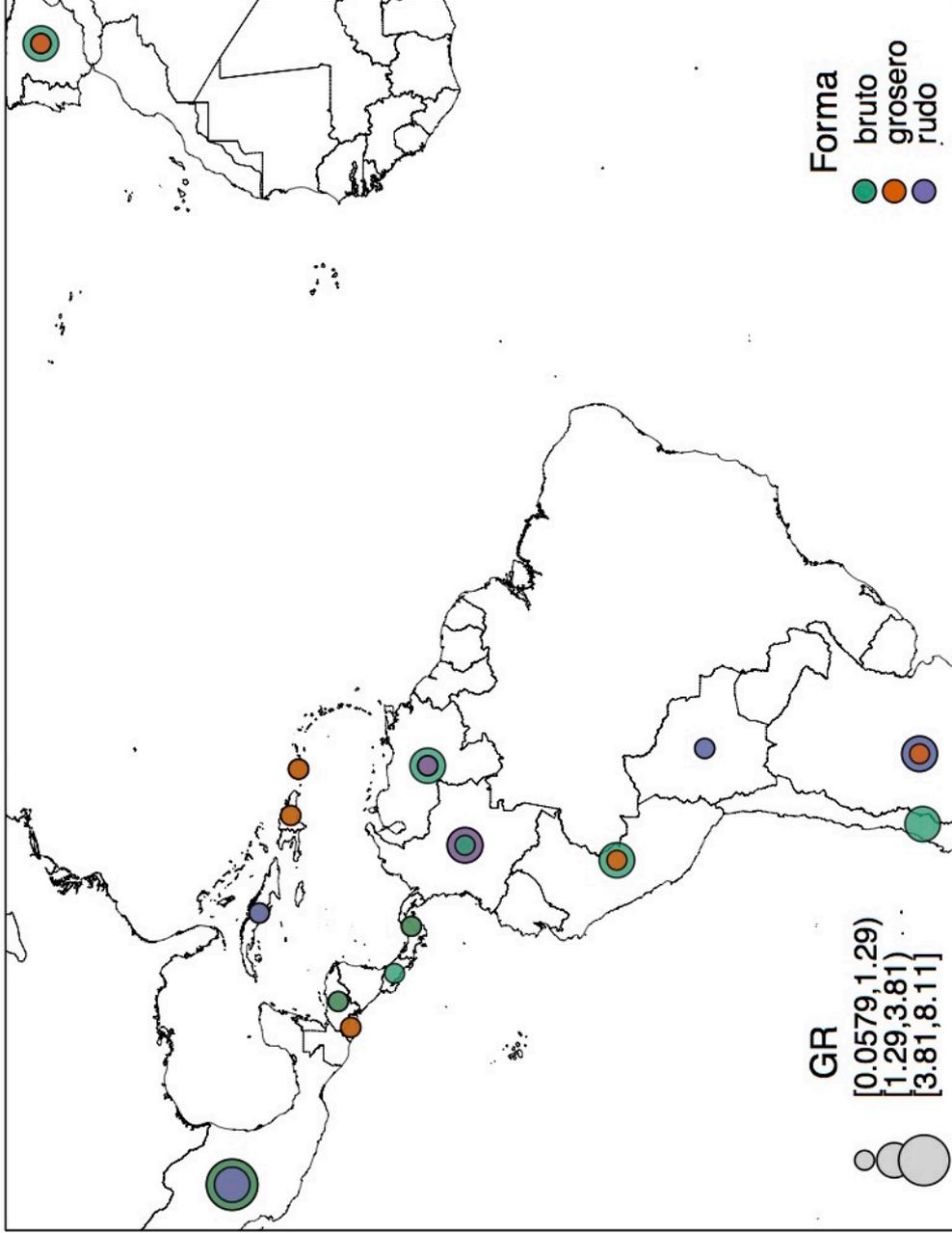
CHAT



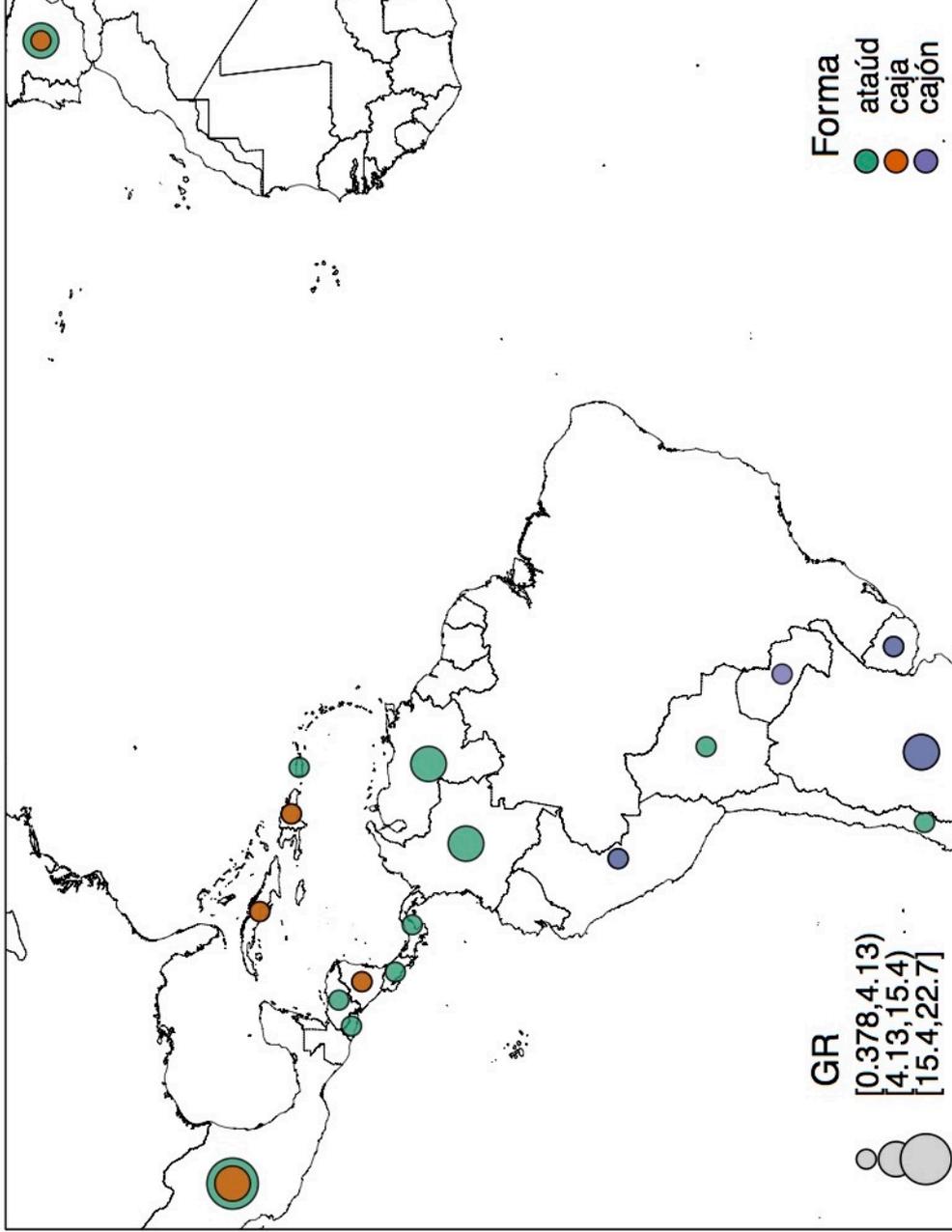
CLOSE FRIEND



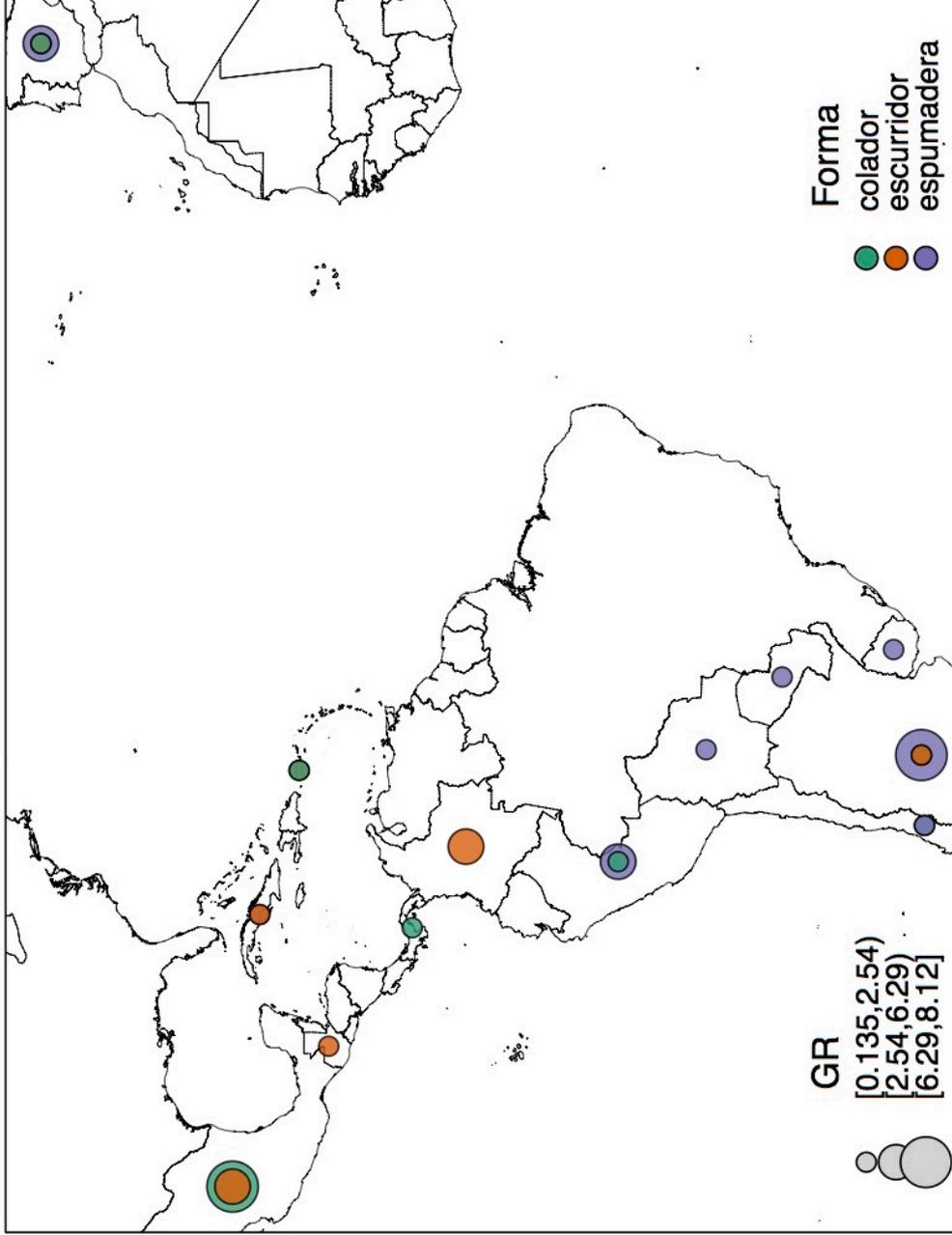
COARSE



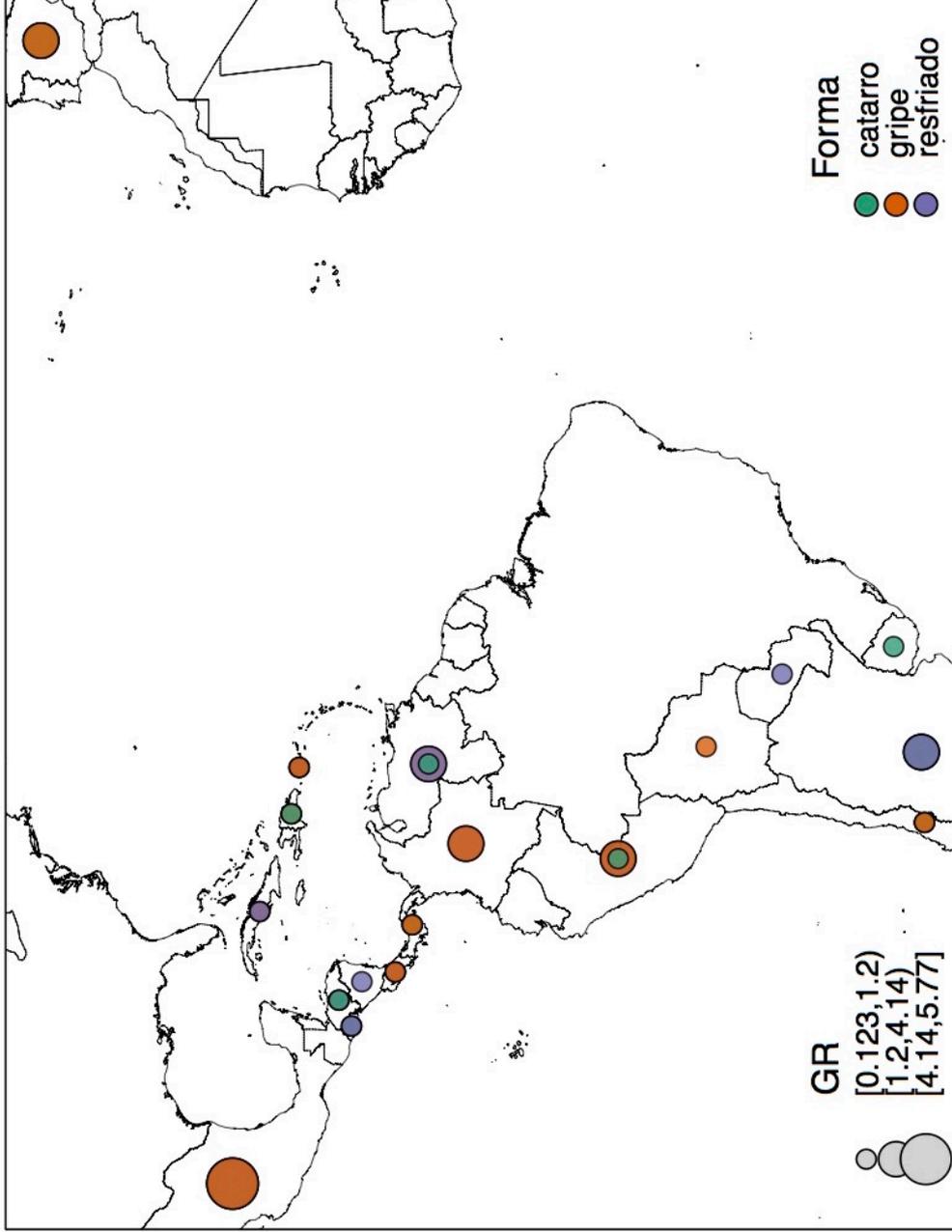
COFFIN



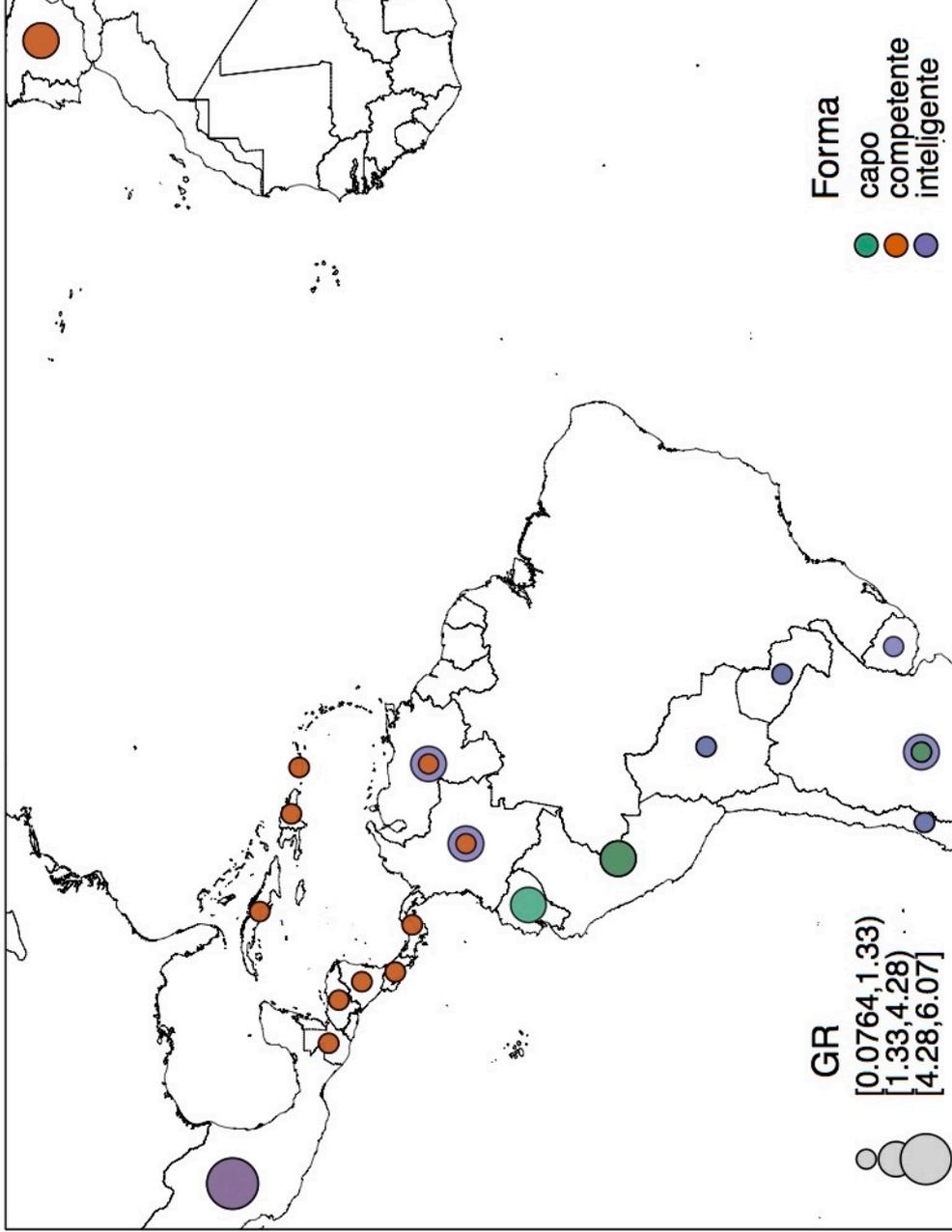
COLANDER



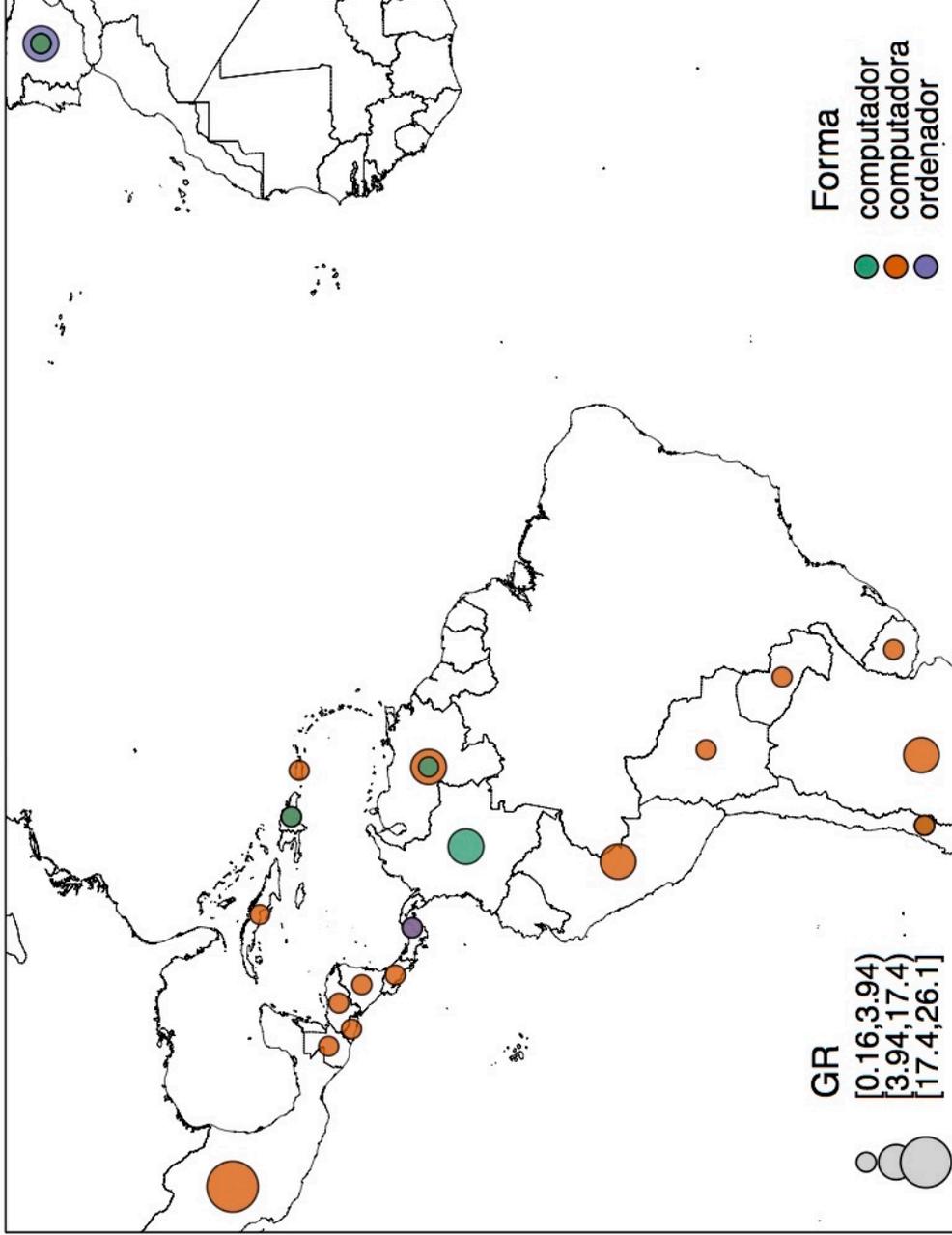
COLD



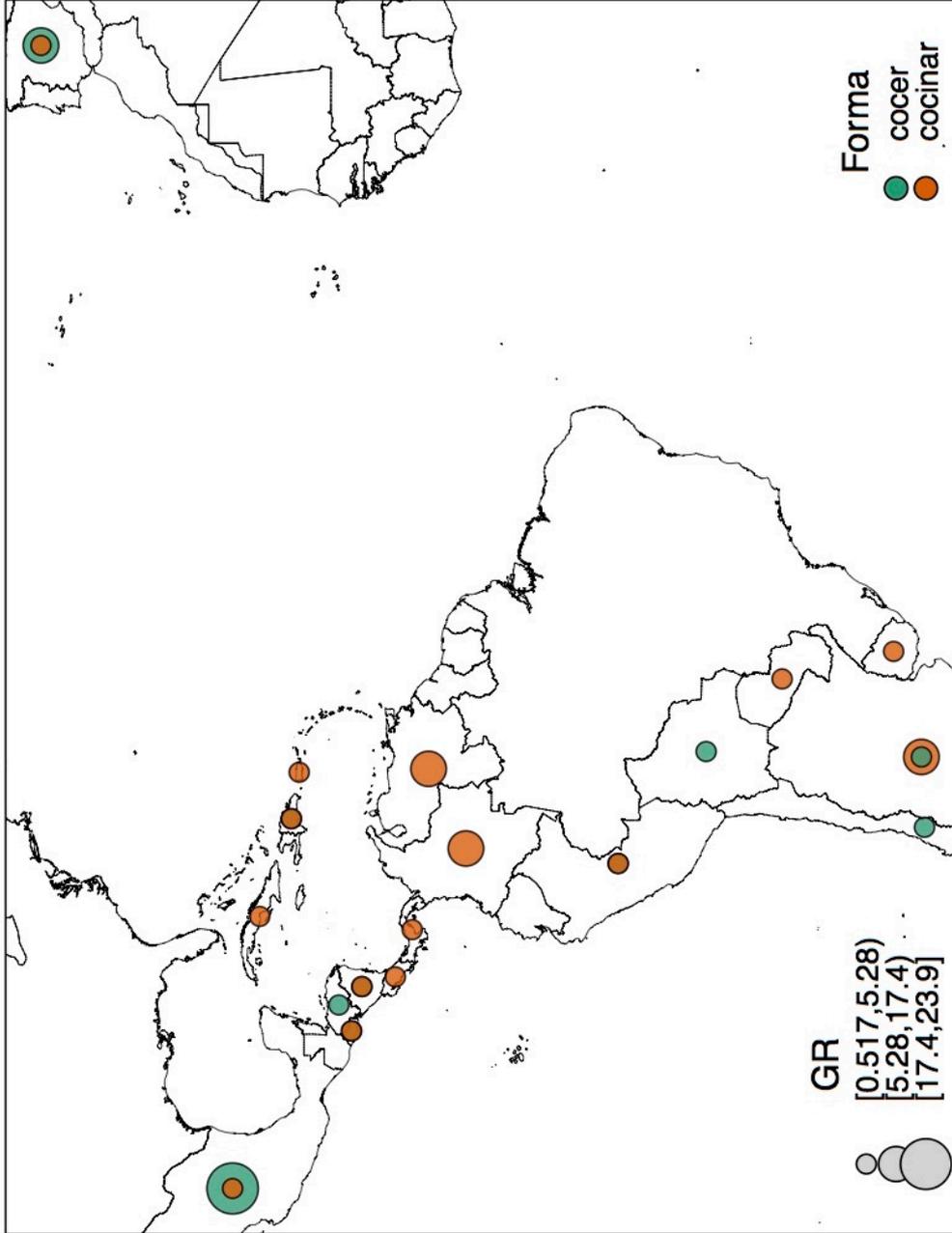
COMPETENT



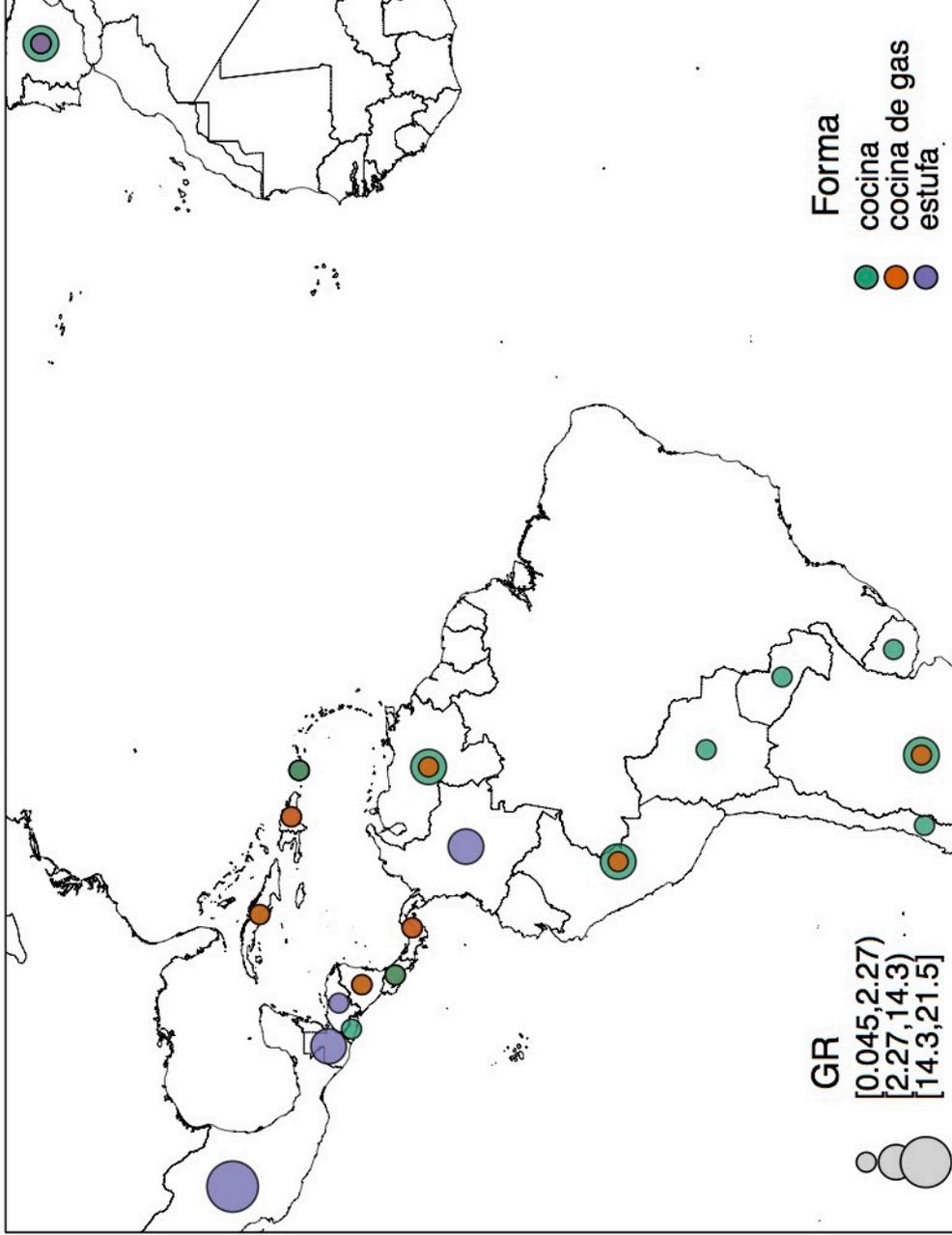
COMPUTER



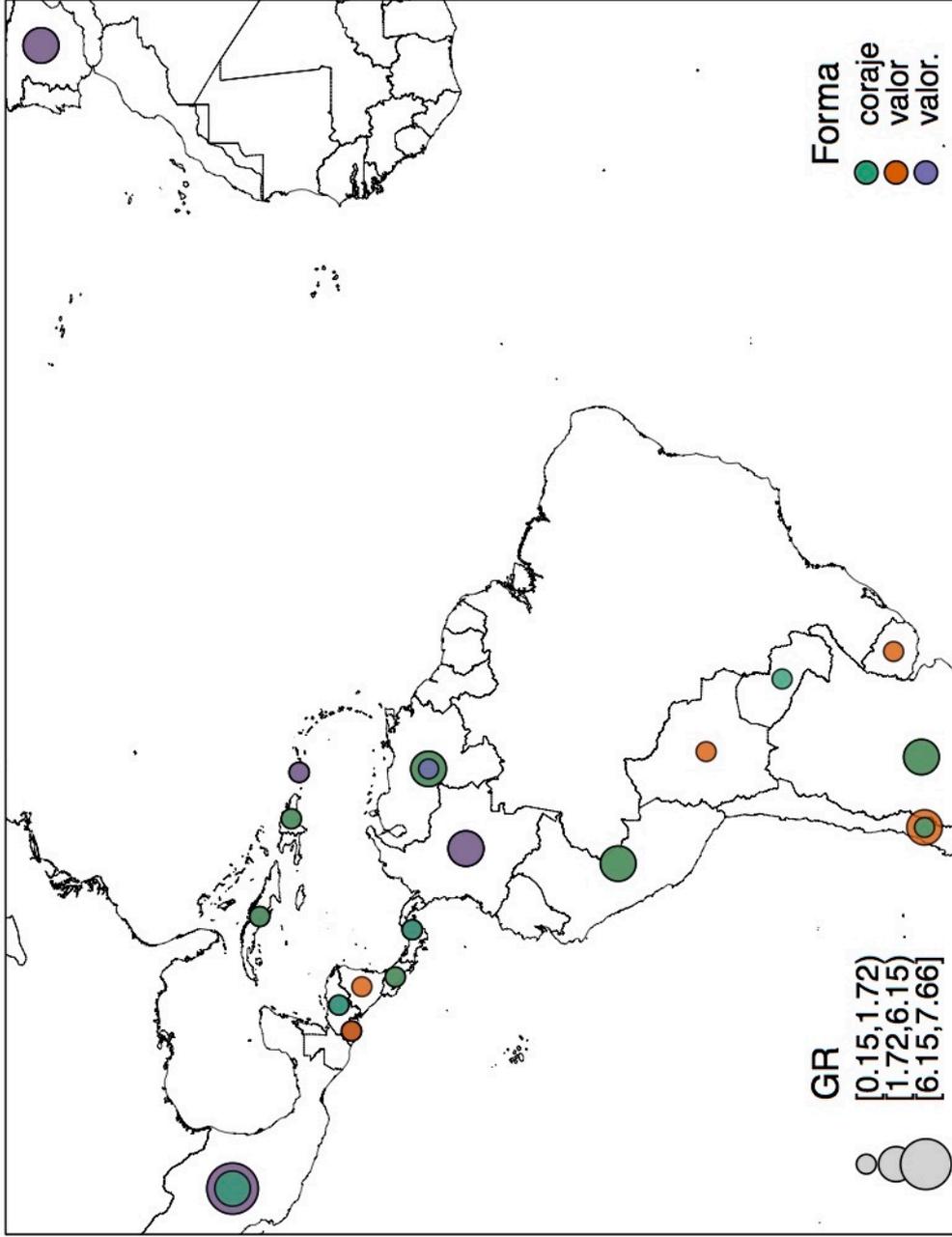
COOK



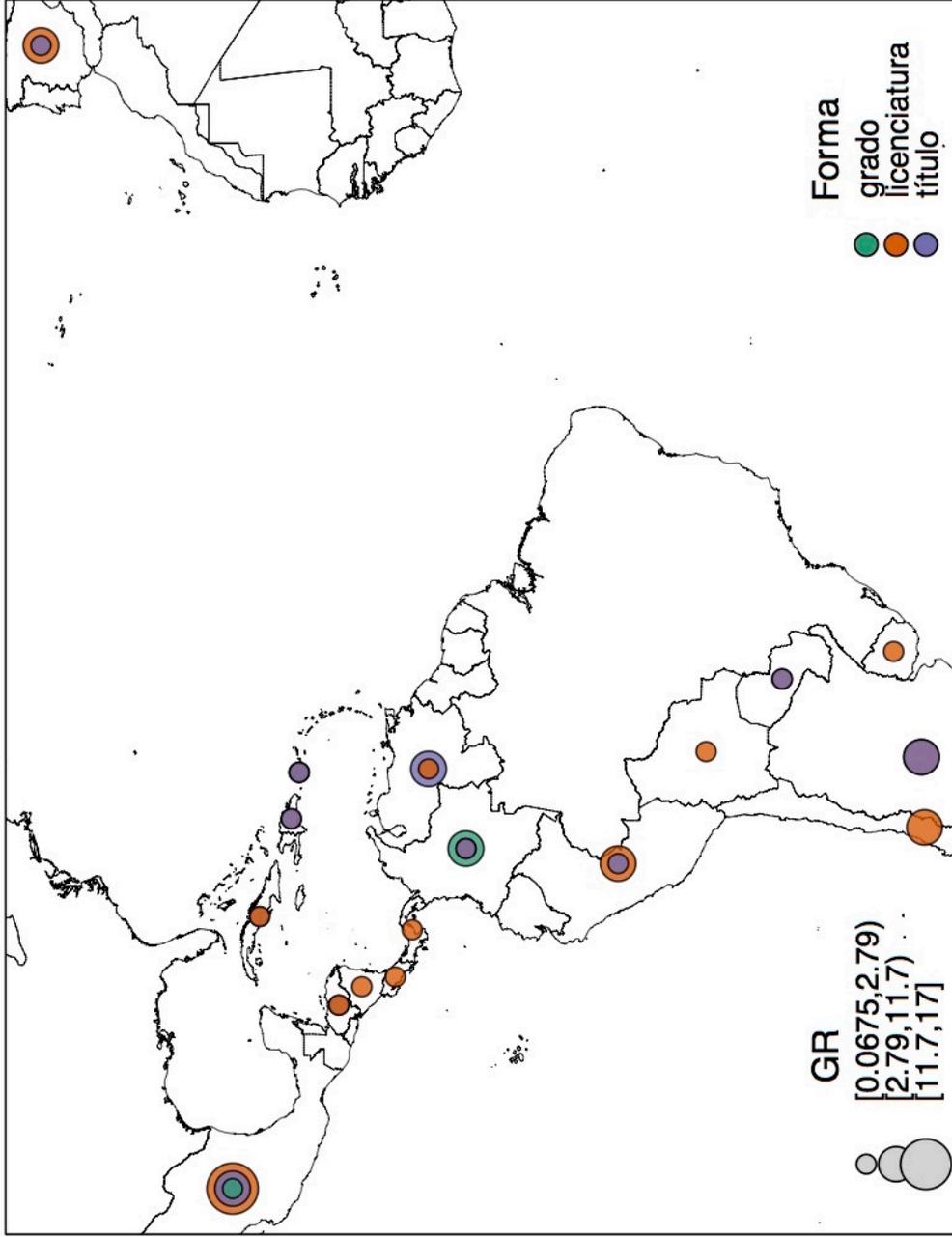
COOKER (STOVE)



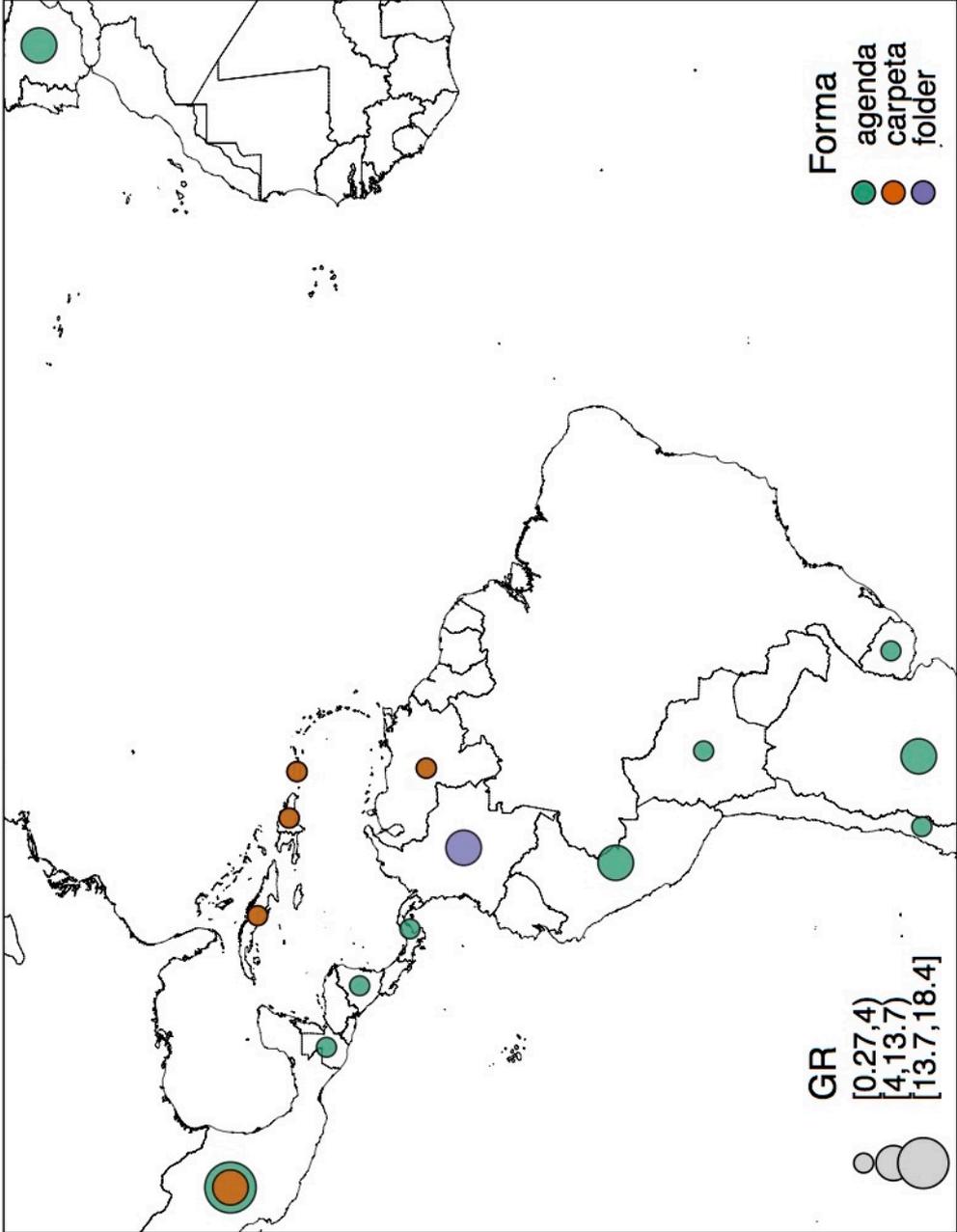
COURAGE



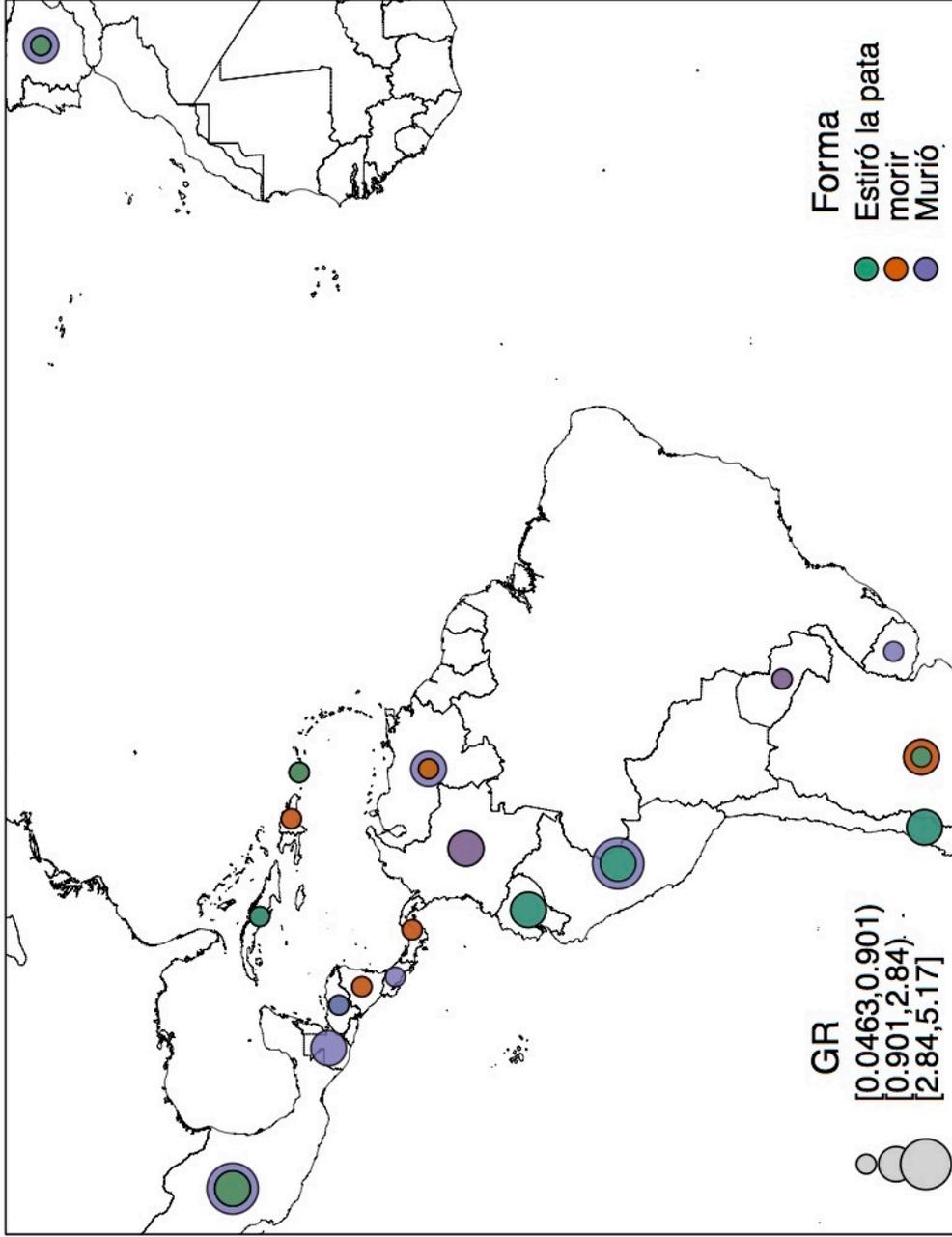
DEGREE



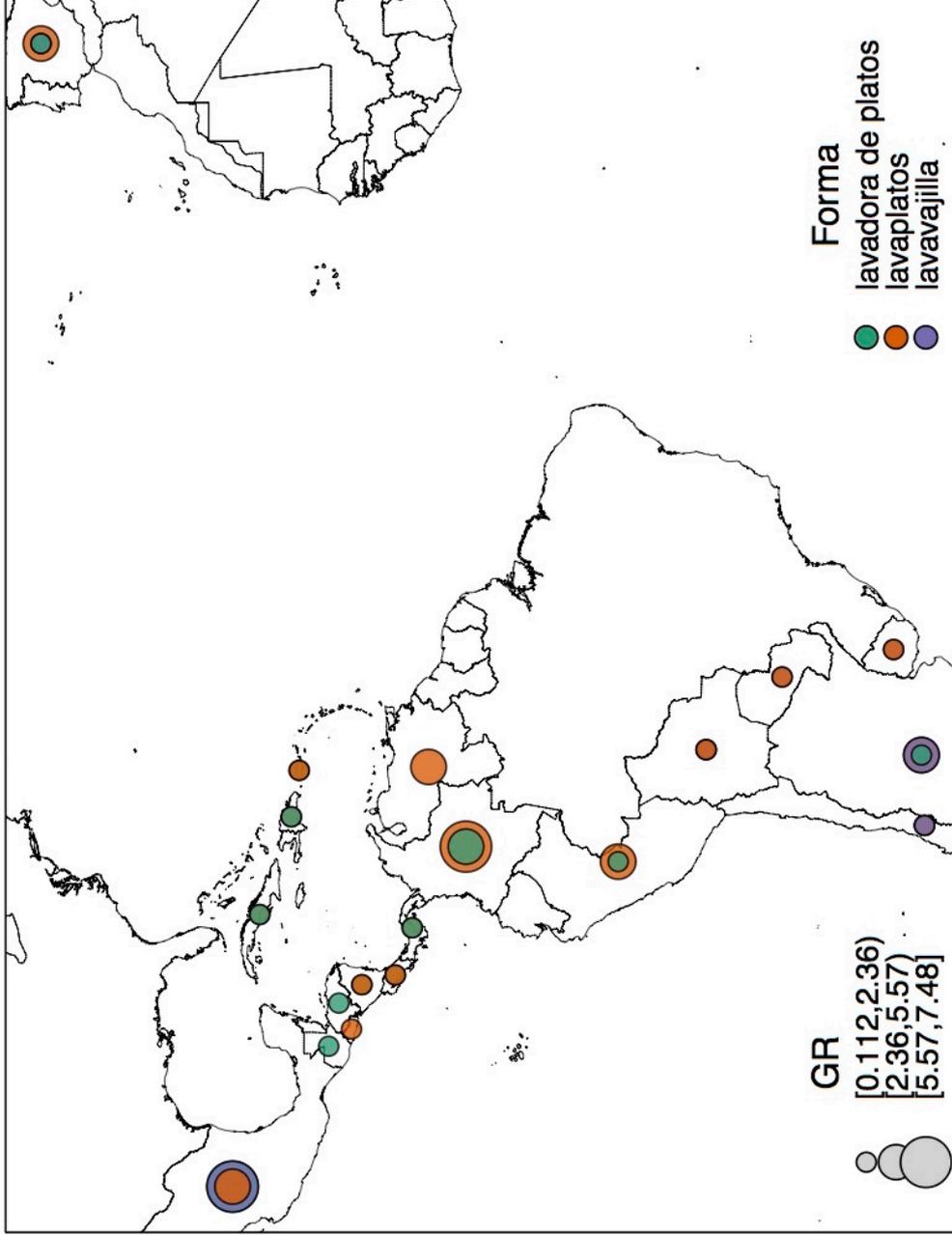
DIARY



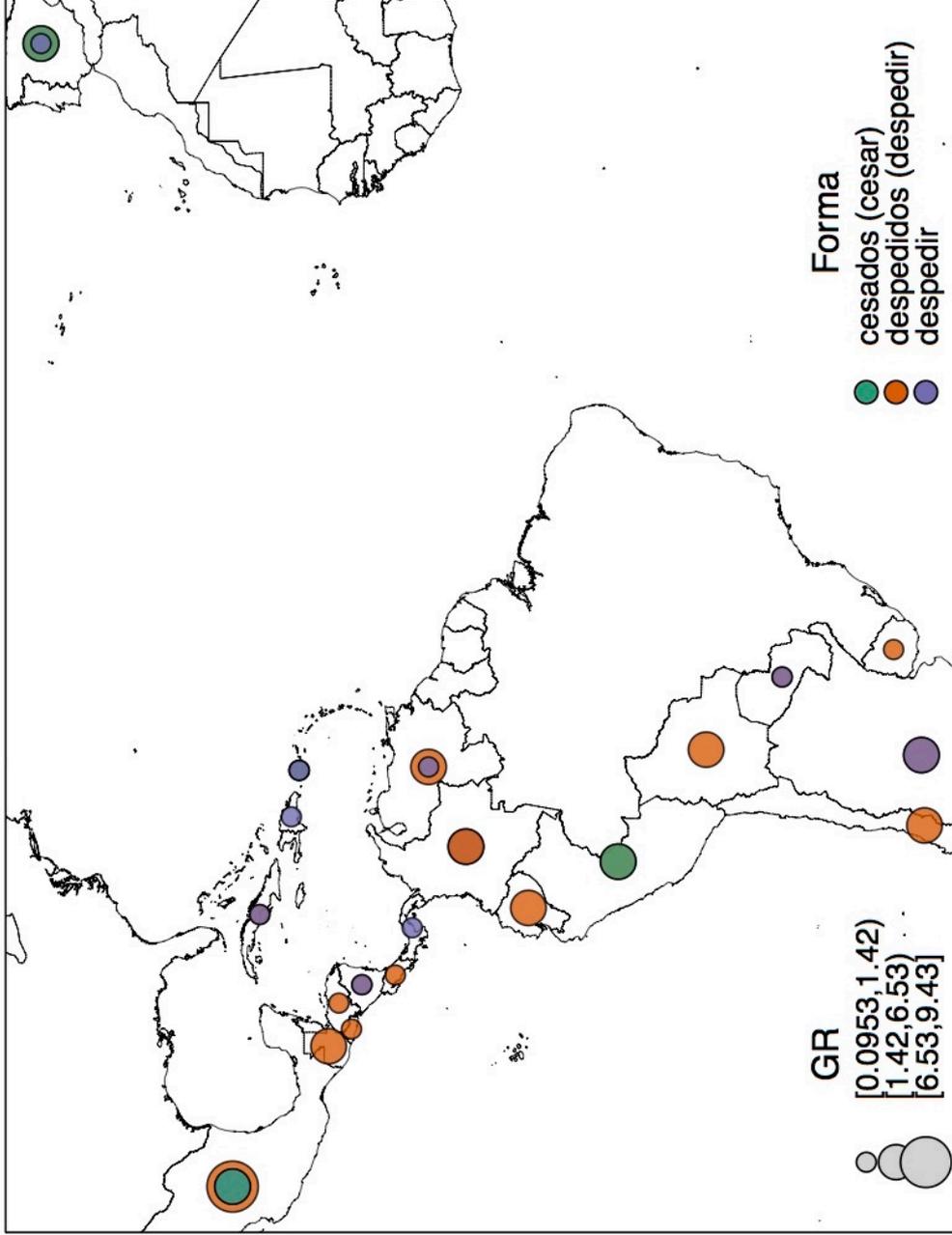
DIE



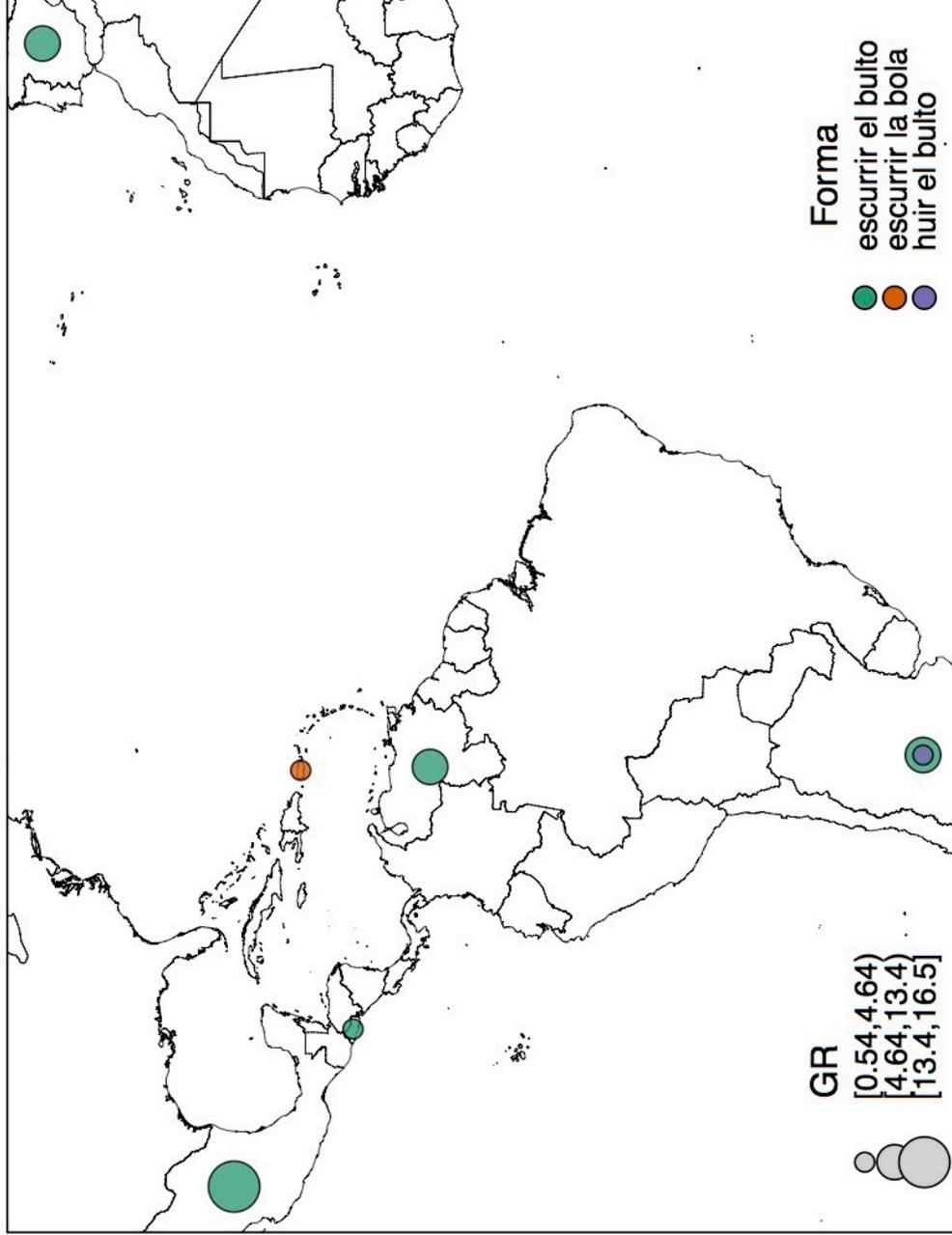
DISHWASHER



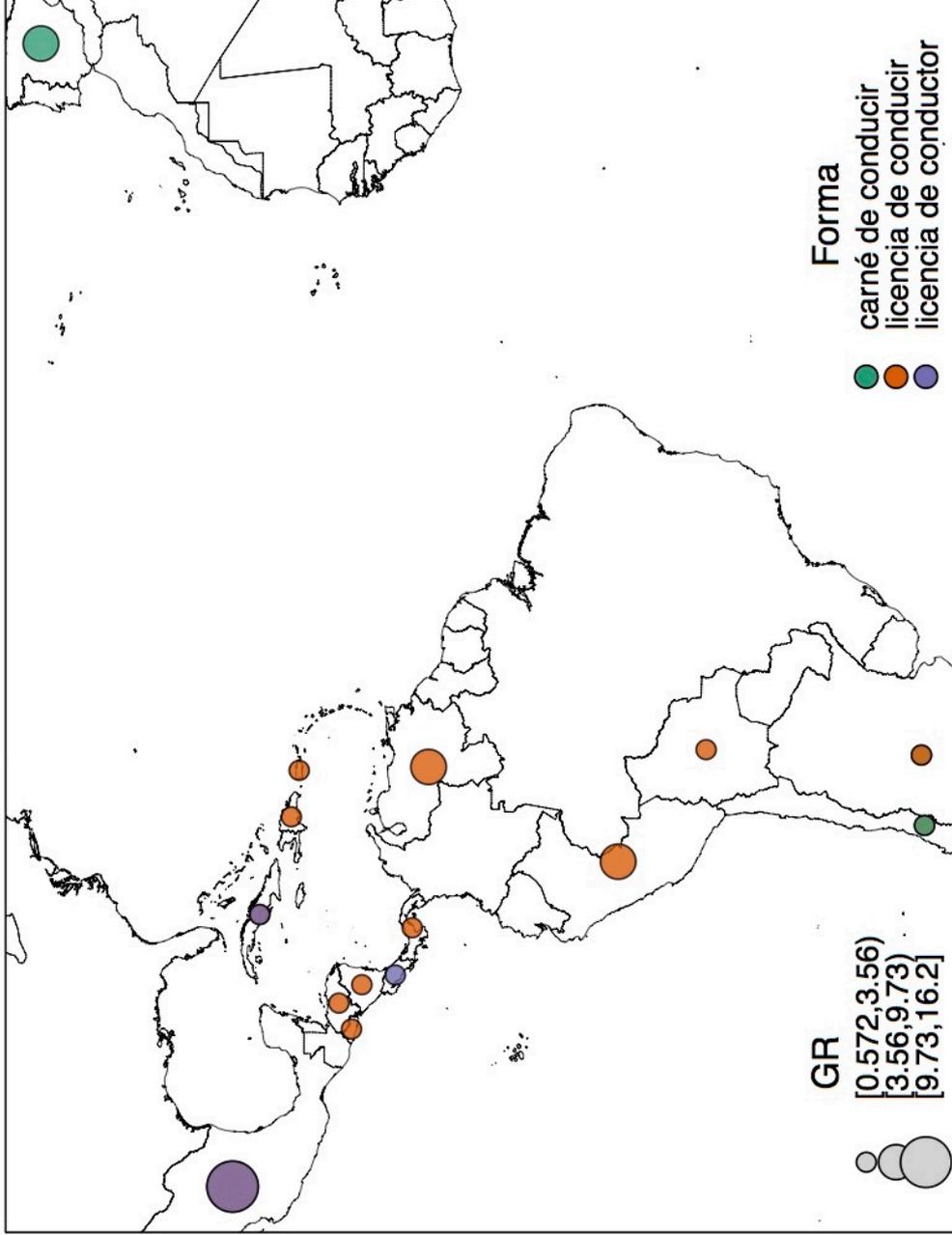
DISMISS



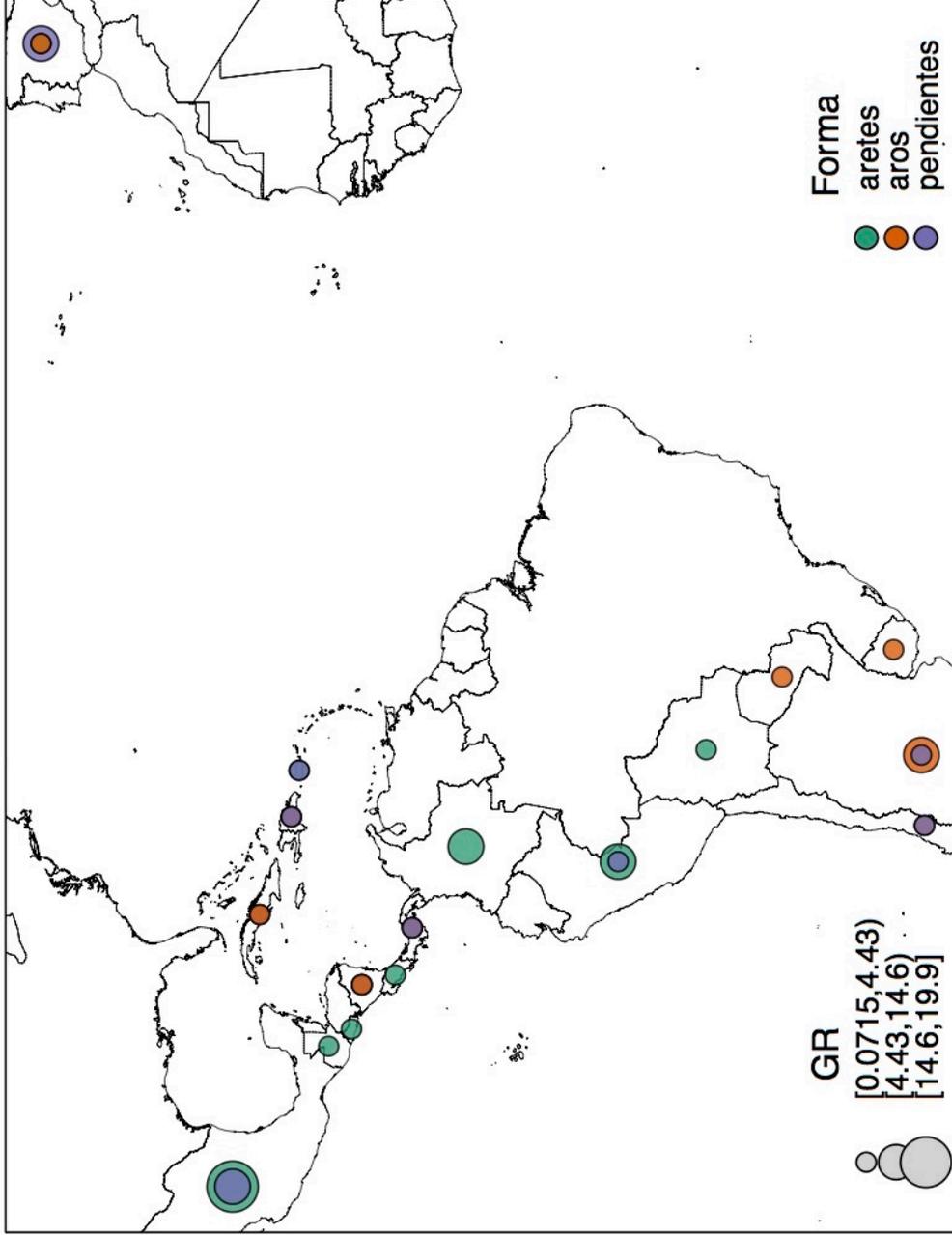
DODGE THE ISSUE



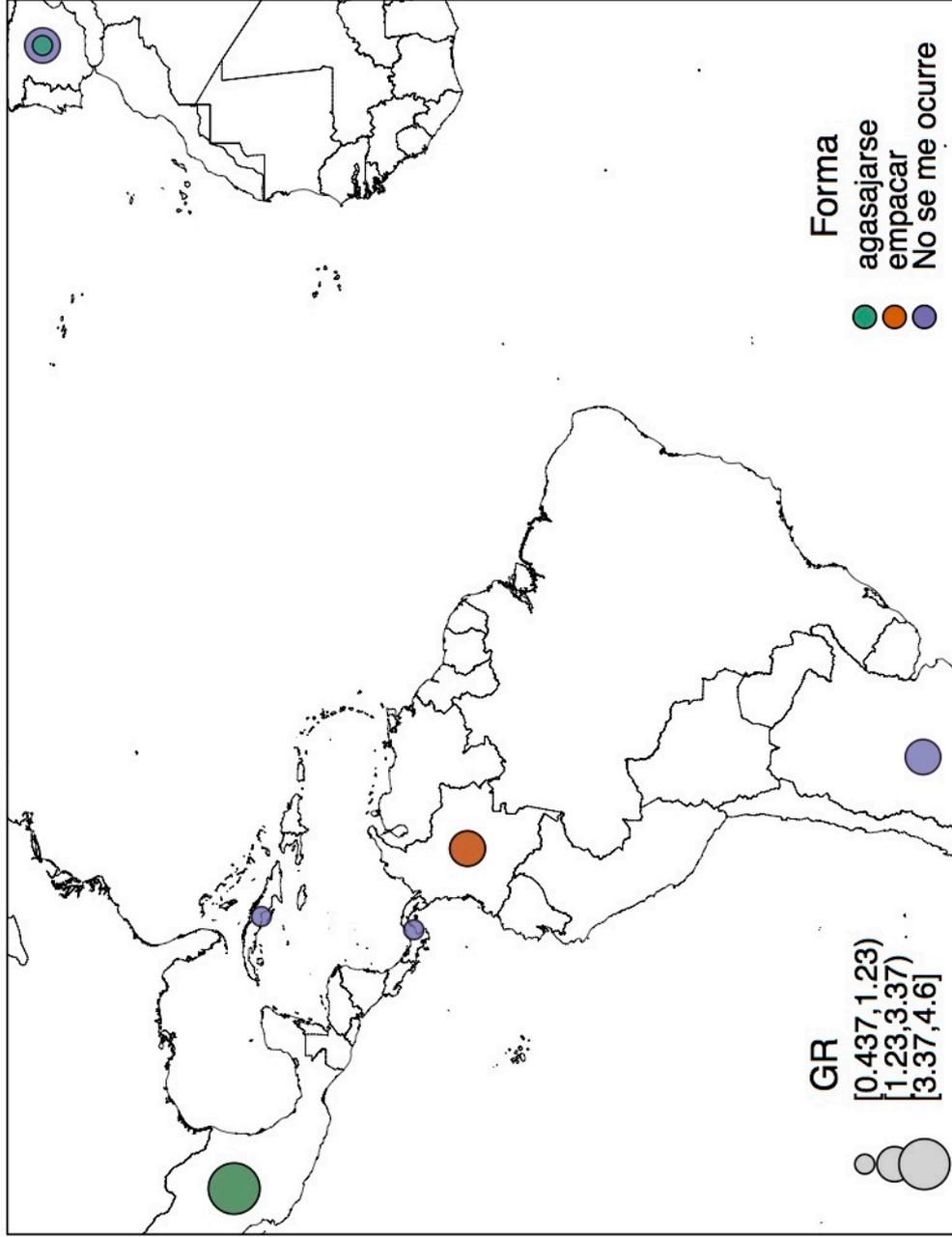
DRIVING LICENCE



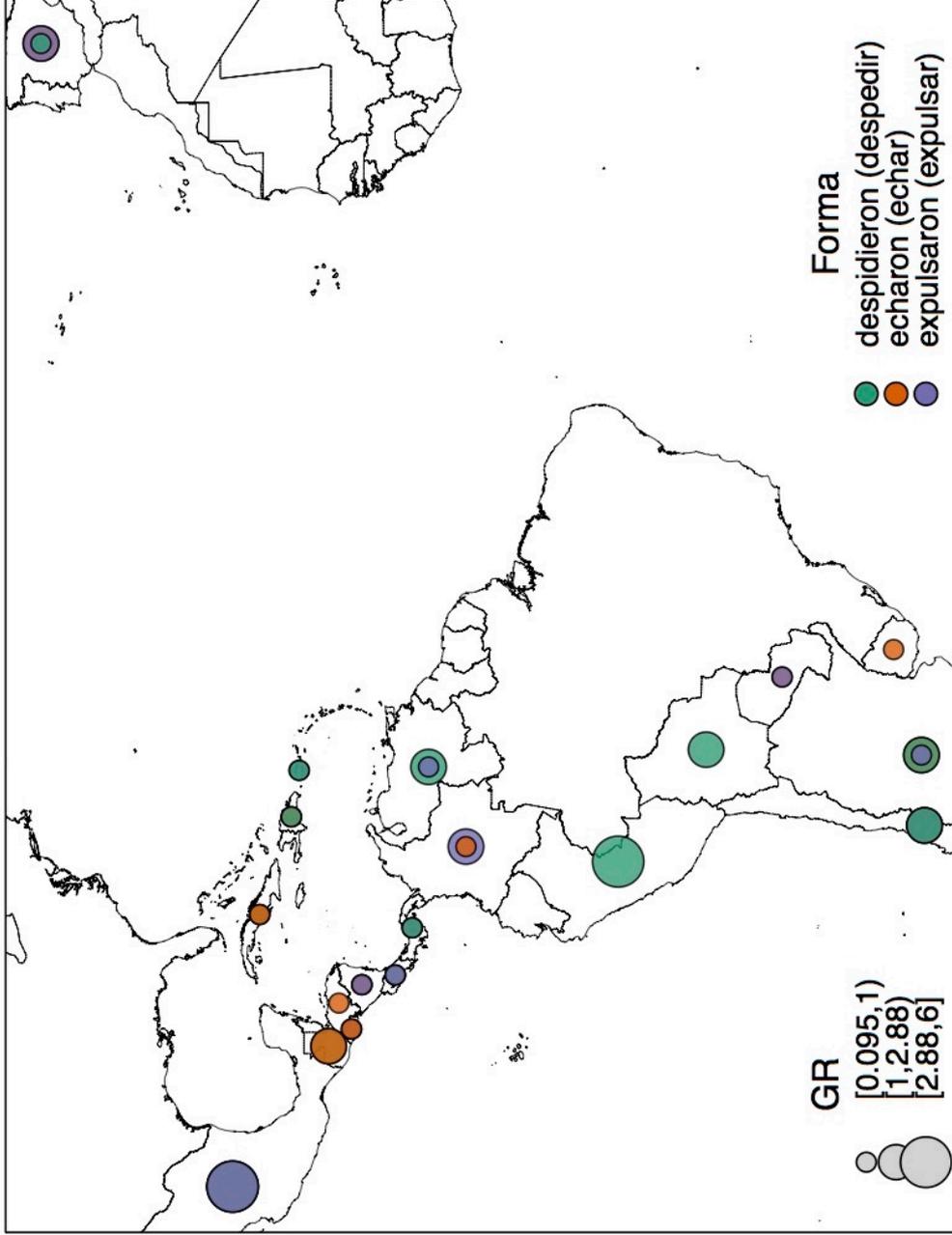
EARRINGS



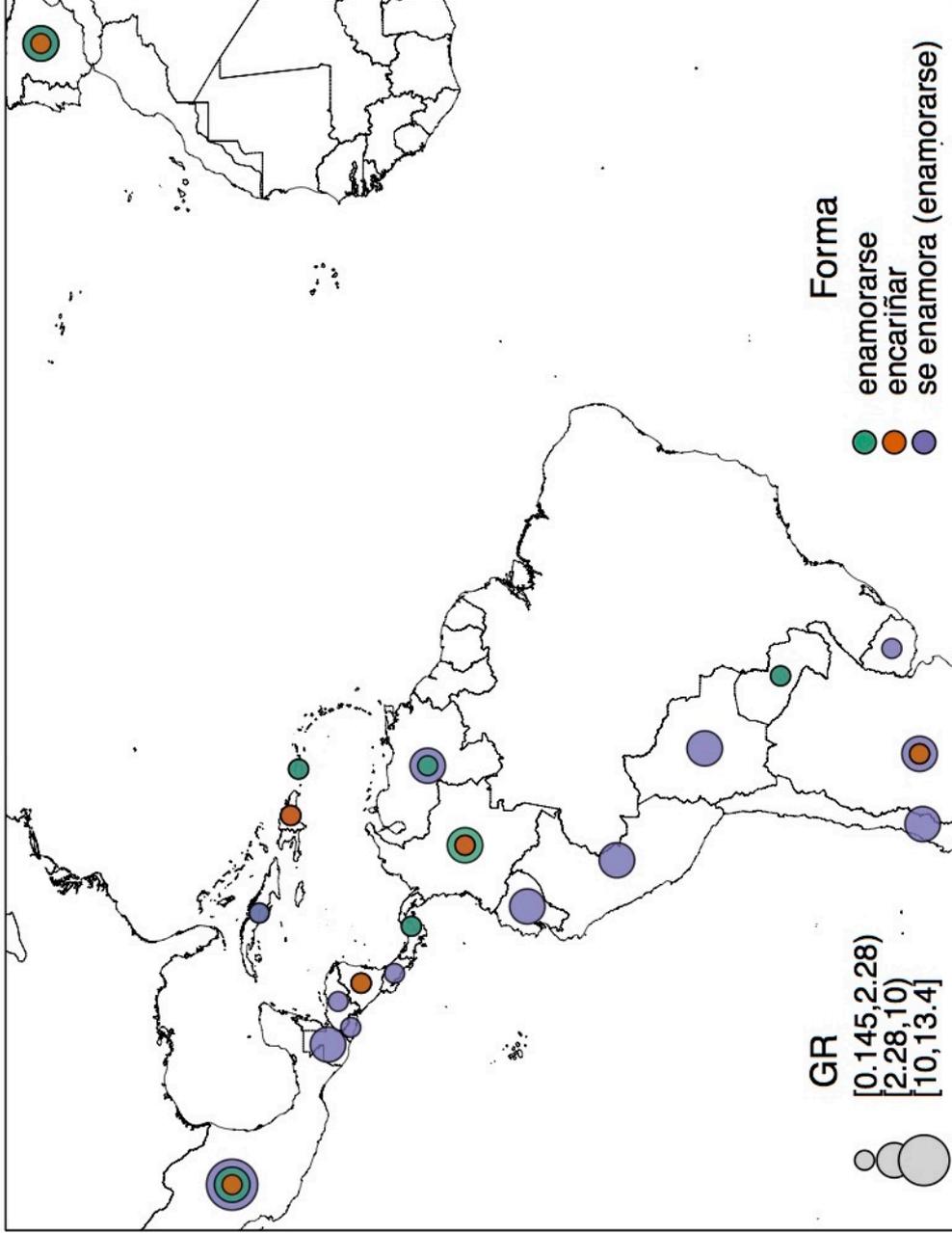
EAT TOO MUCH



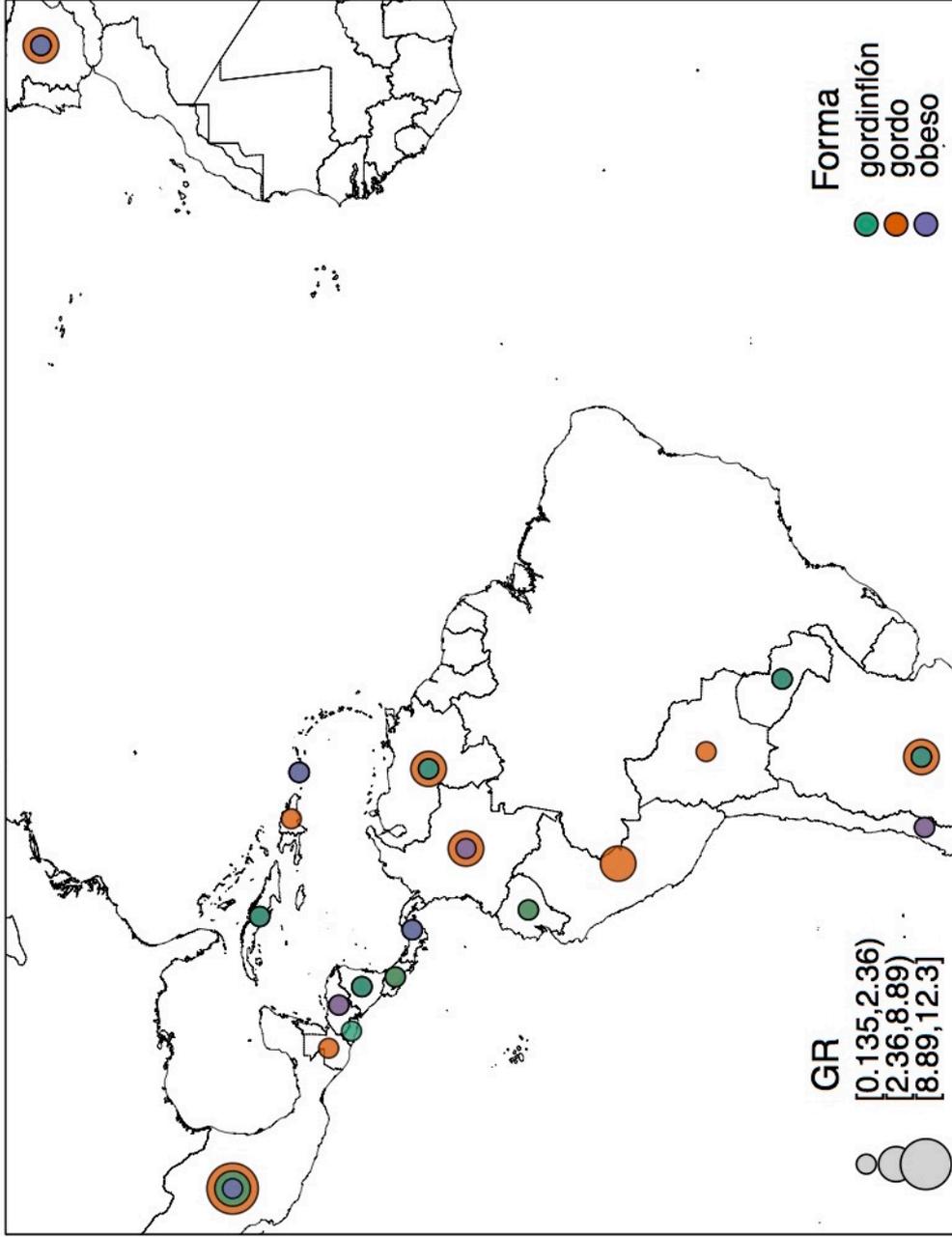
EXPEL



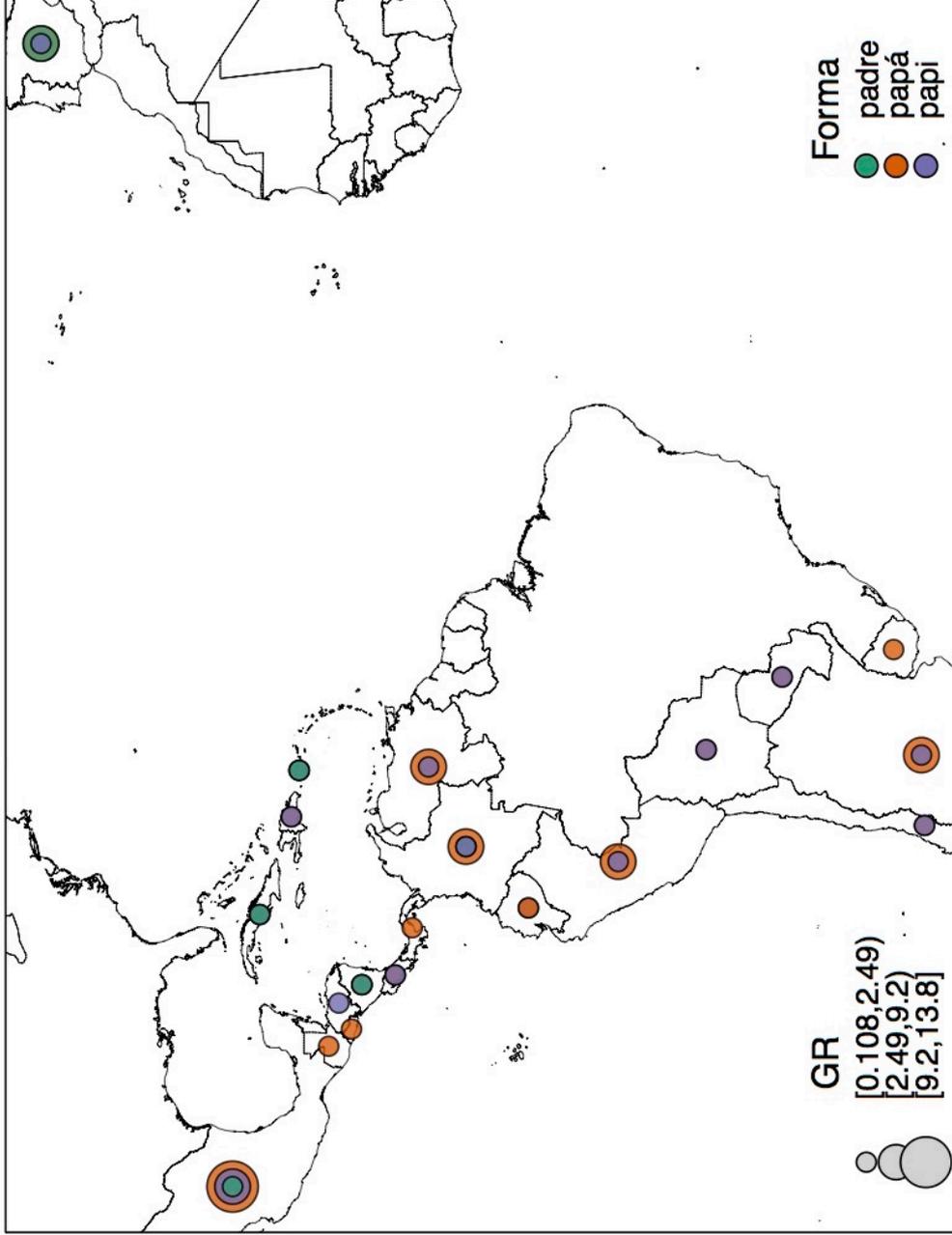
FALL IN LOVE



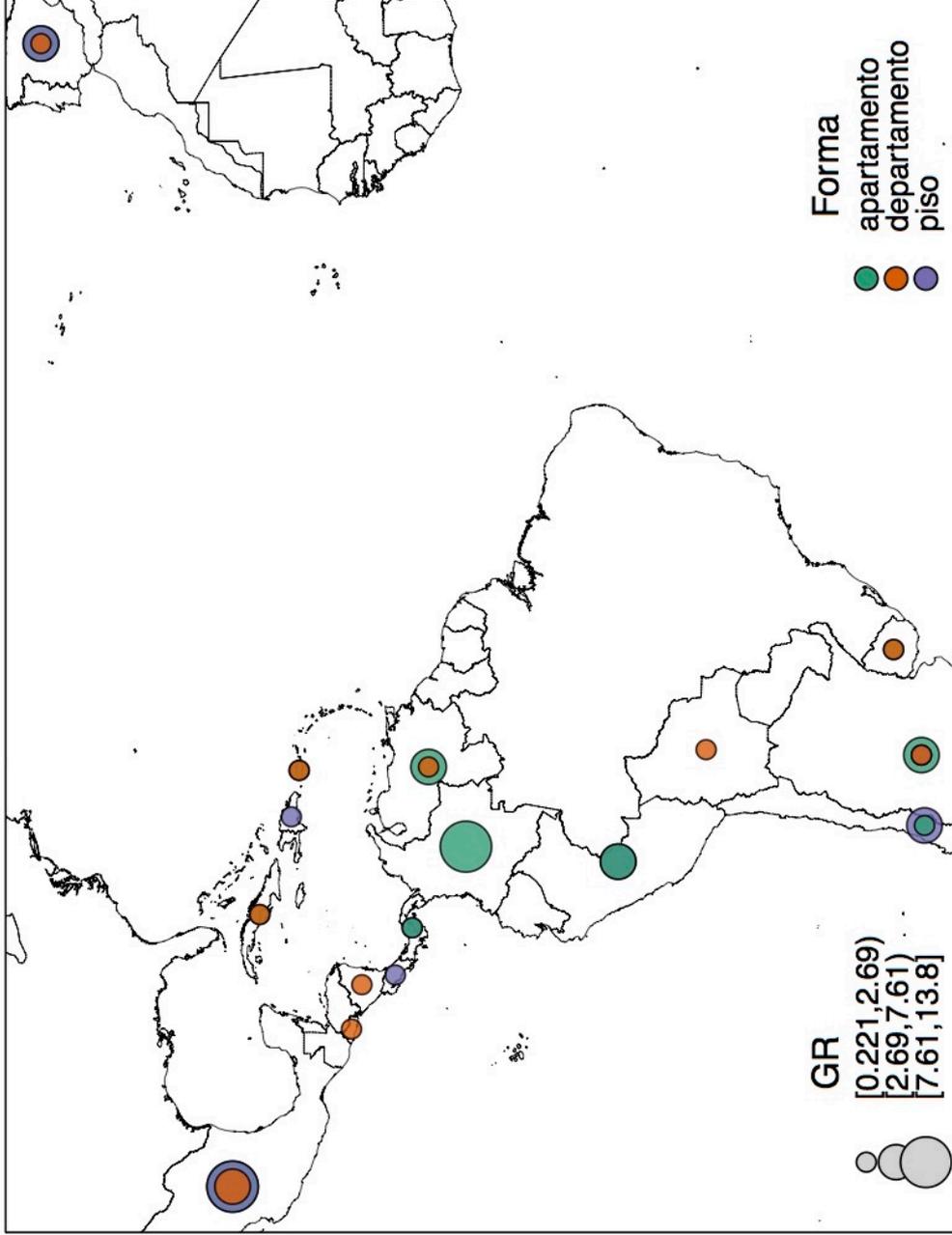
FAT (PLUMP)



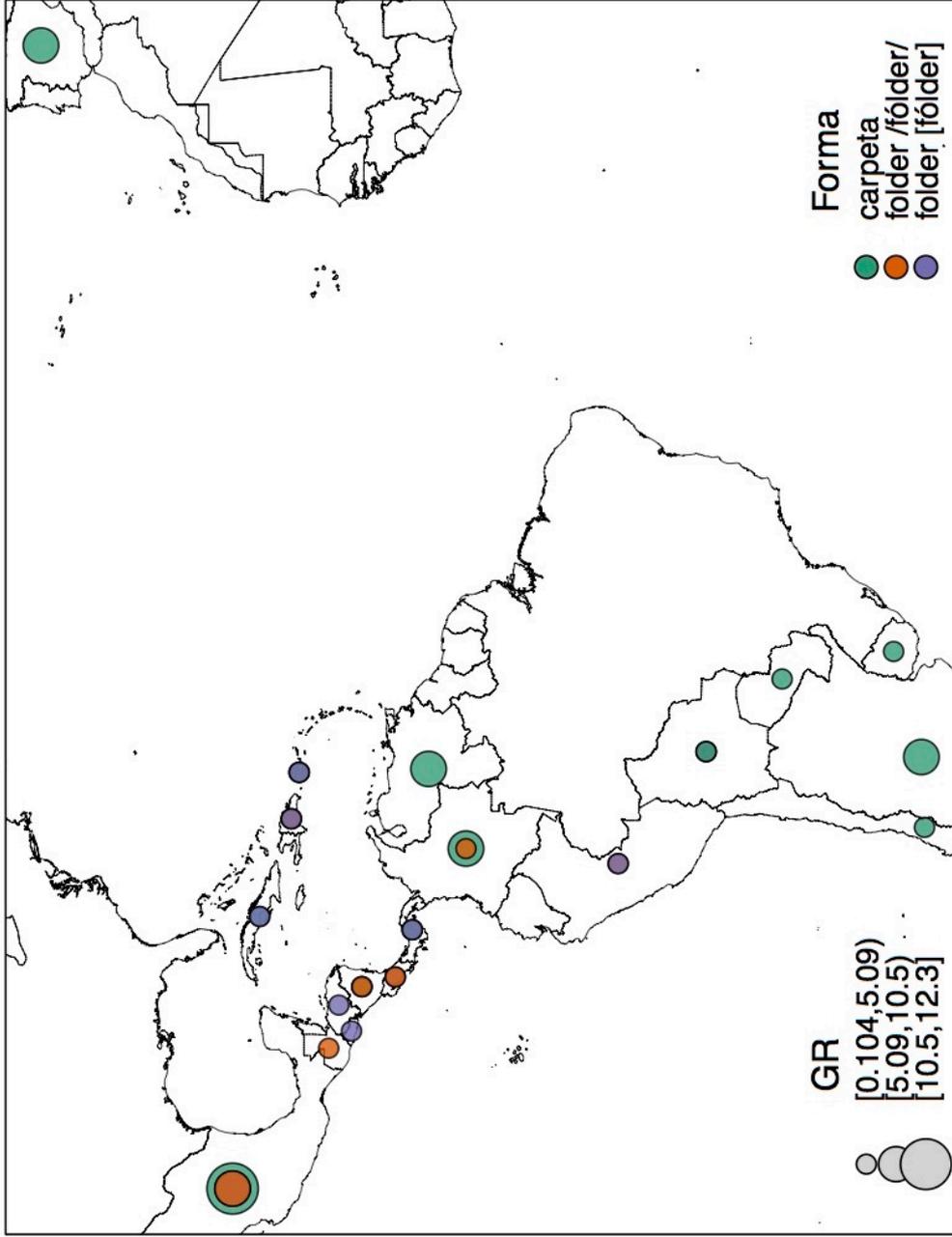
FATHER (DADDY)



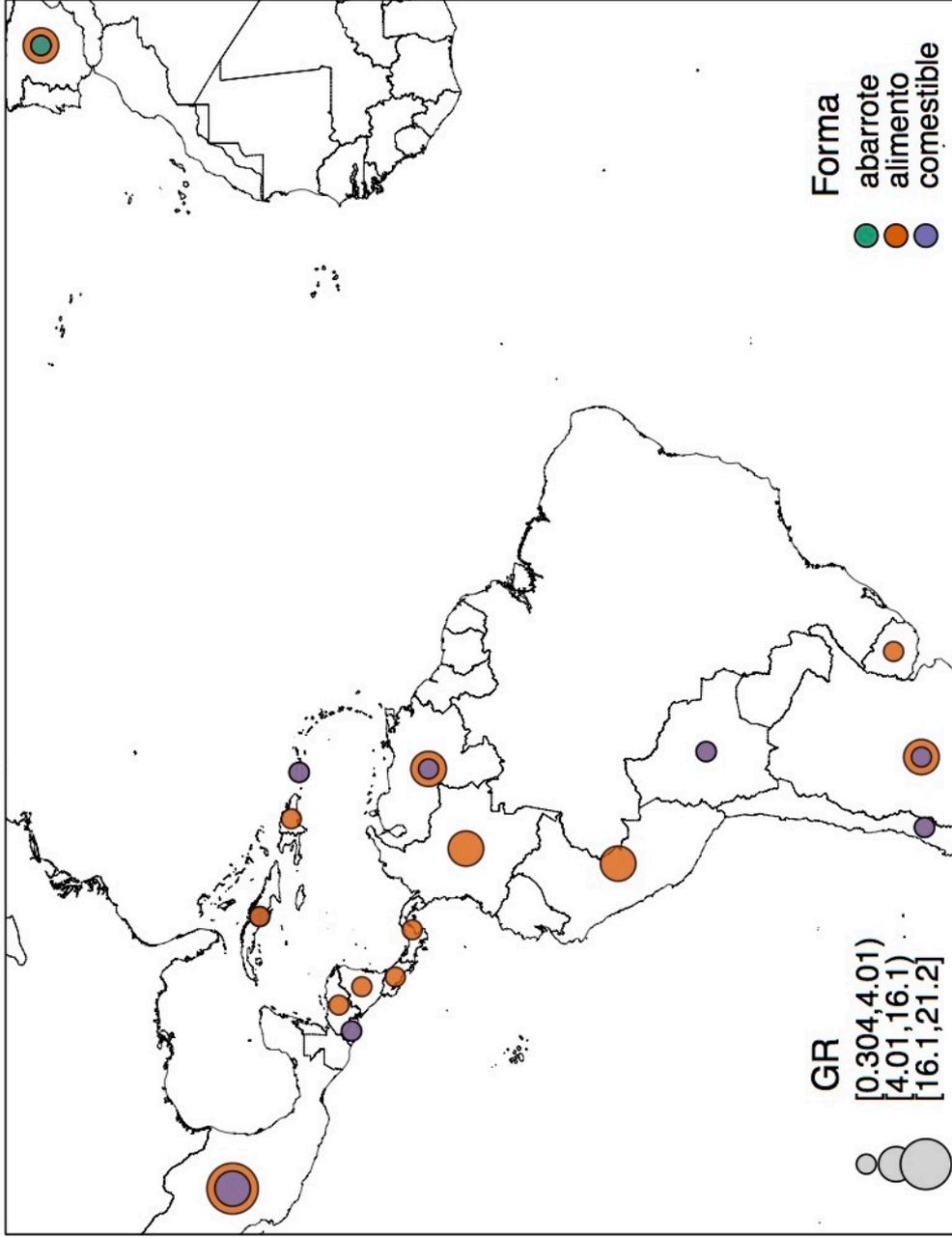
FLAT (US: APPARTMENT)



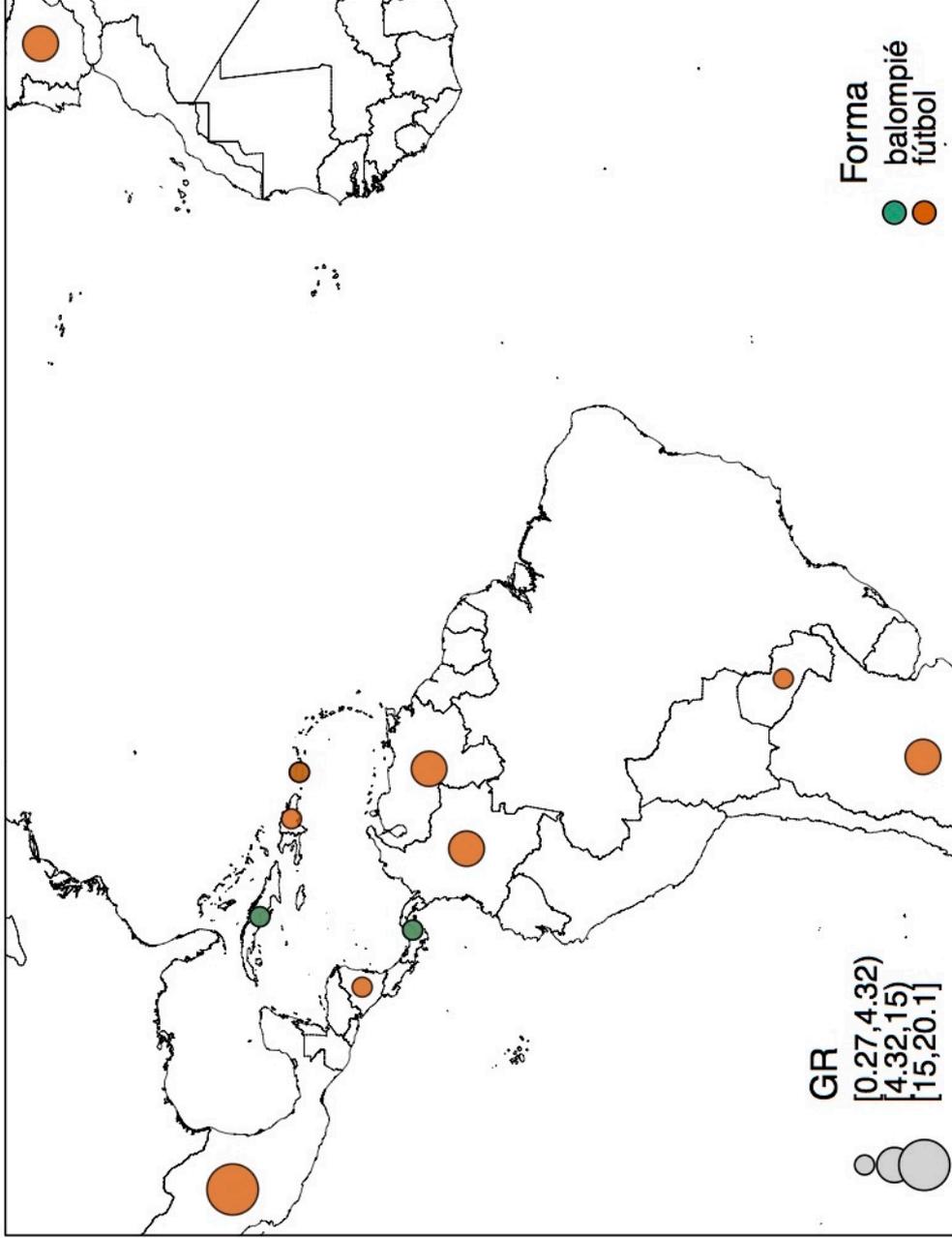
FOLDER



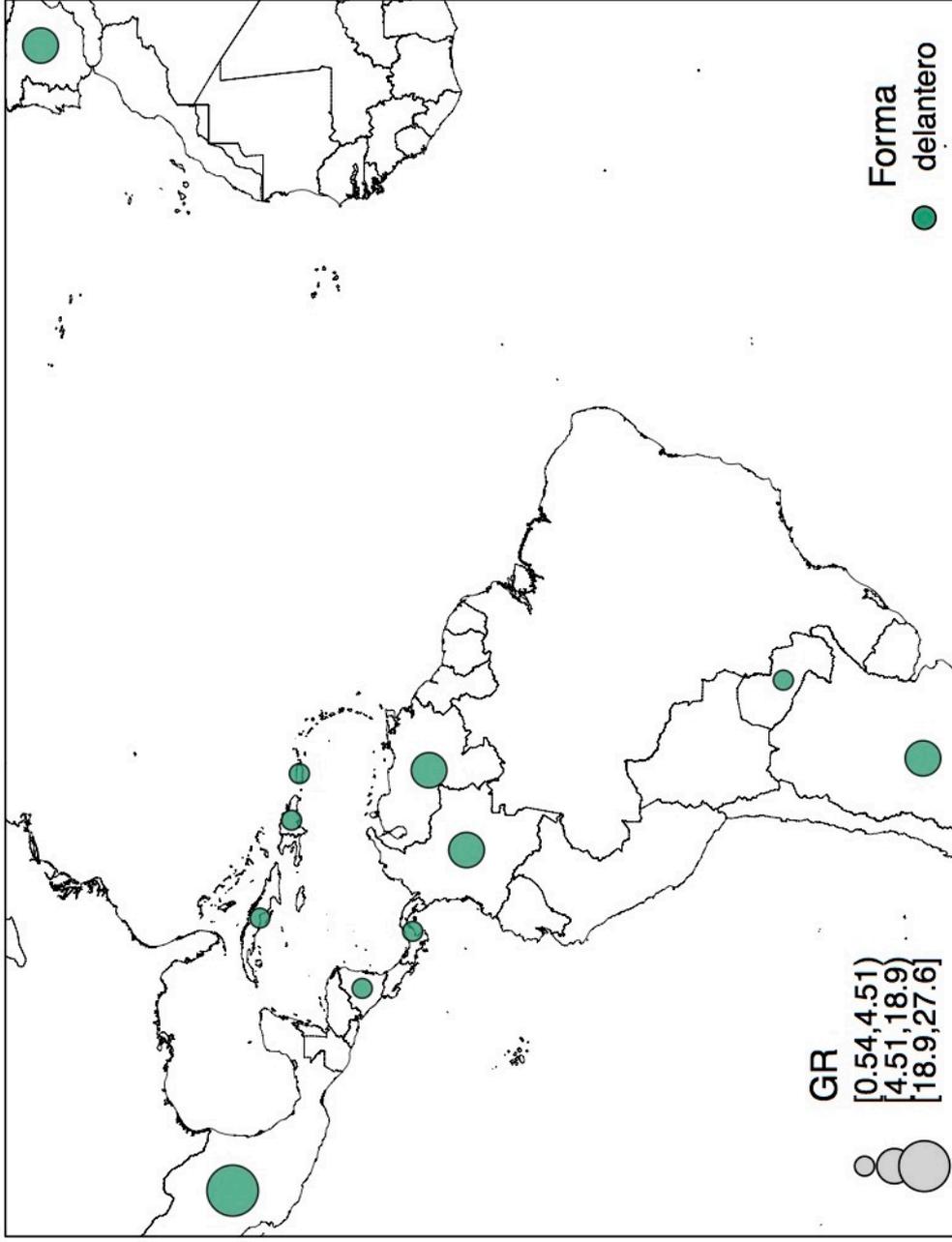
FOOD



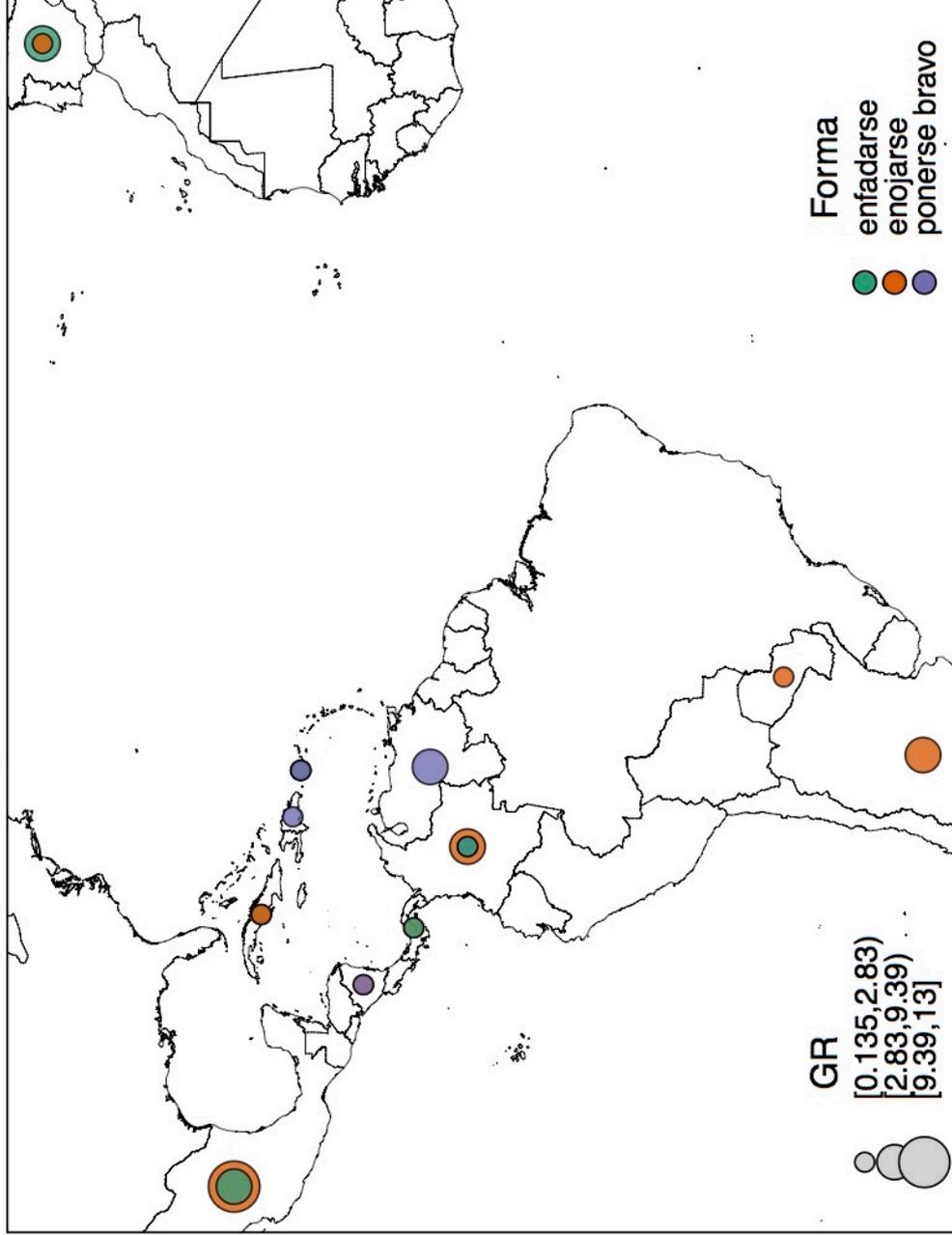
FOOTBALL



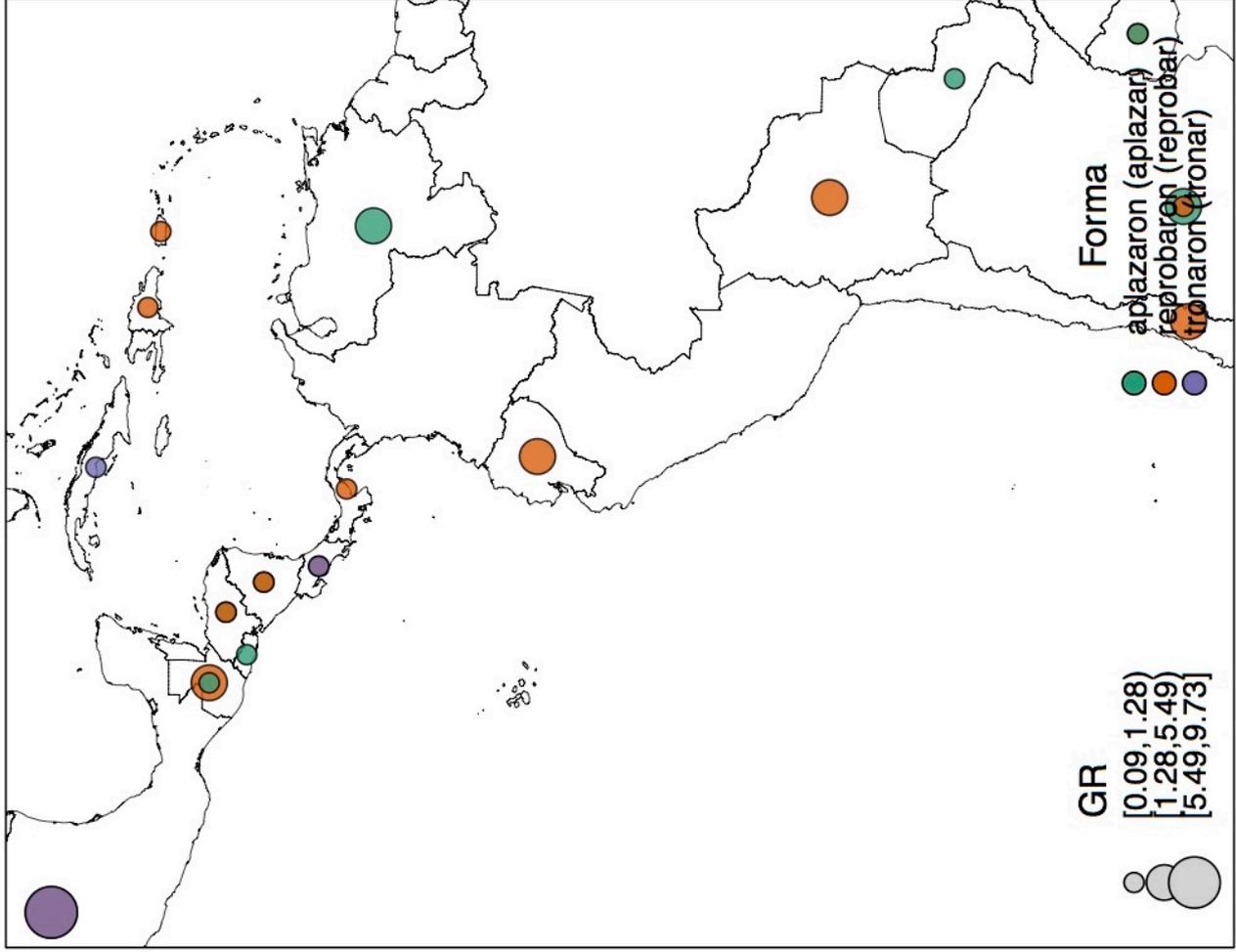
FORWARD



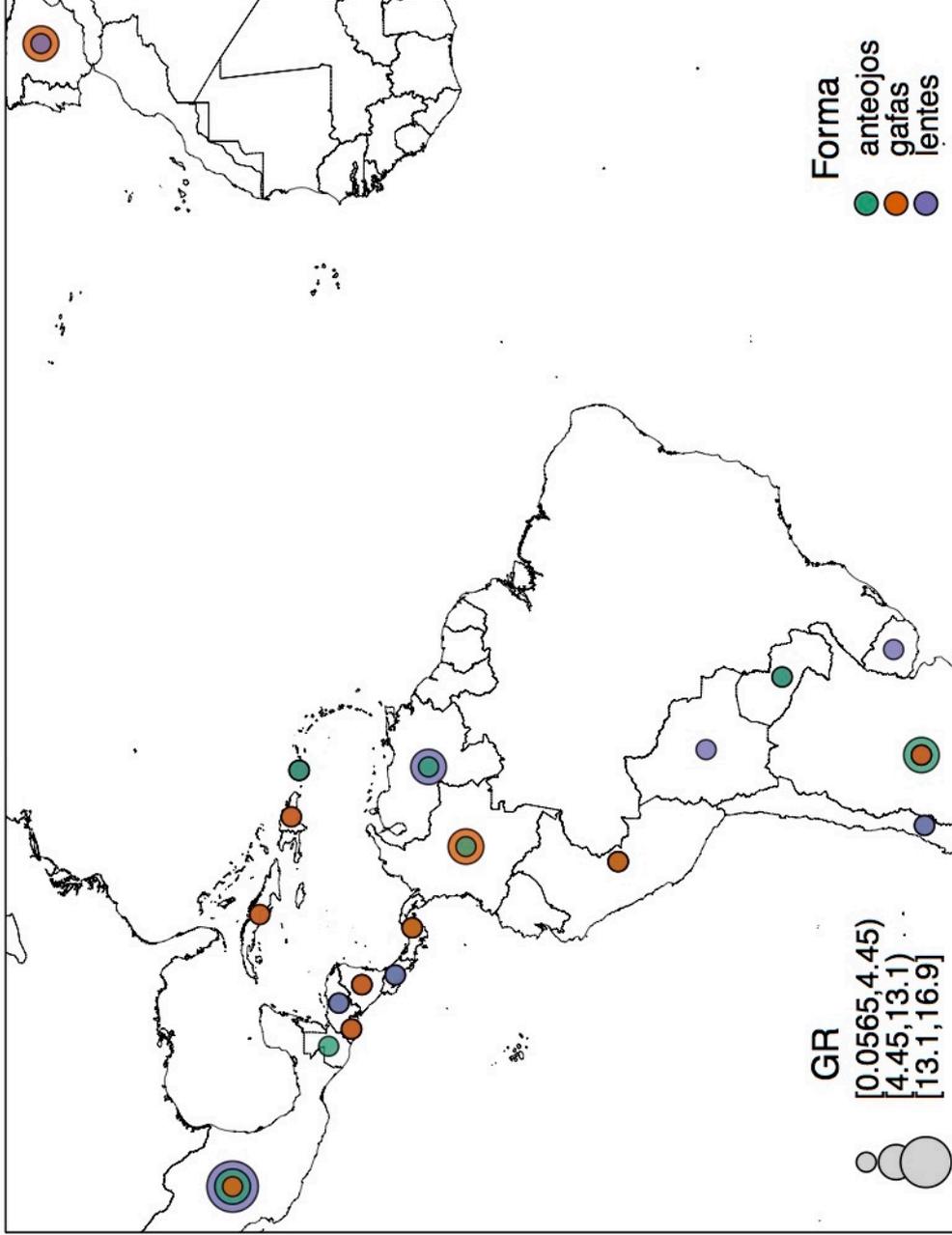
GET ANGRY



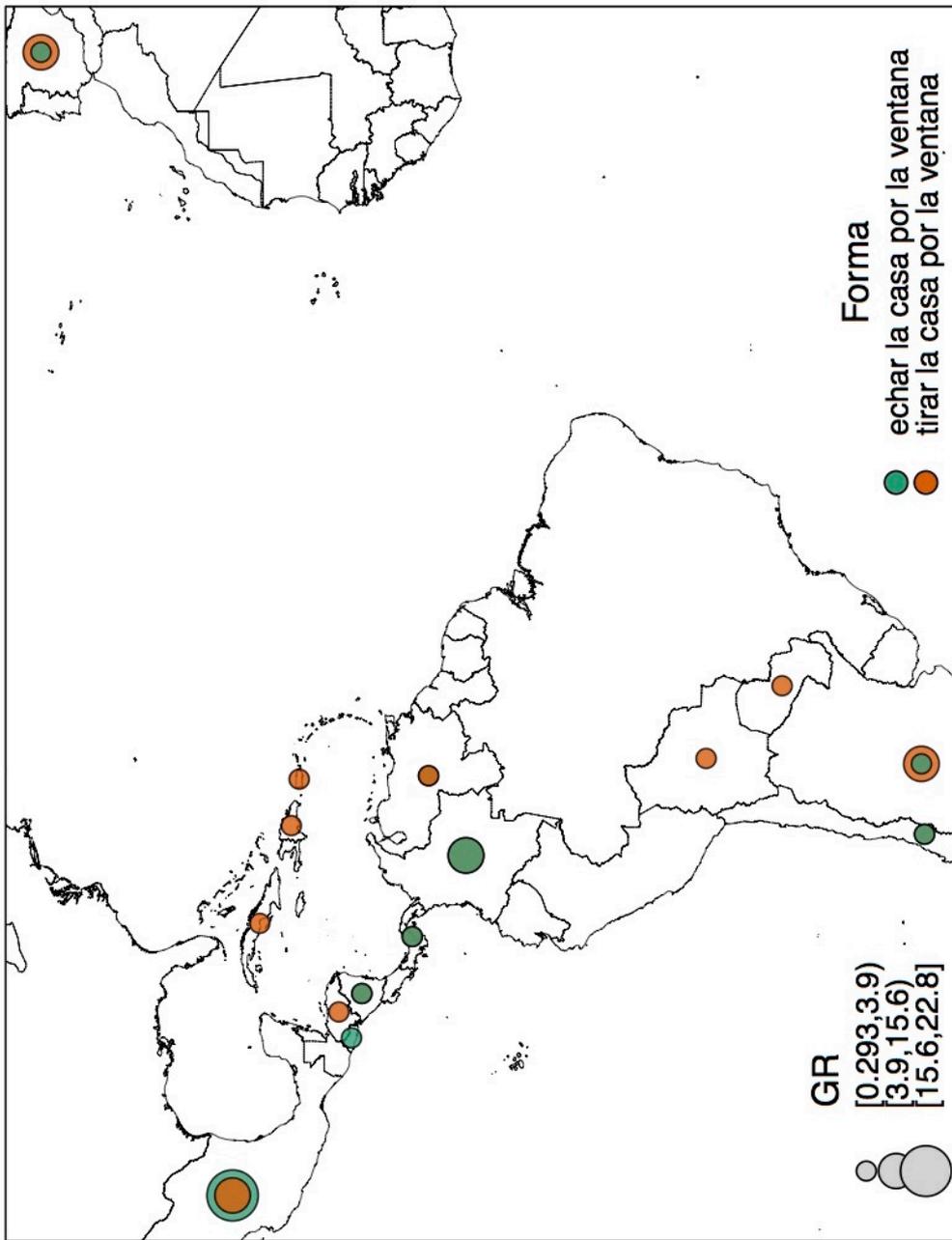
GIVE A
FAILING
GRADE



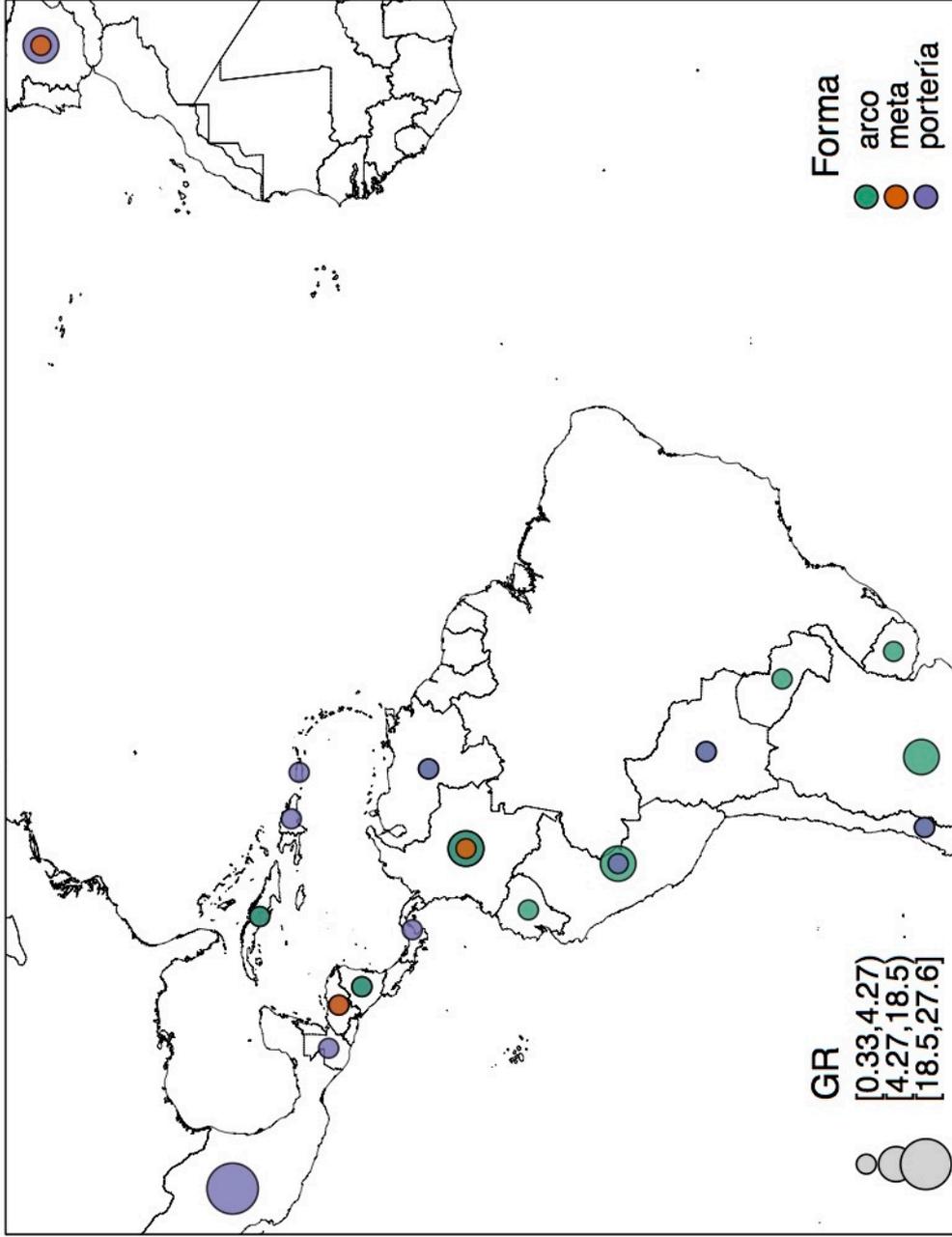
GLASSES



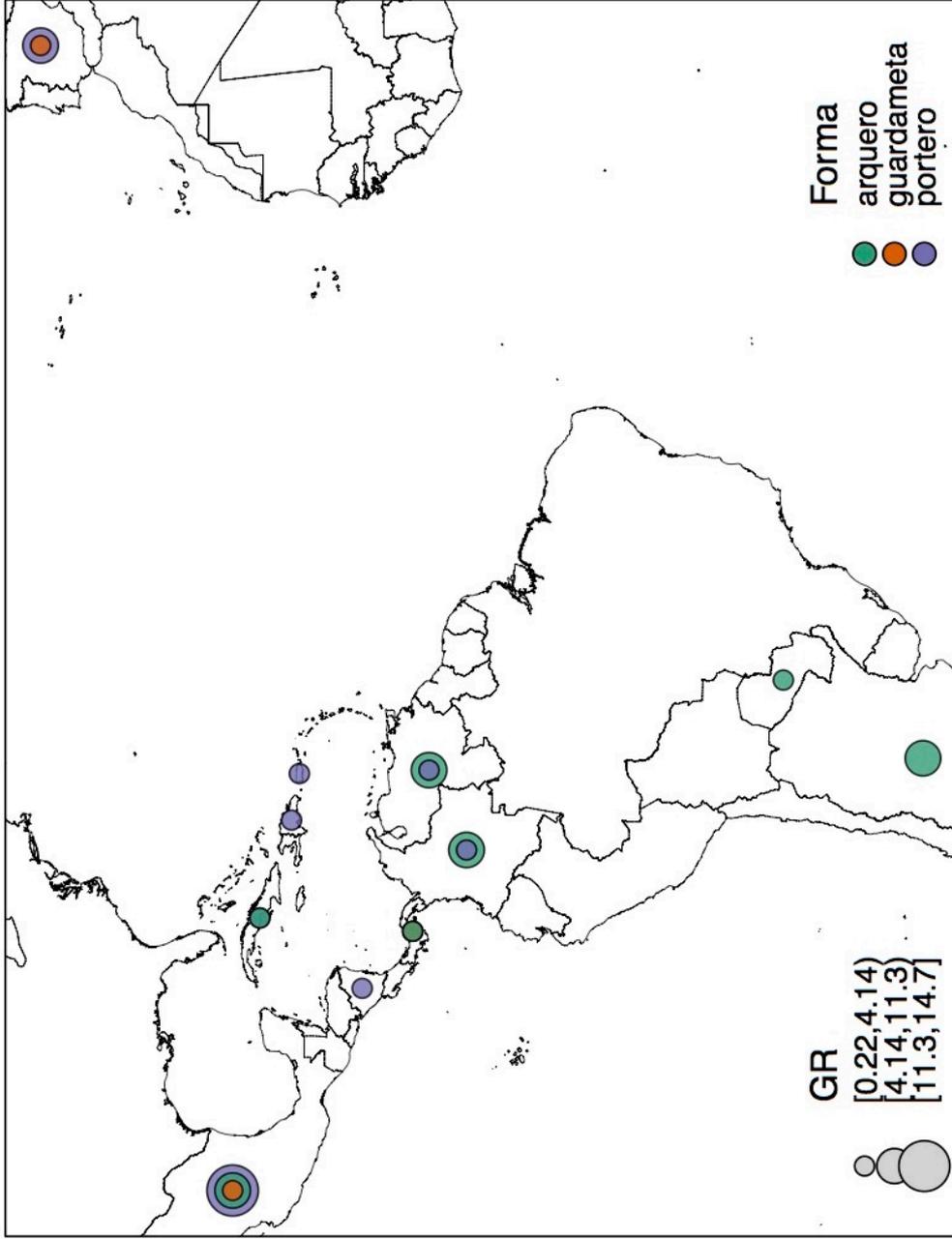
GO TO ENORMOUS EXPENSE



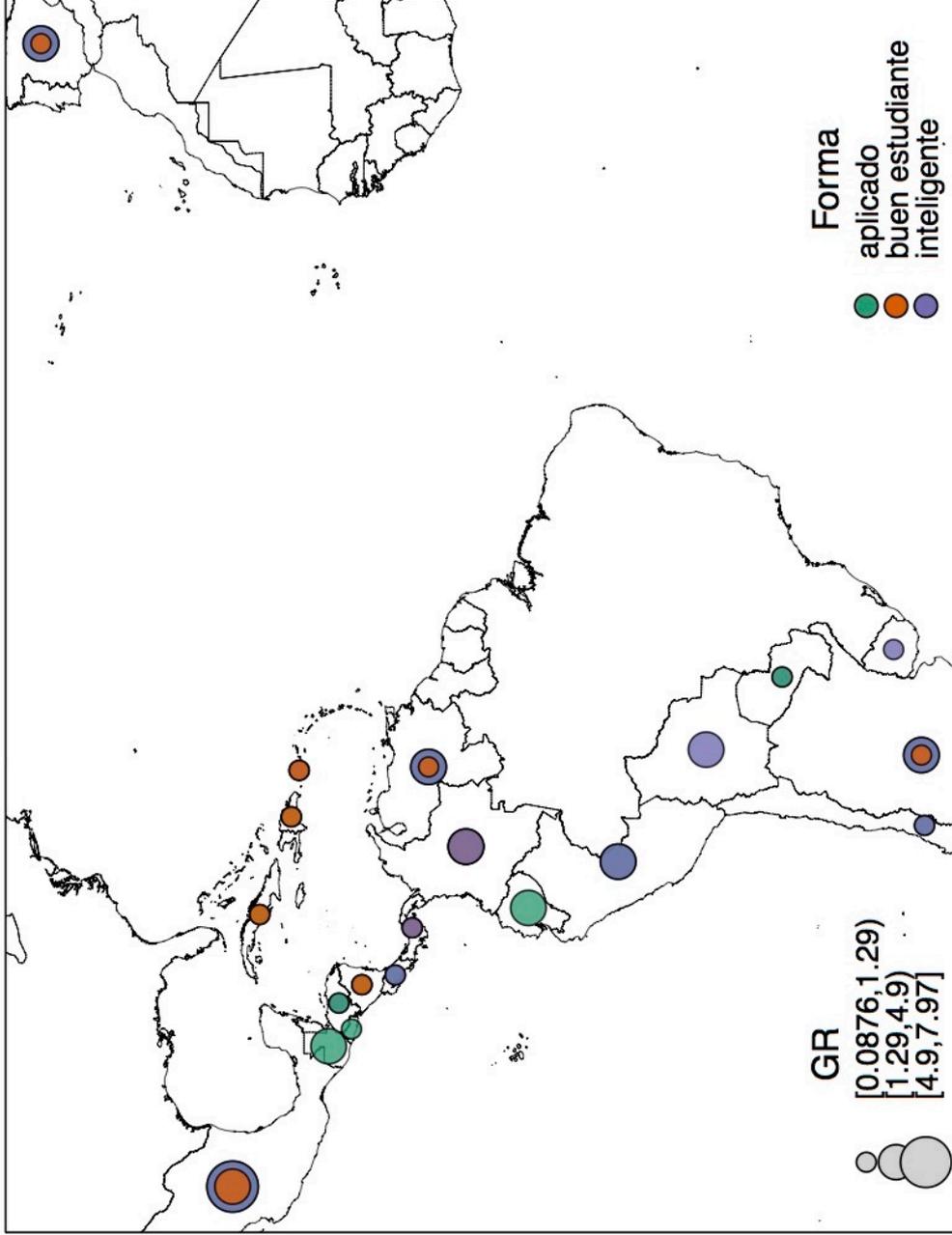
GOAL



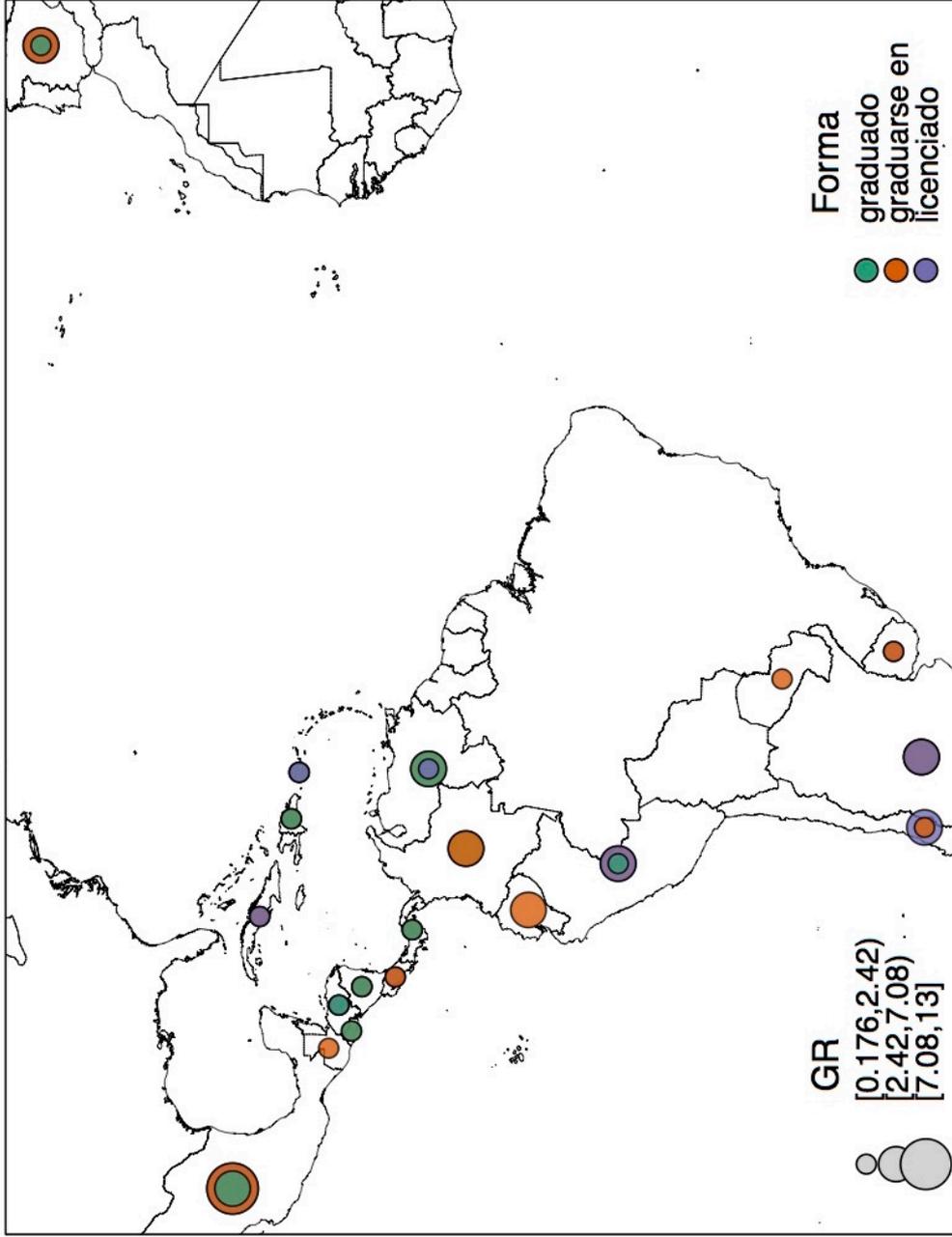
GOALKEEPER



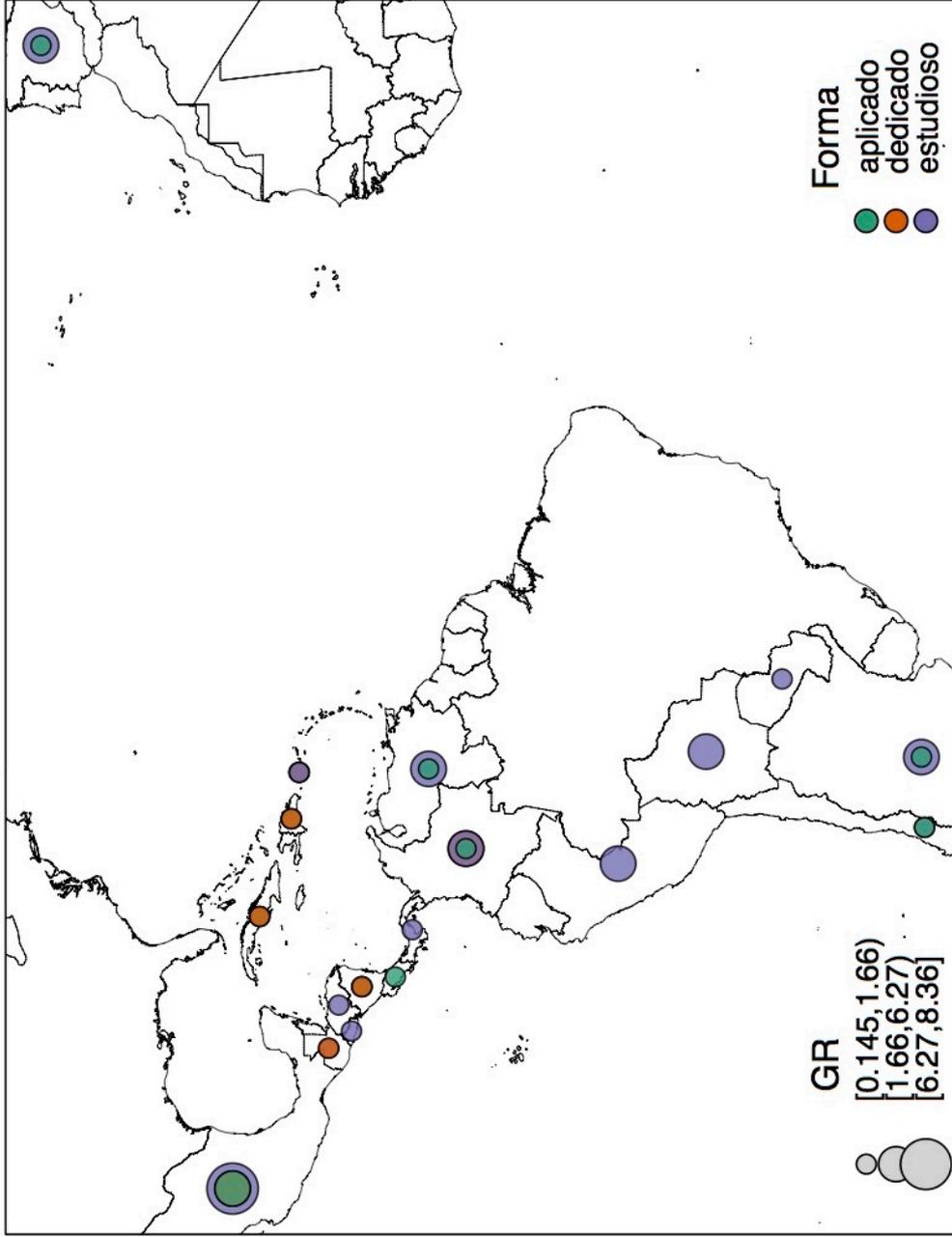
GOOD STUDENT



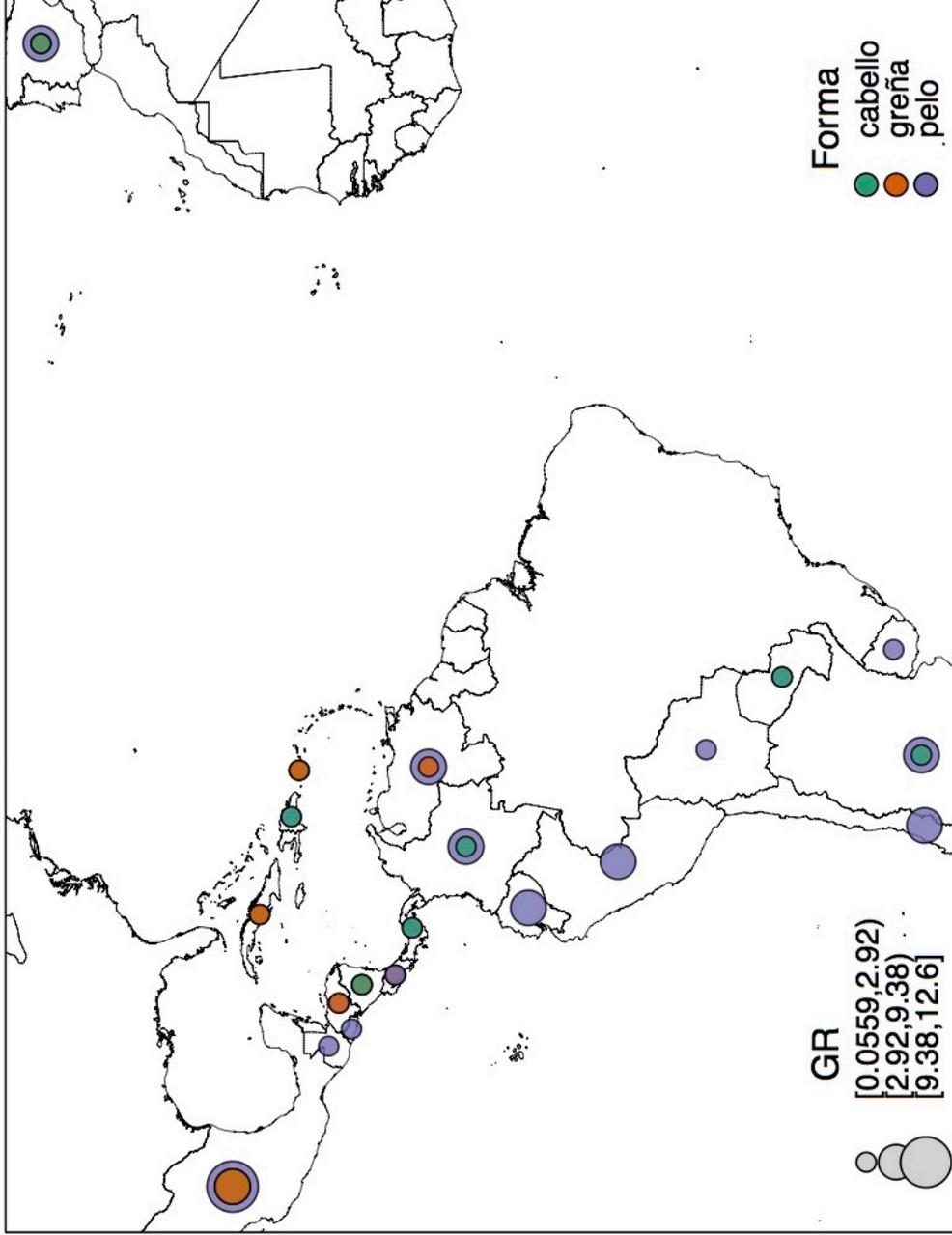
GRADUATE



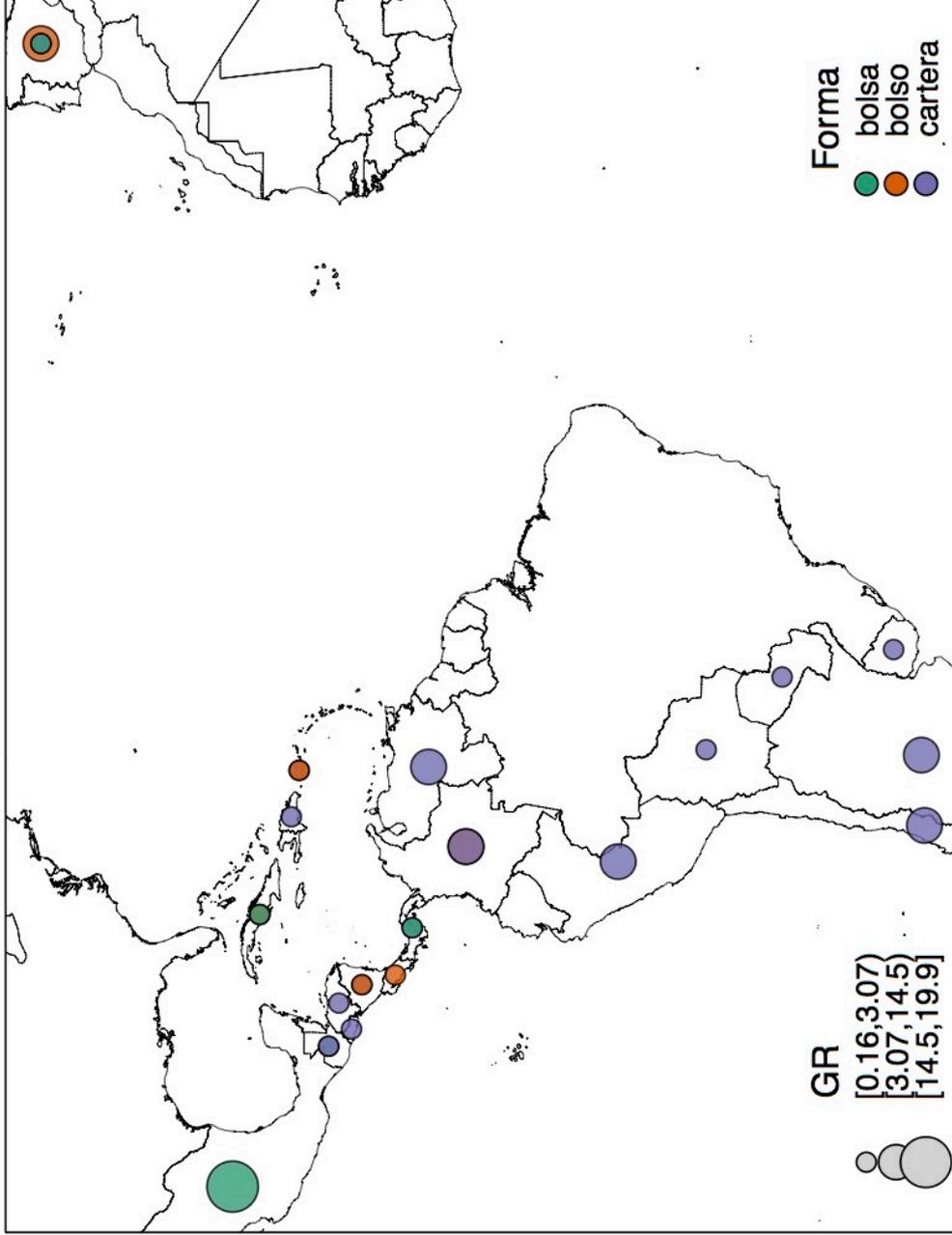
GRIND



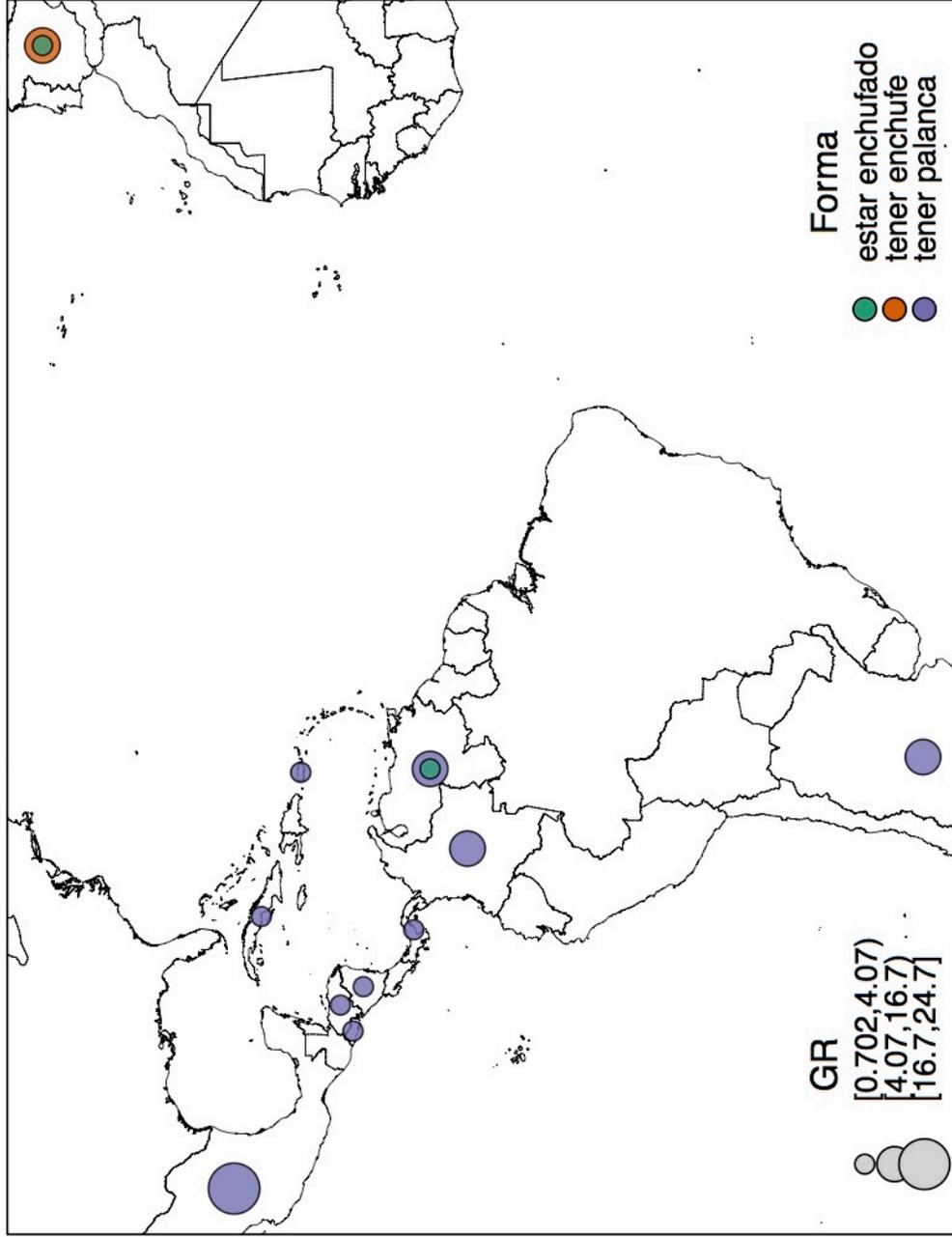
HAIR



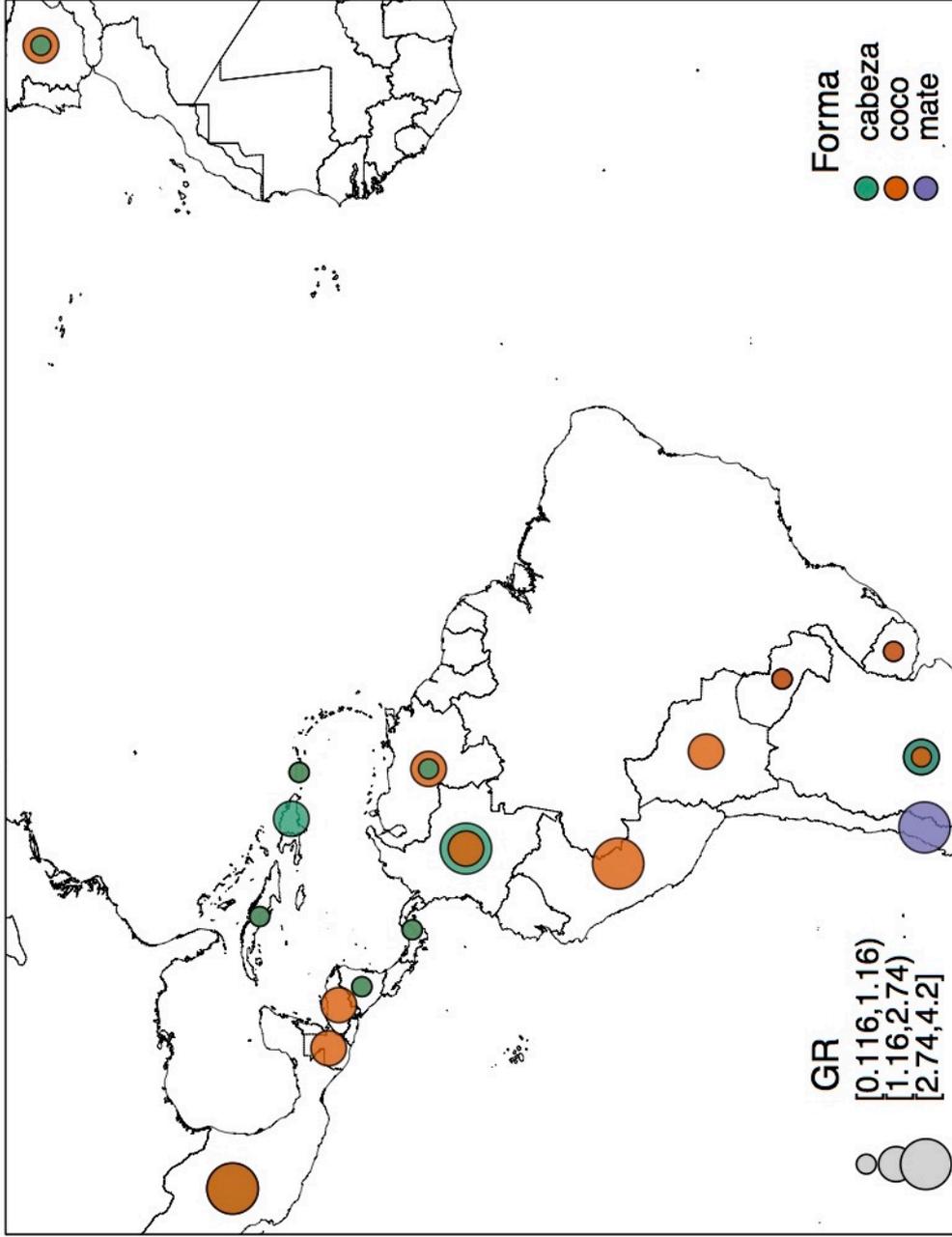
HANDBAG



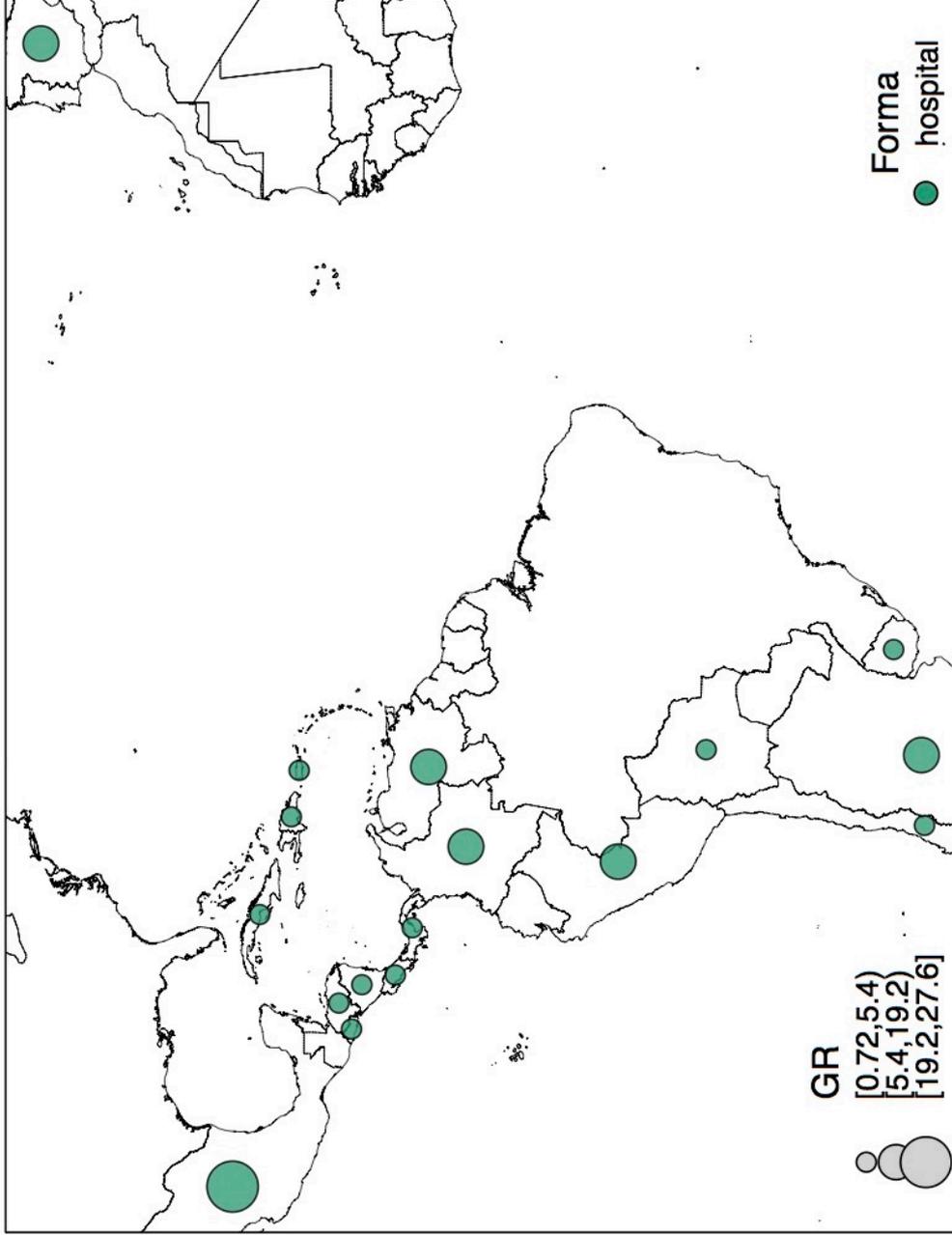
HAVE PULL



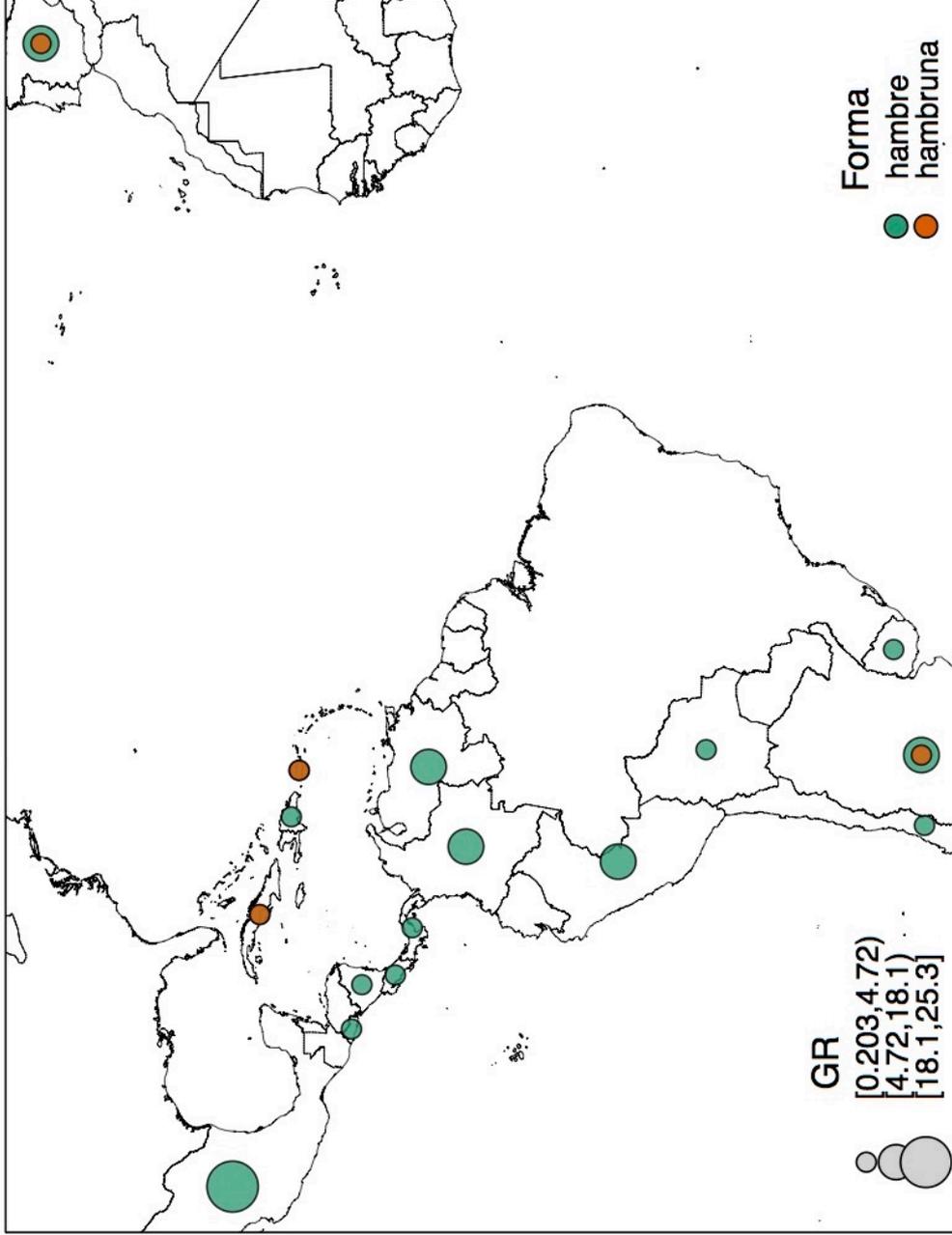
HEAD



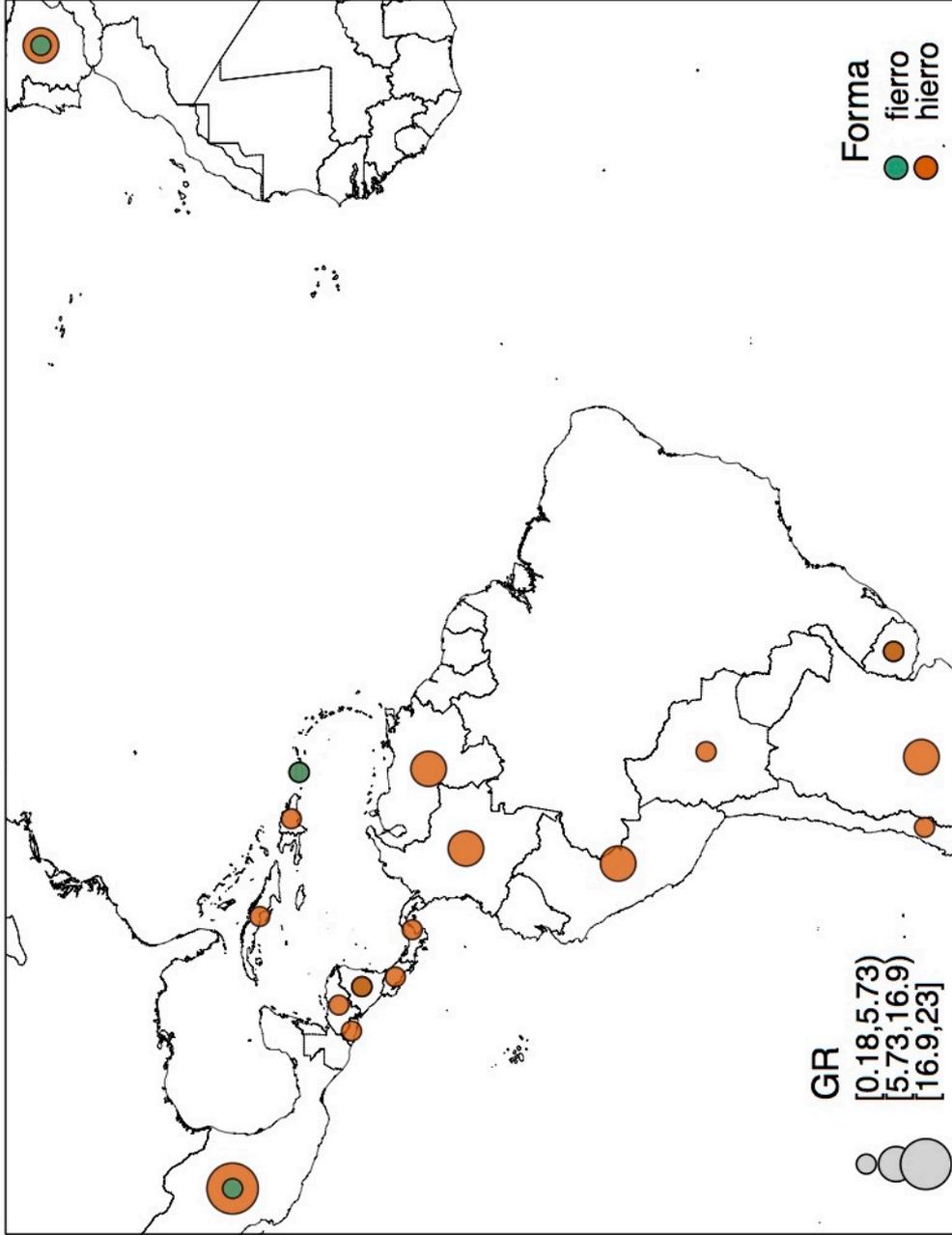
HOSPITAL



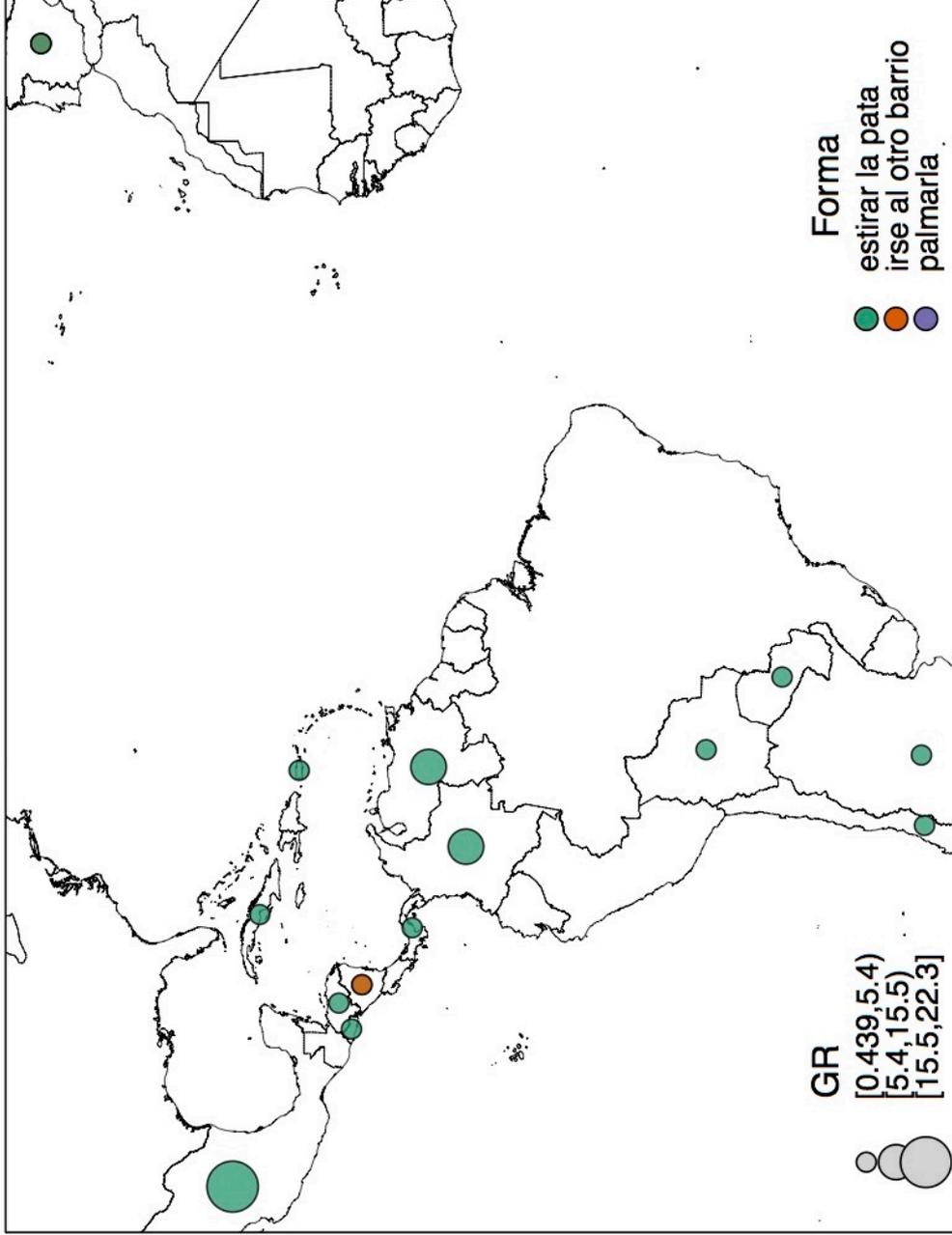
HUNGER



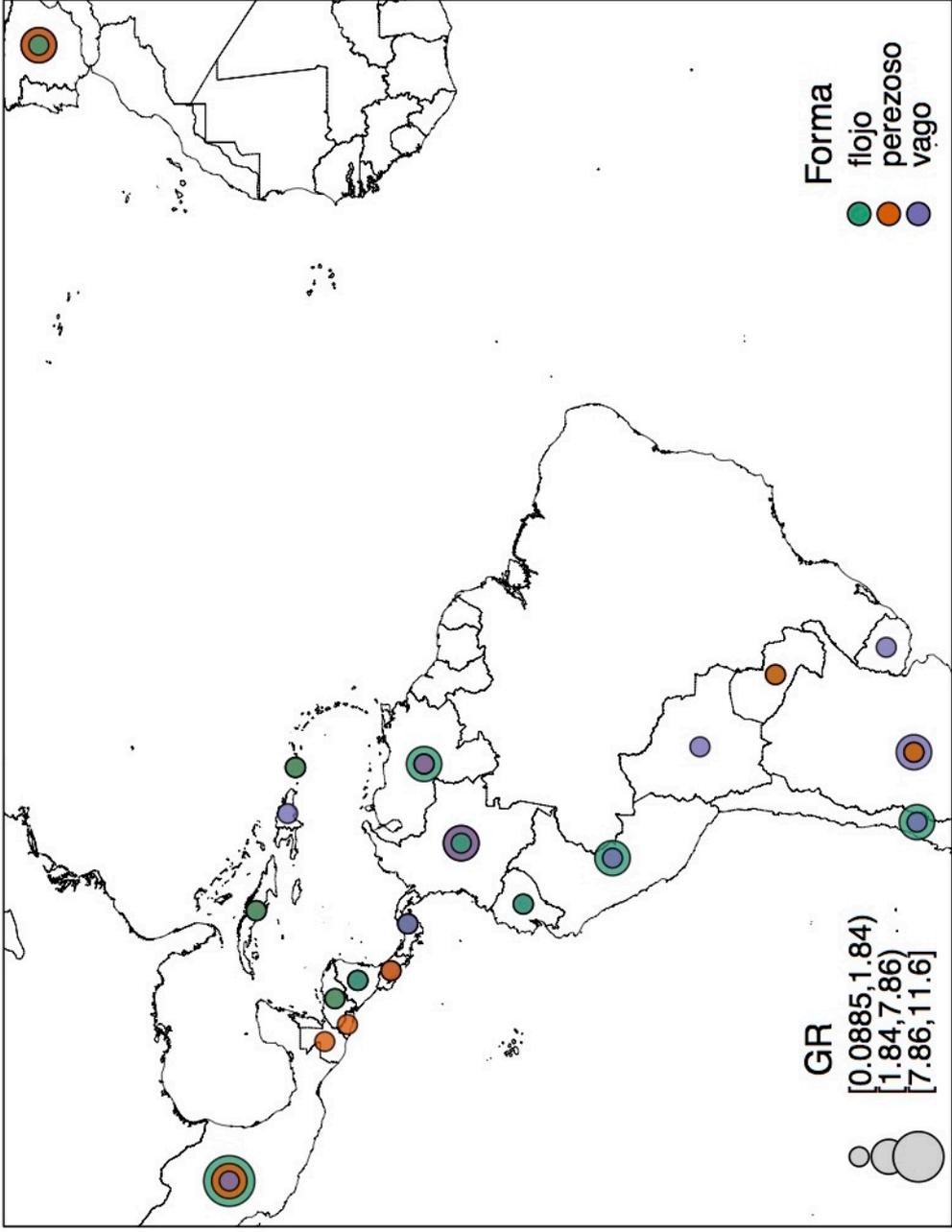
IRON



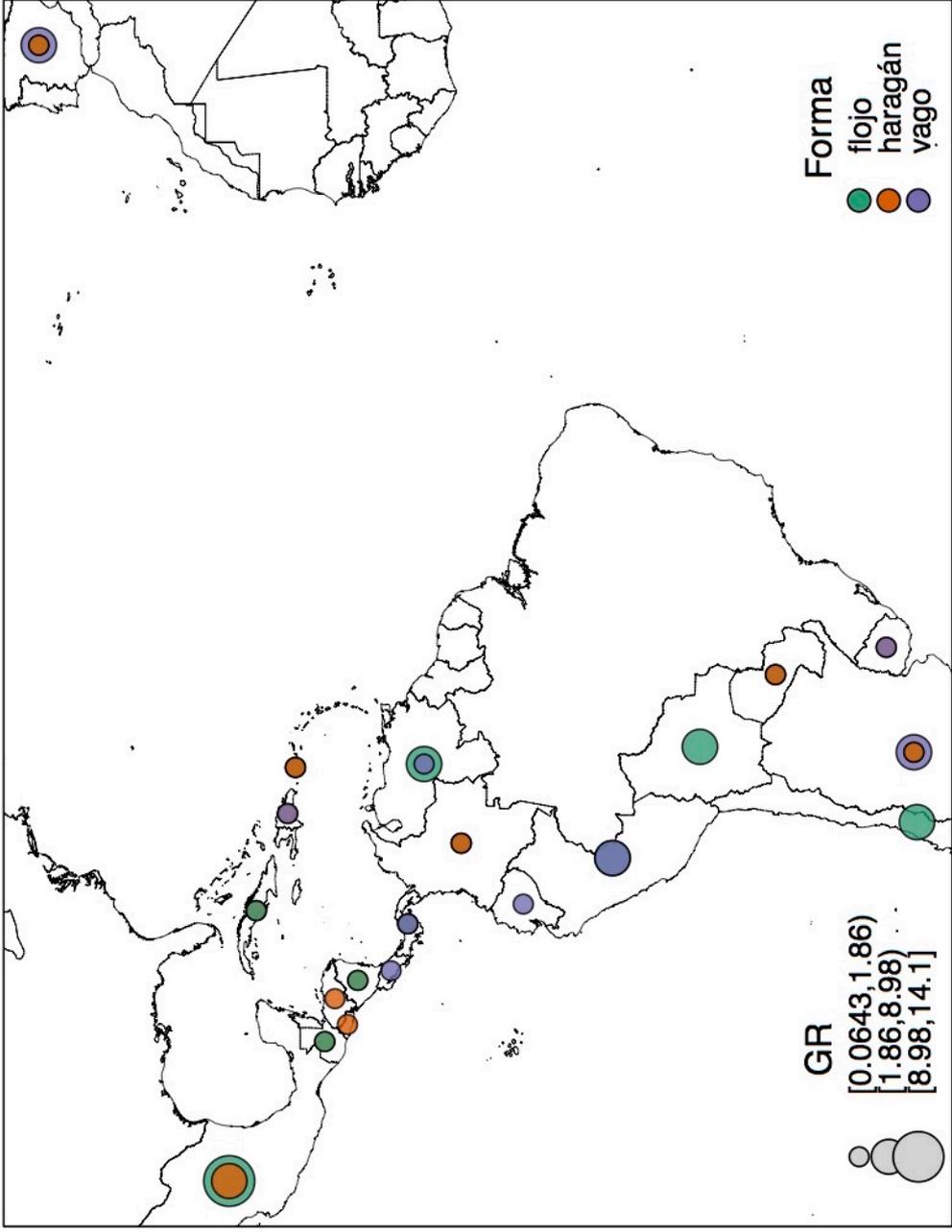
KICK THE BULLET



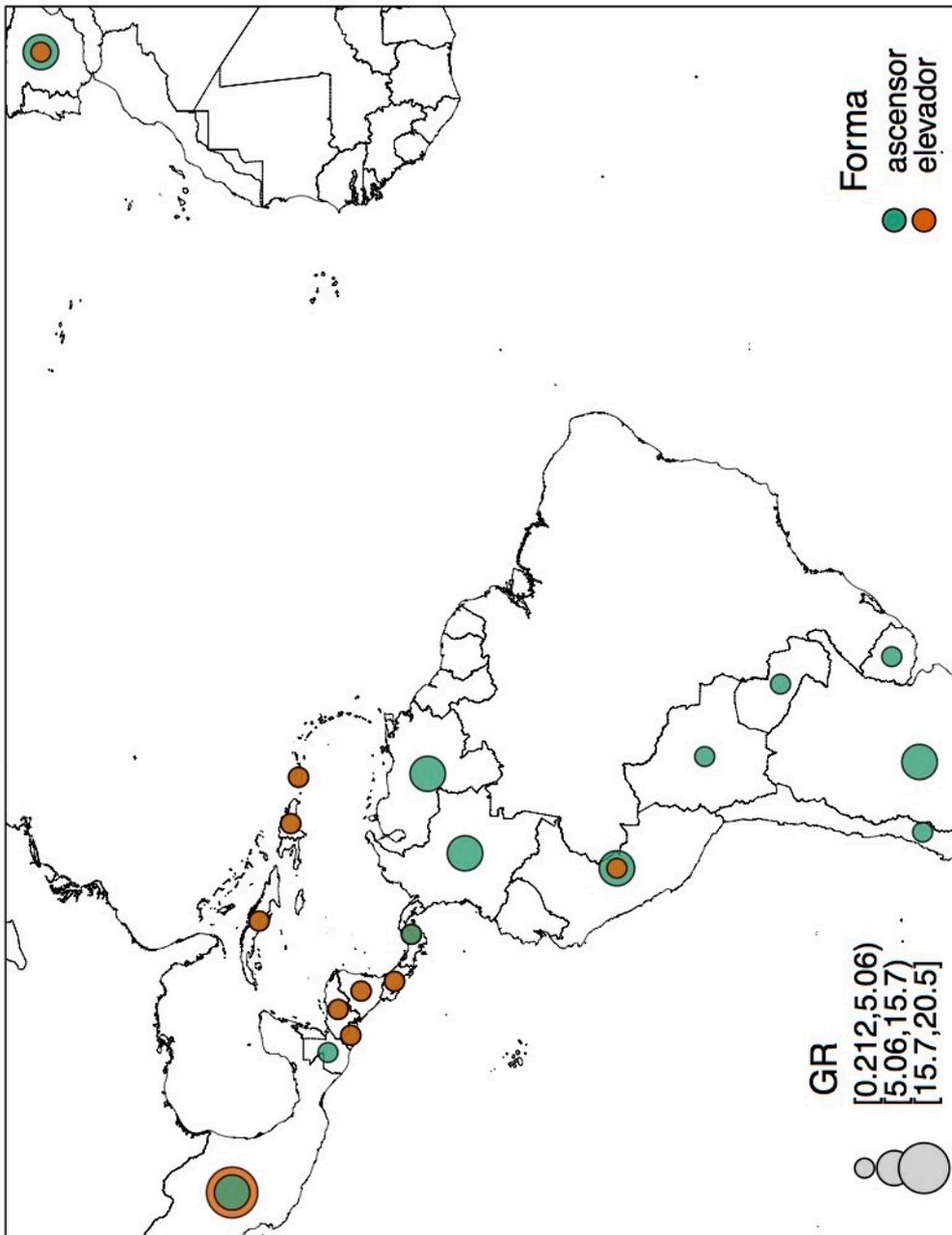
LAZY



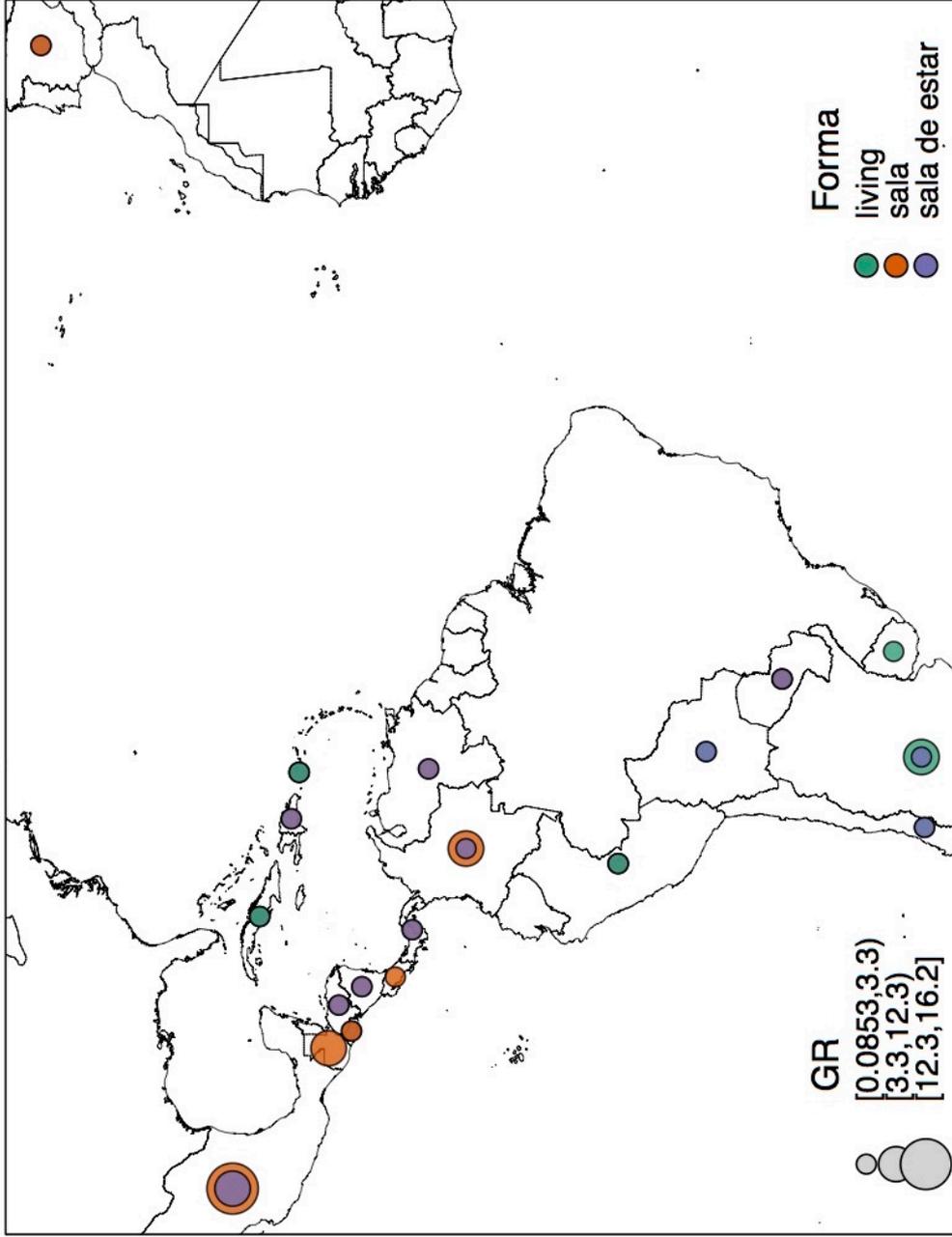
LAZY STUDENT



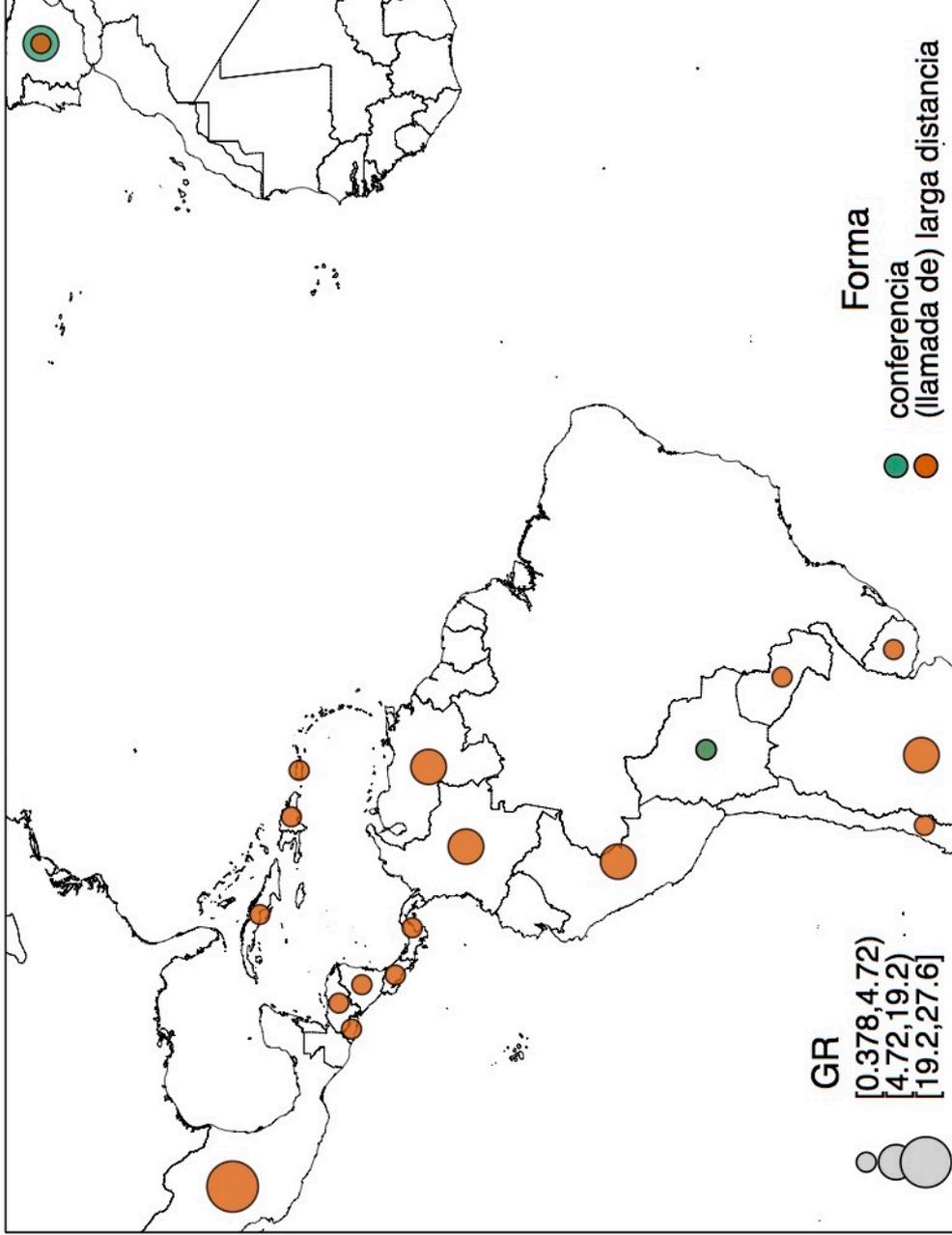
LIFT (US: ELEVATOR)



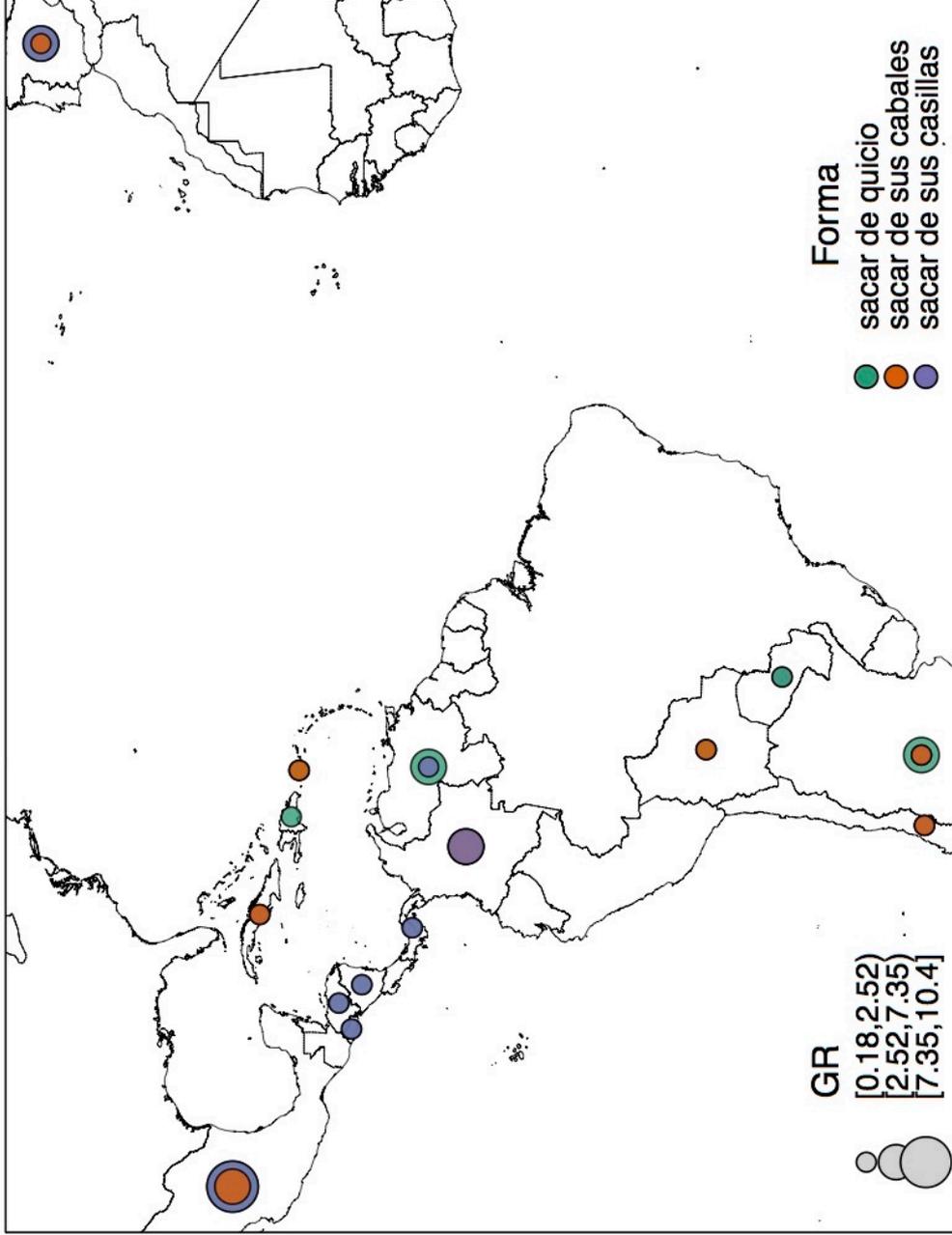
LIVING ROOM



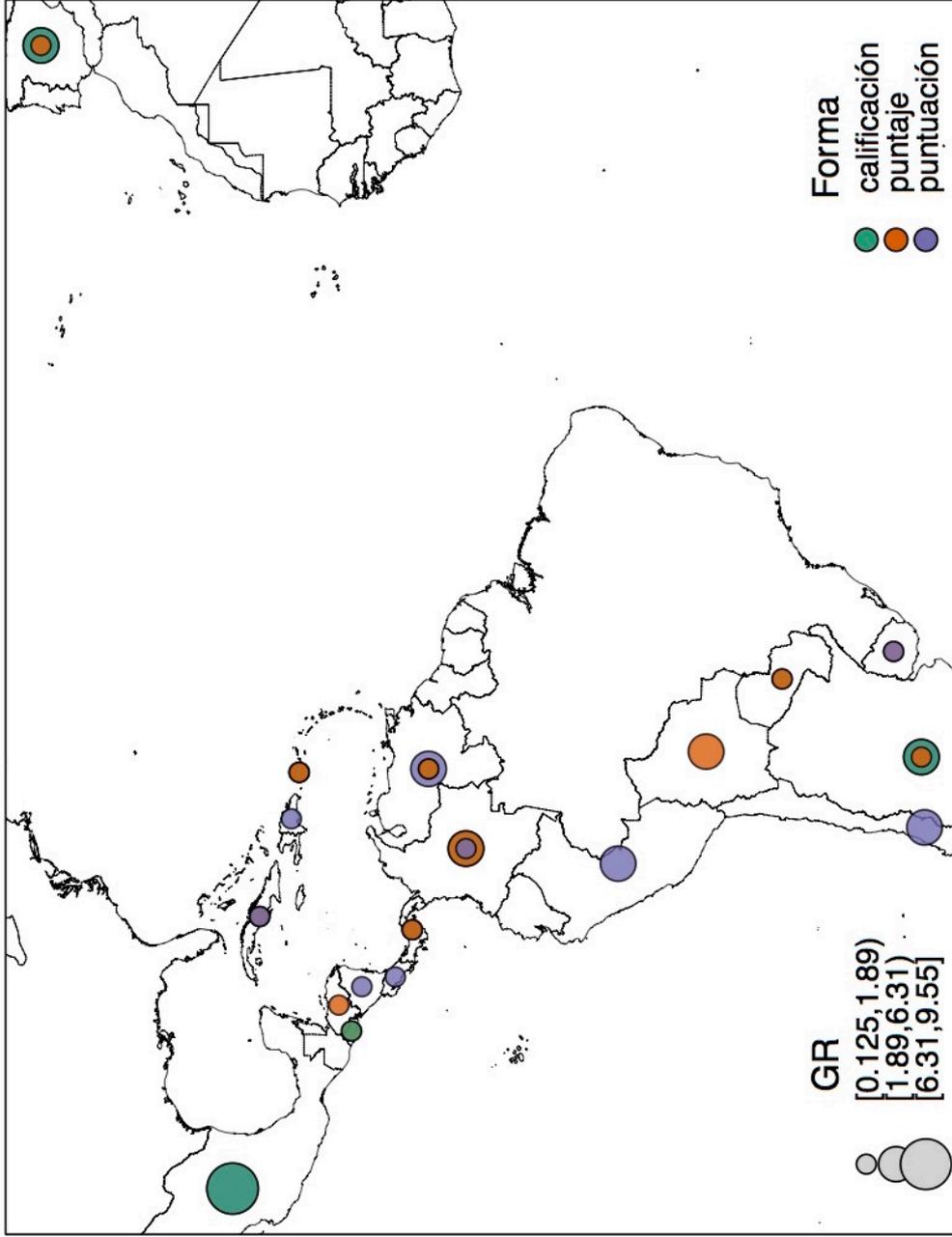
LONG DISTANCE CALL



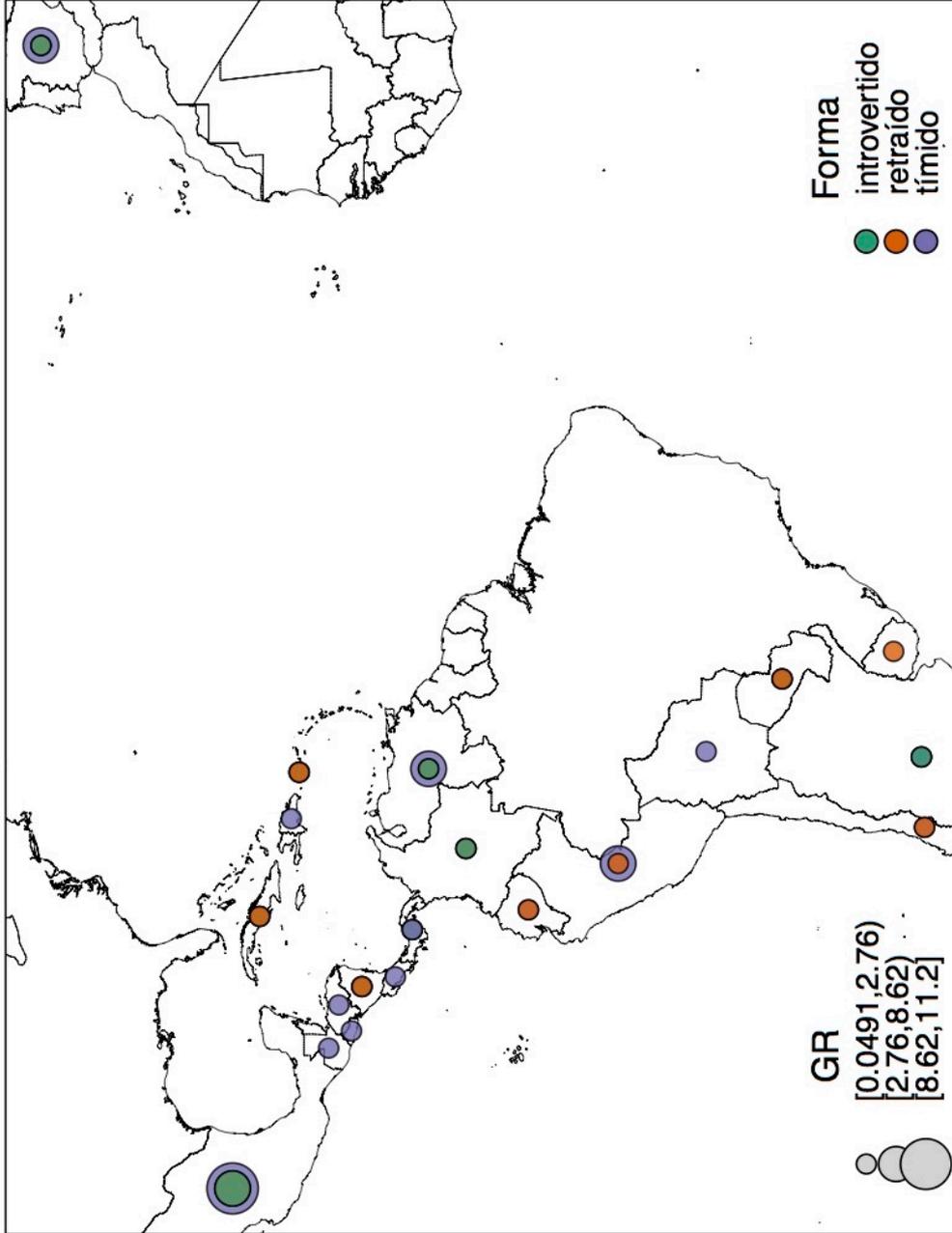
MAKE FLY OFF THE HANDLE



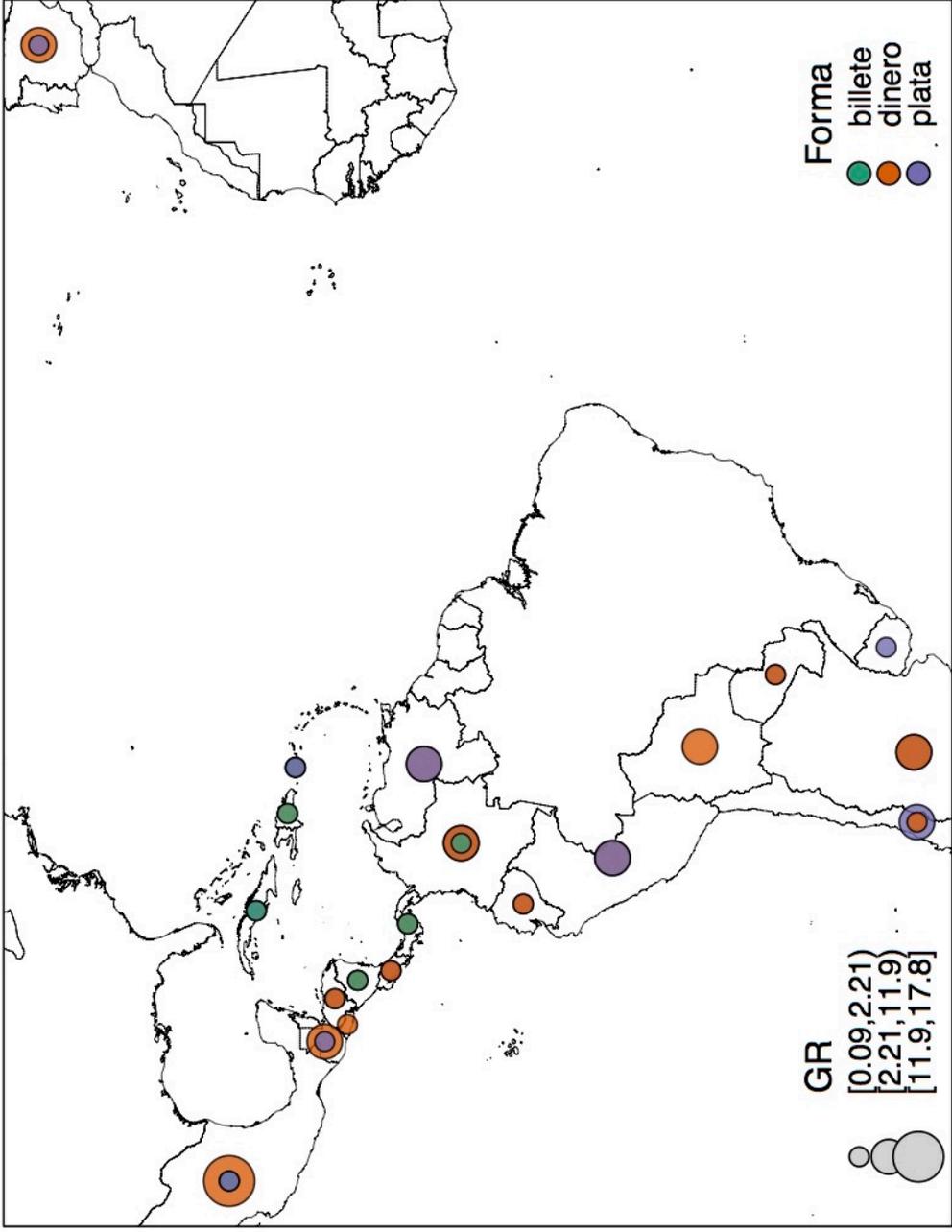
MARK (US: GRADE)



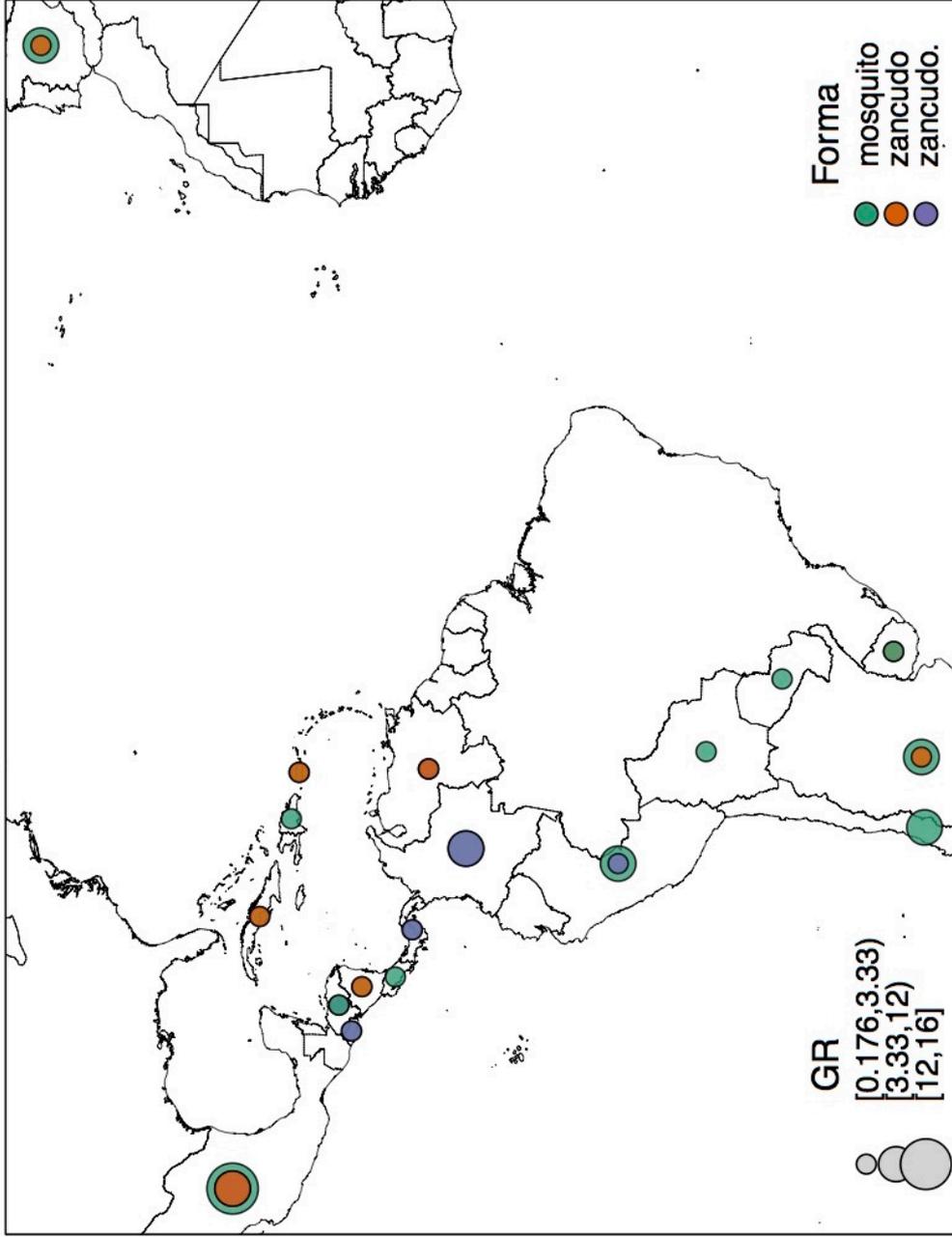
MEEK



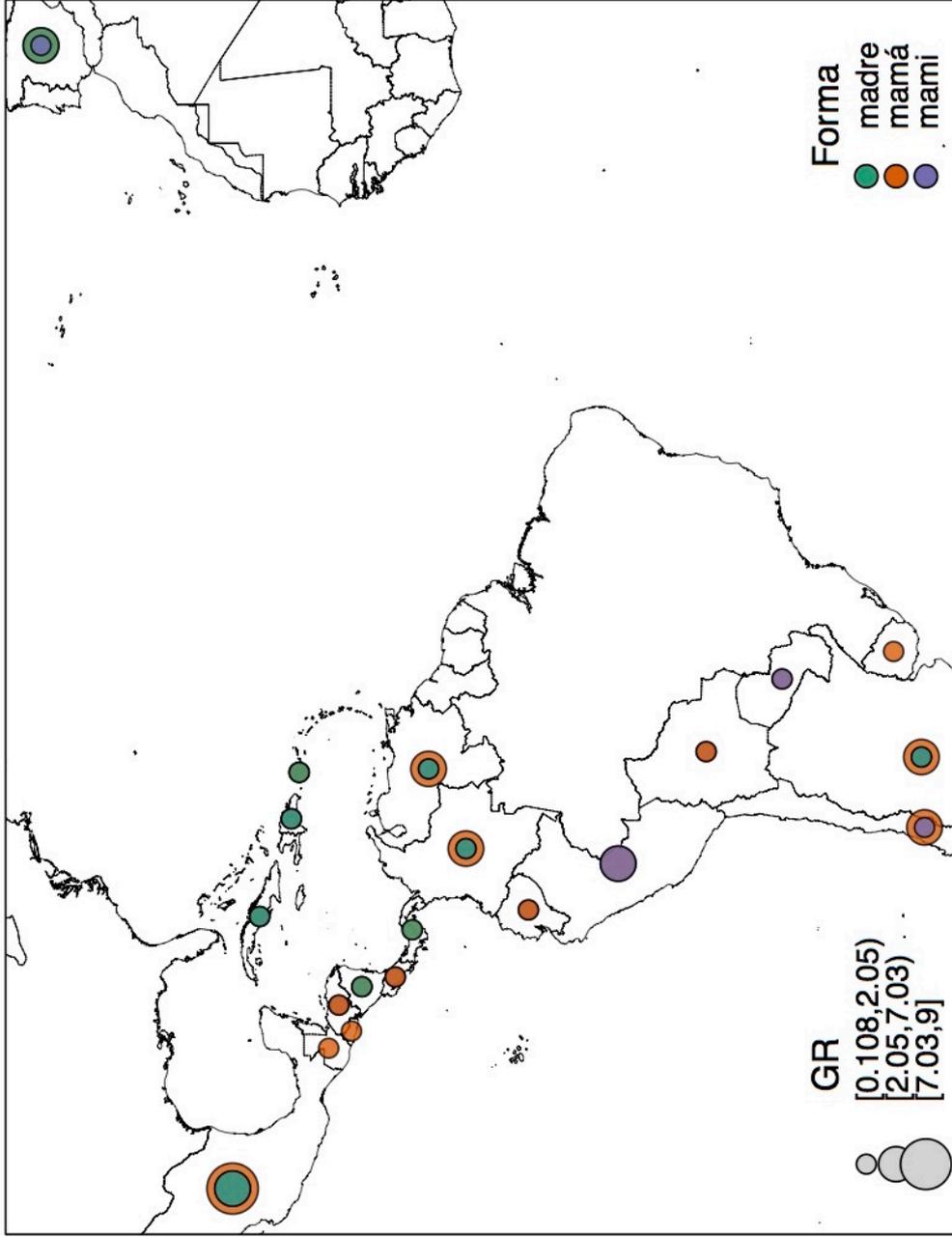
MONEY



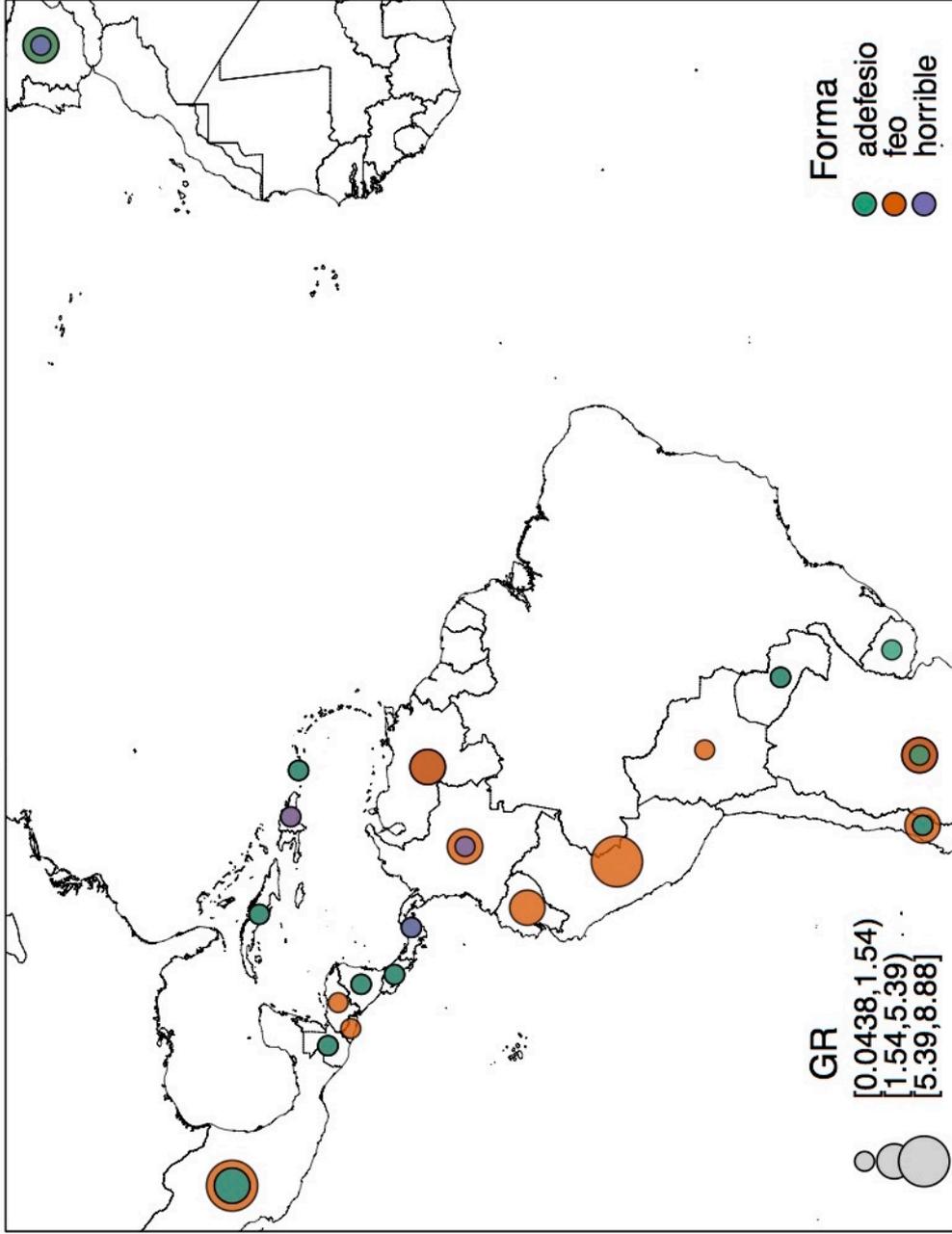
MOSQUITO



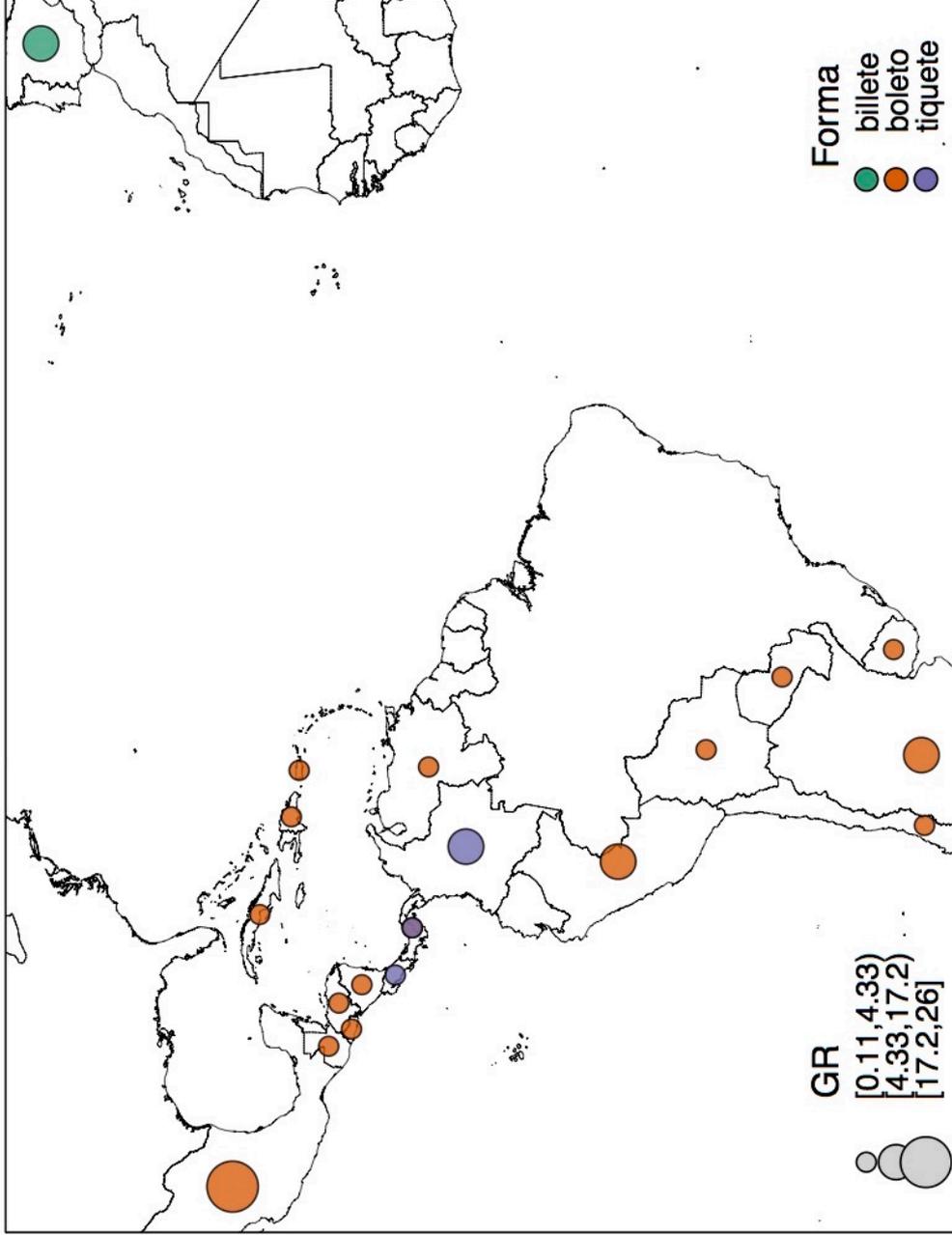
MOTHER (MUMMY)



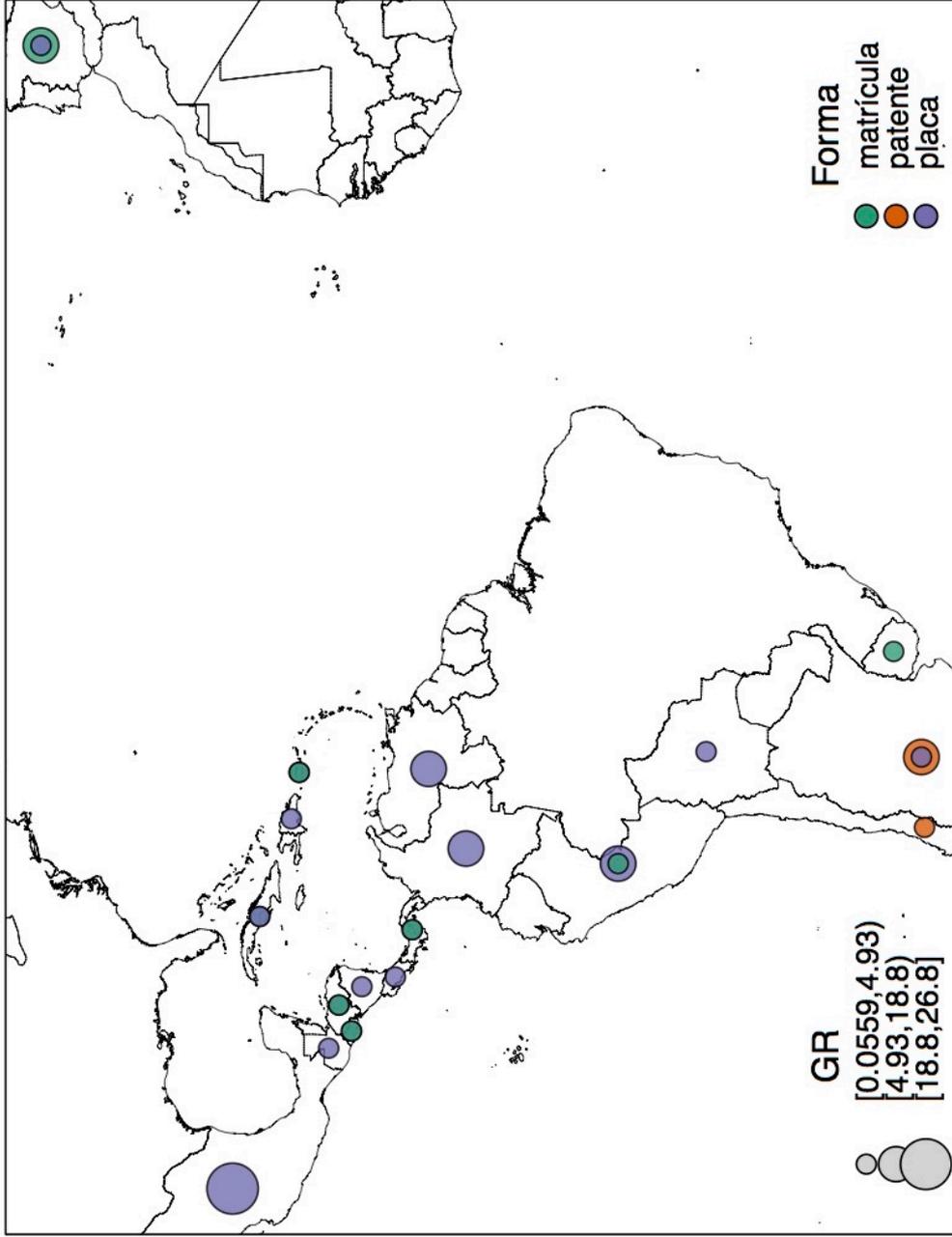
NOT ATTRACTIVE



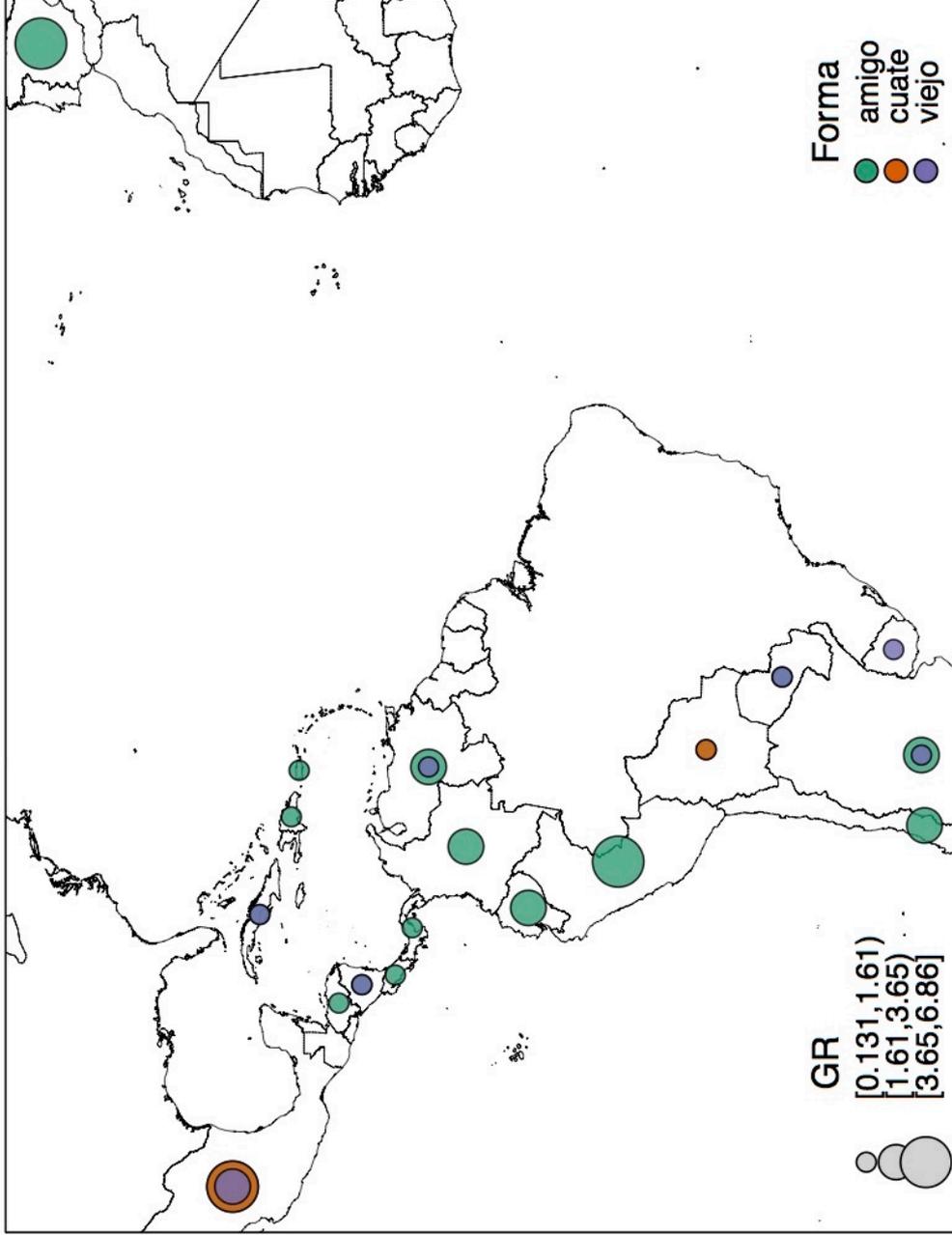
NOTE (US: BILL)



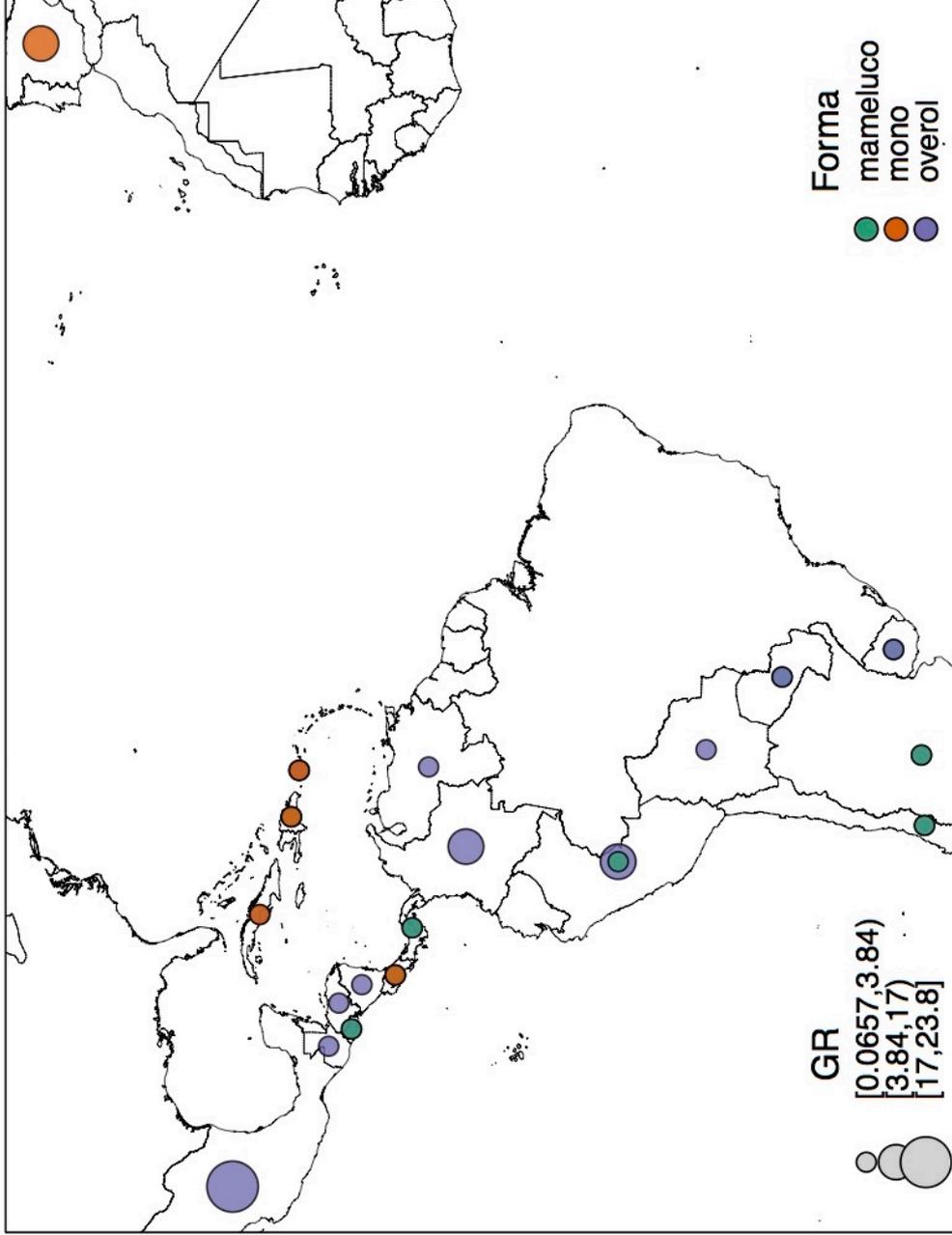
NUMBER PLATE (US: LICENSE PLATE)



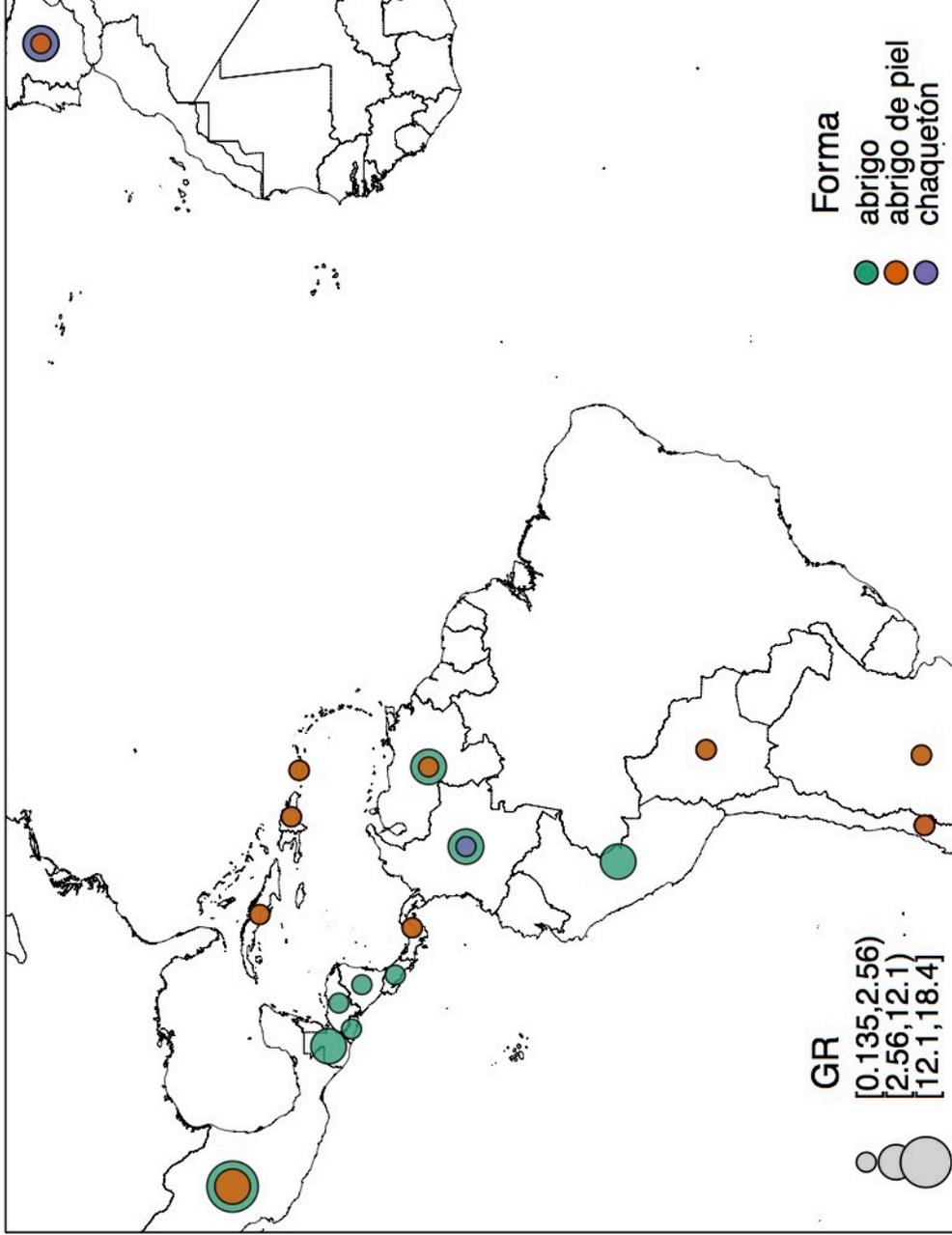
OLD BOY



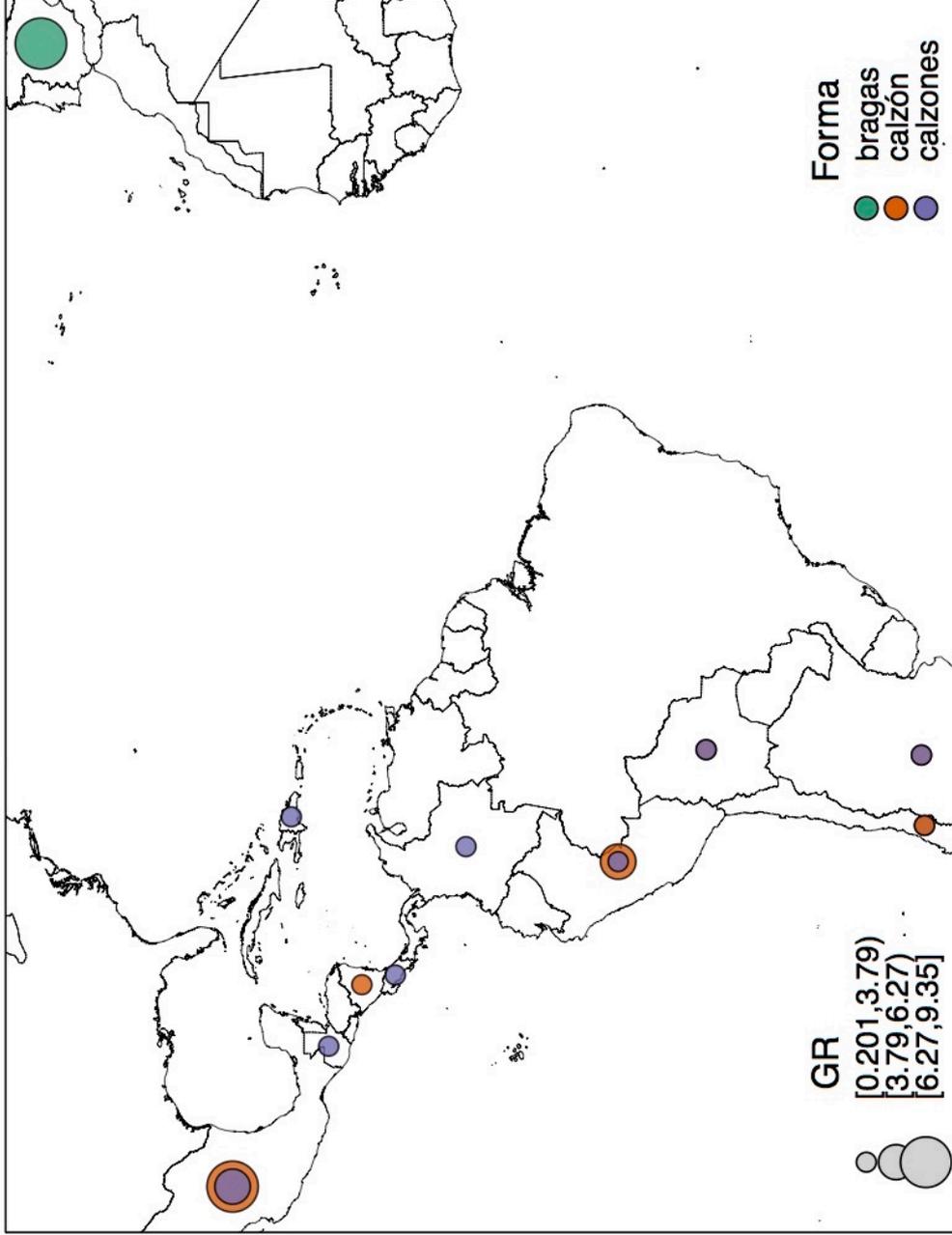
OVERALLS



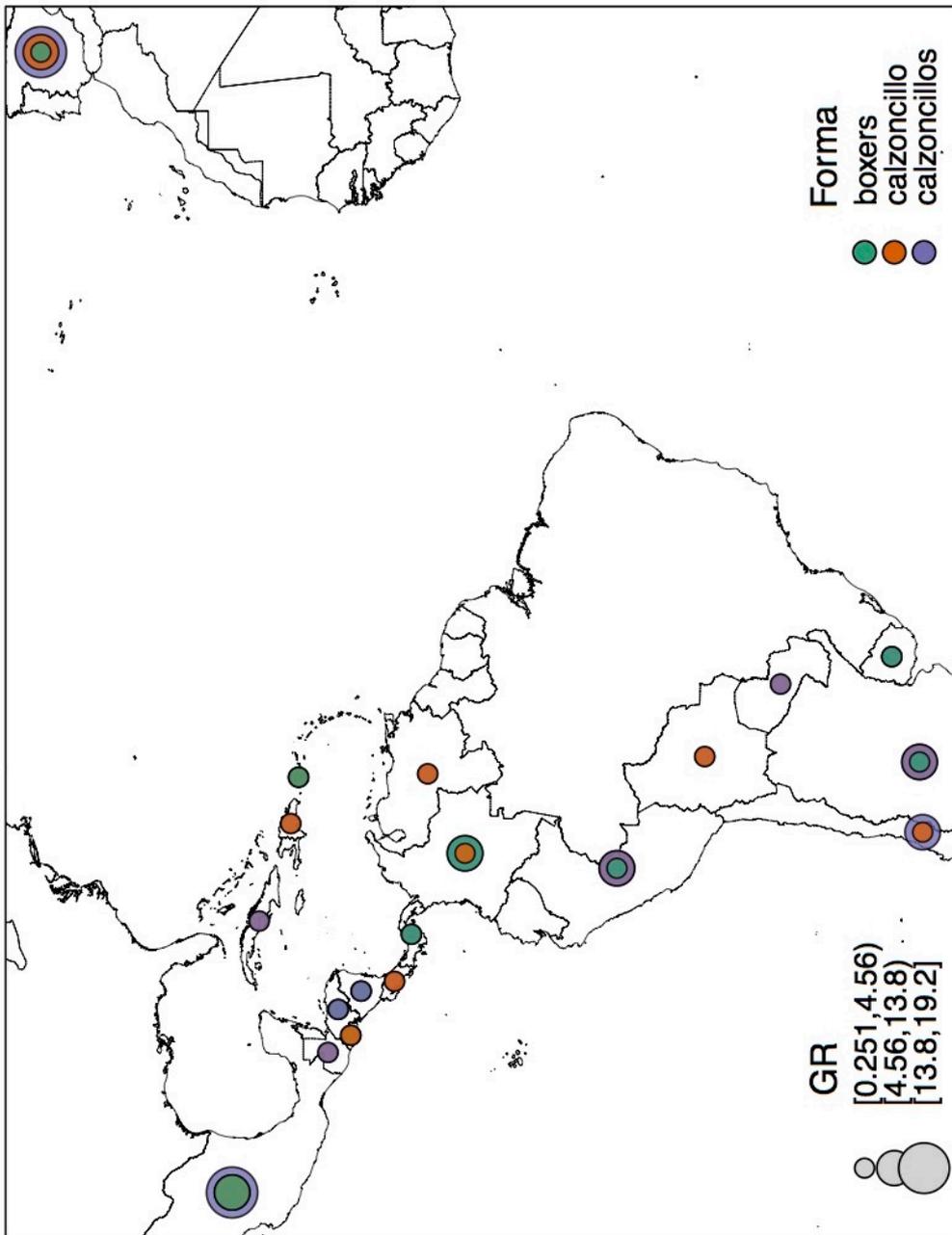
OVERCOAT



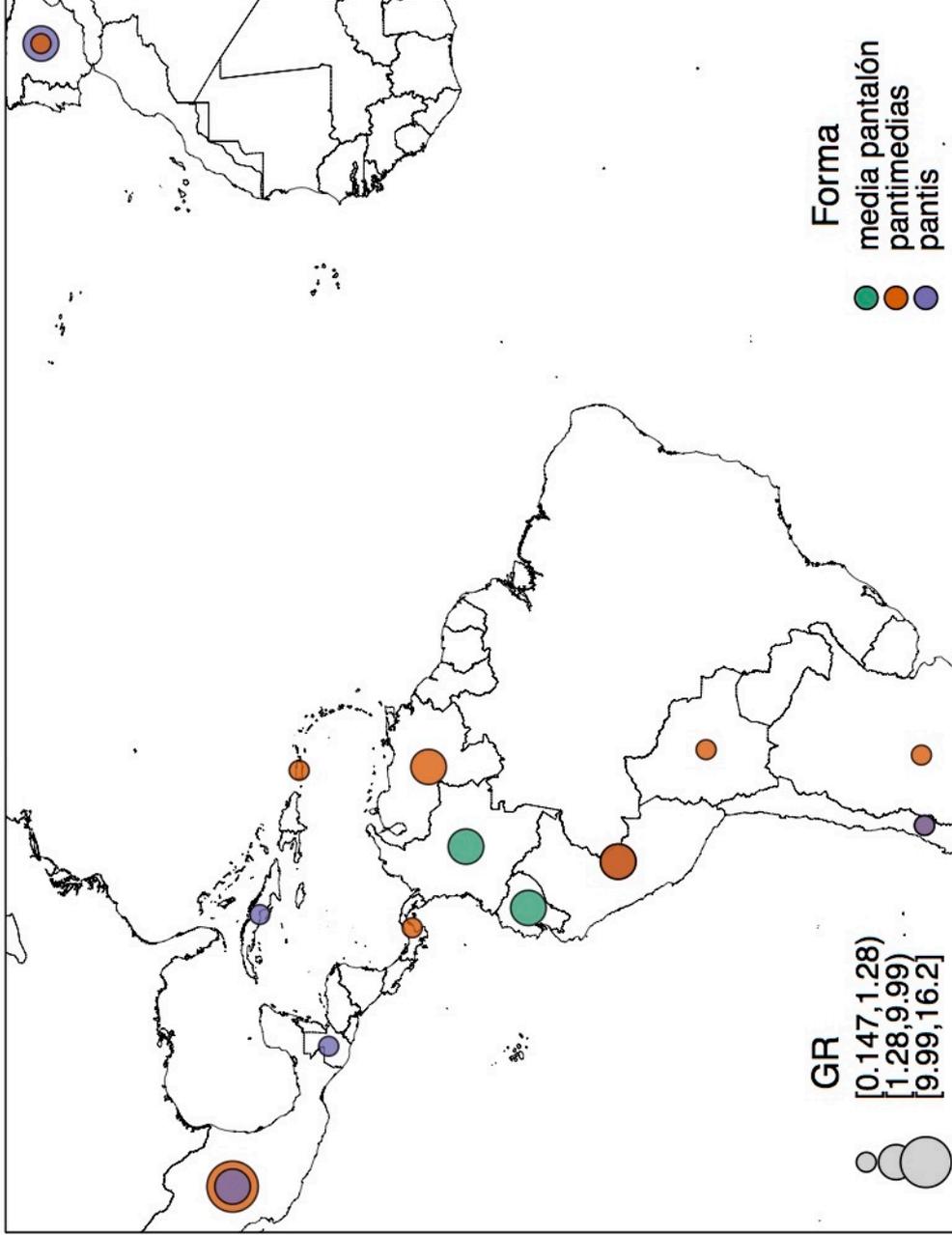
PANTIES



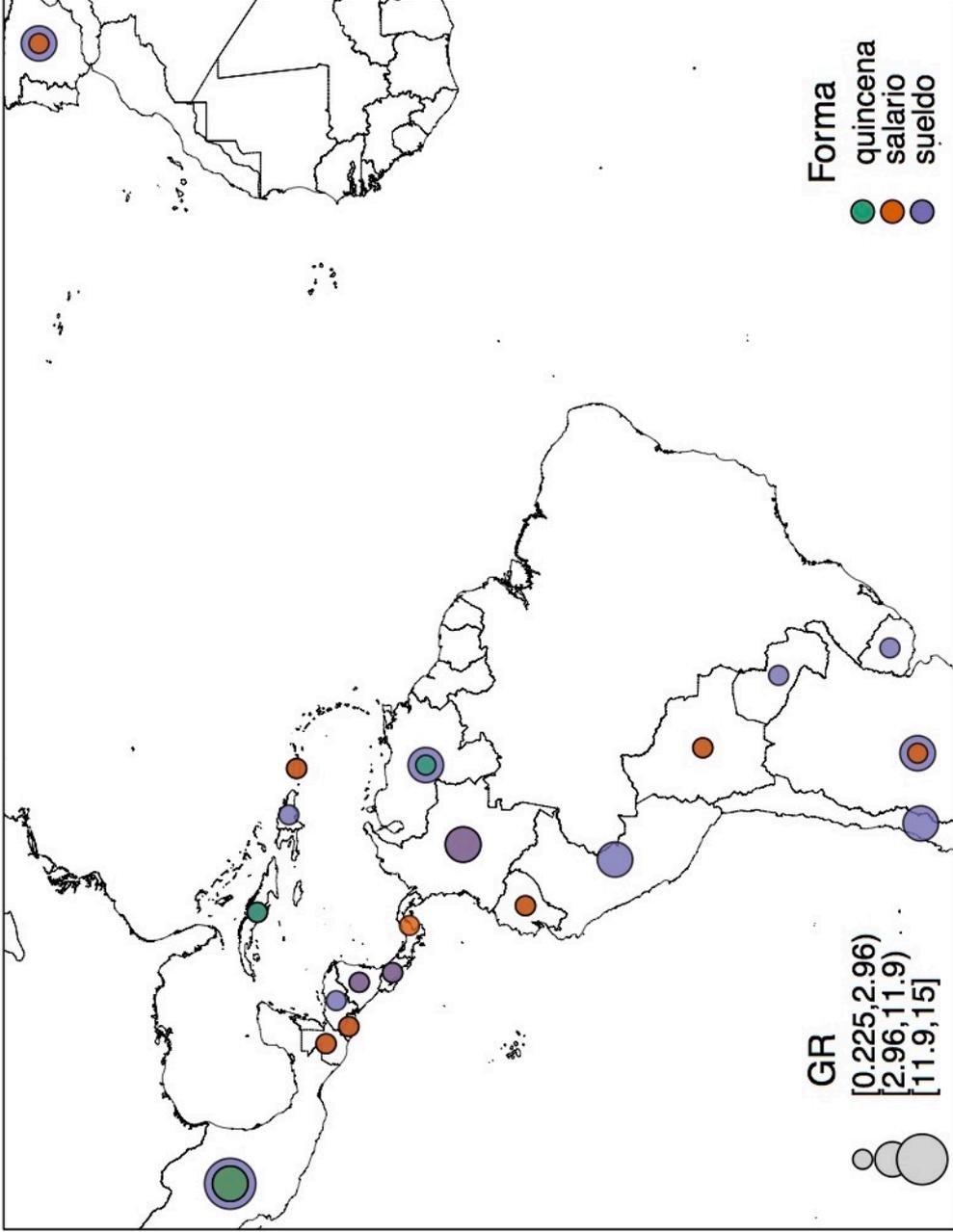
PANTS (US: SHORTS)



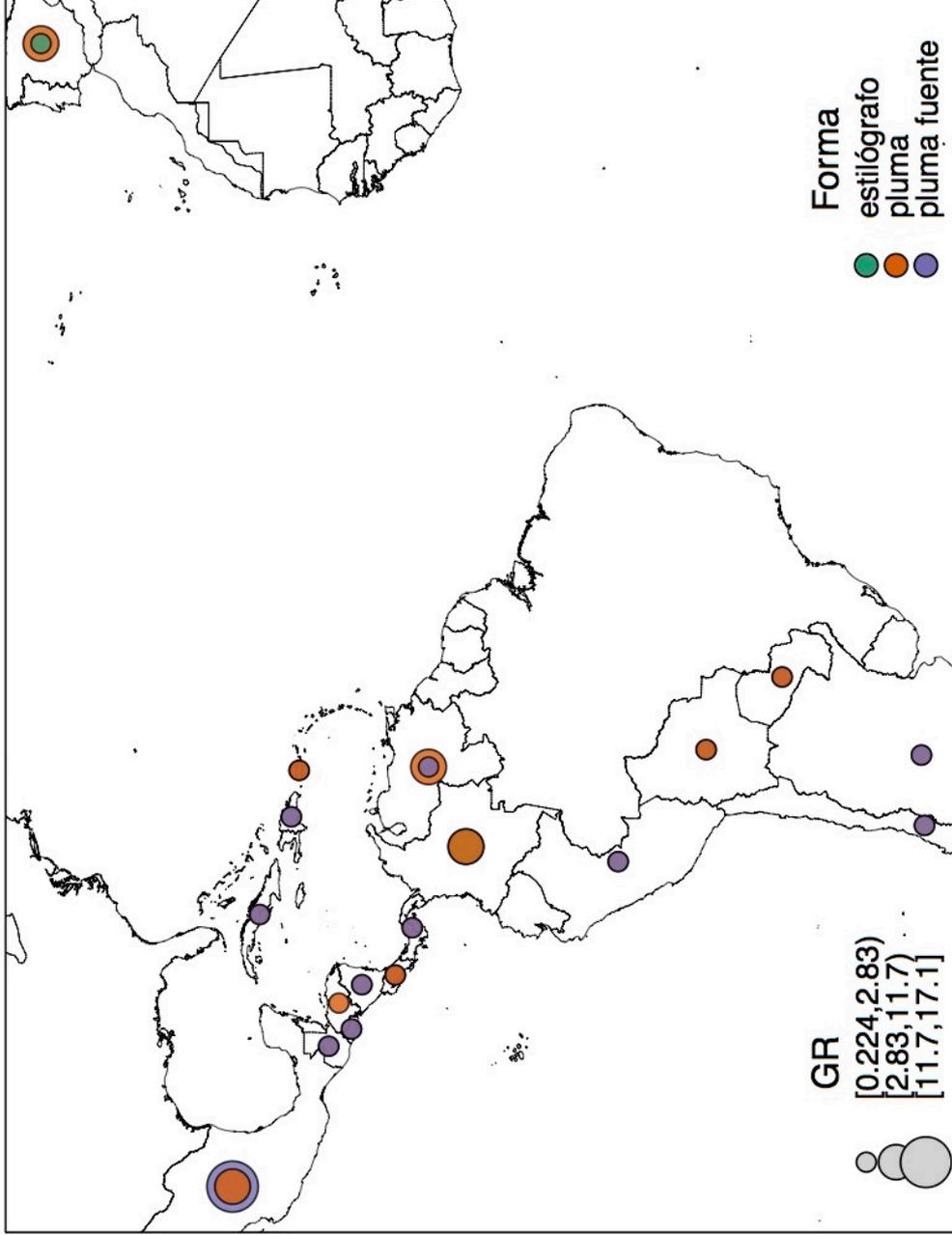
PANTY HOSE



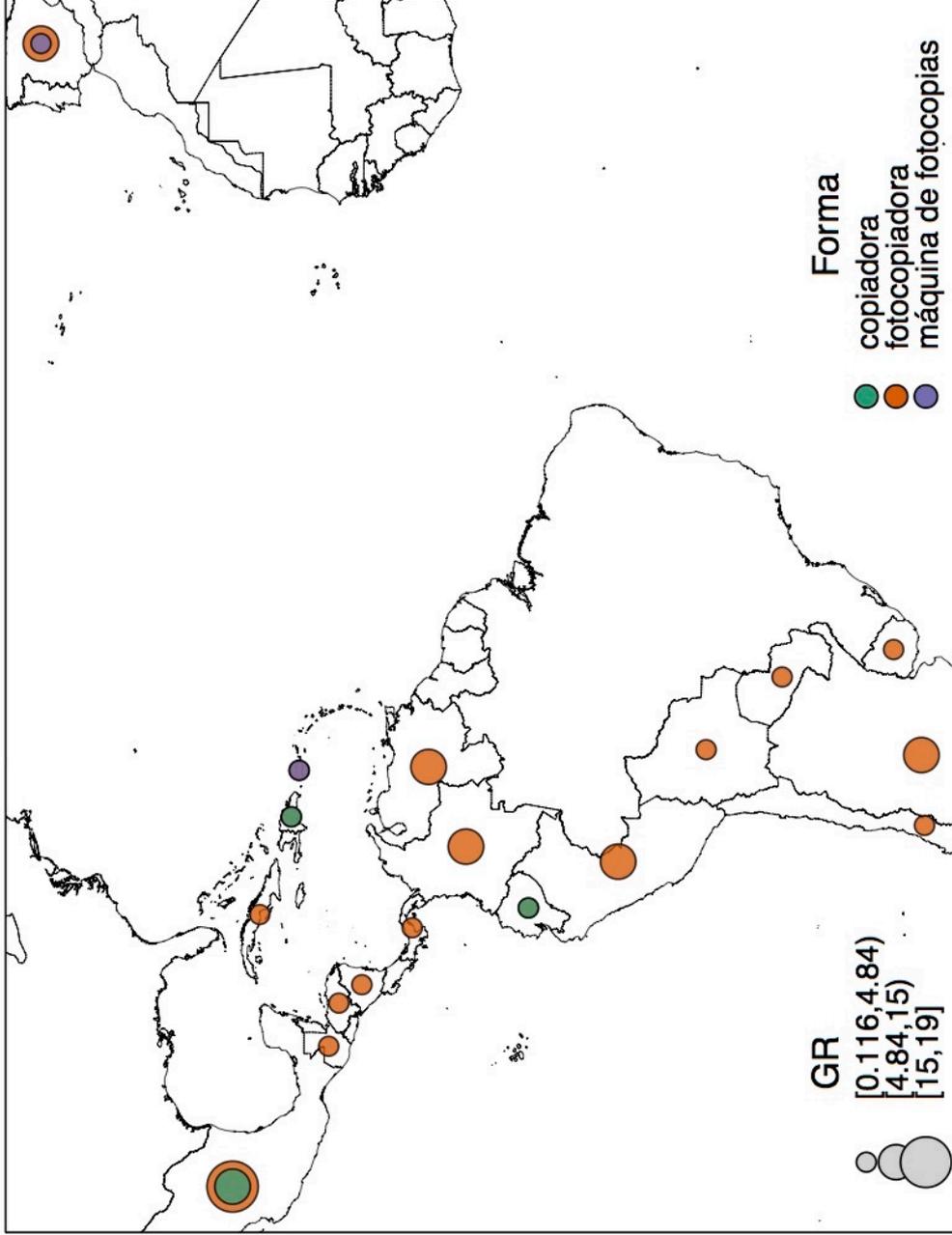
PAY



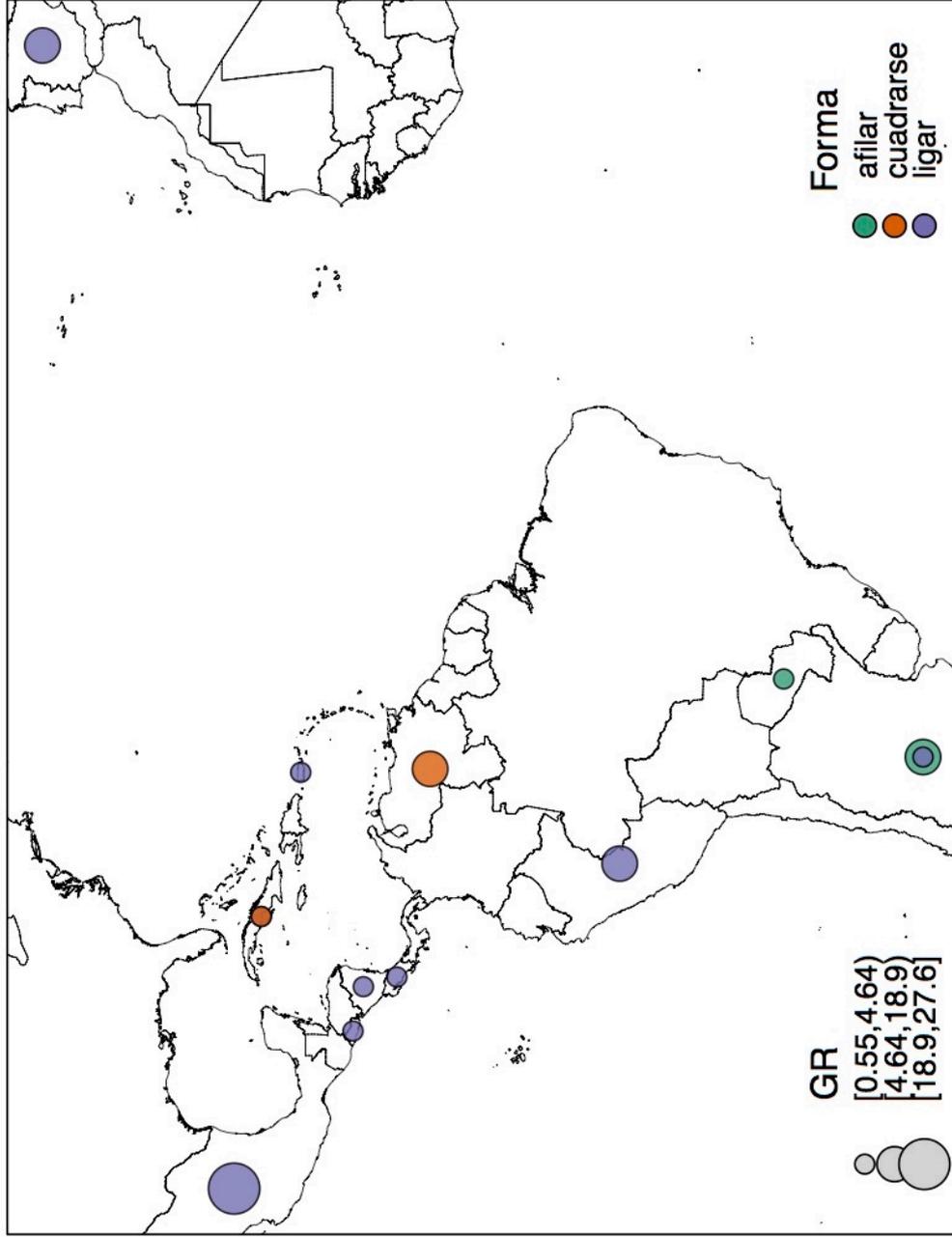
PEN



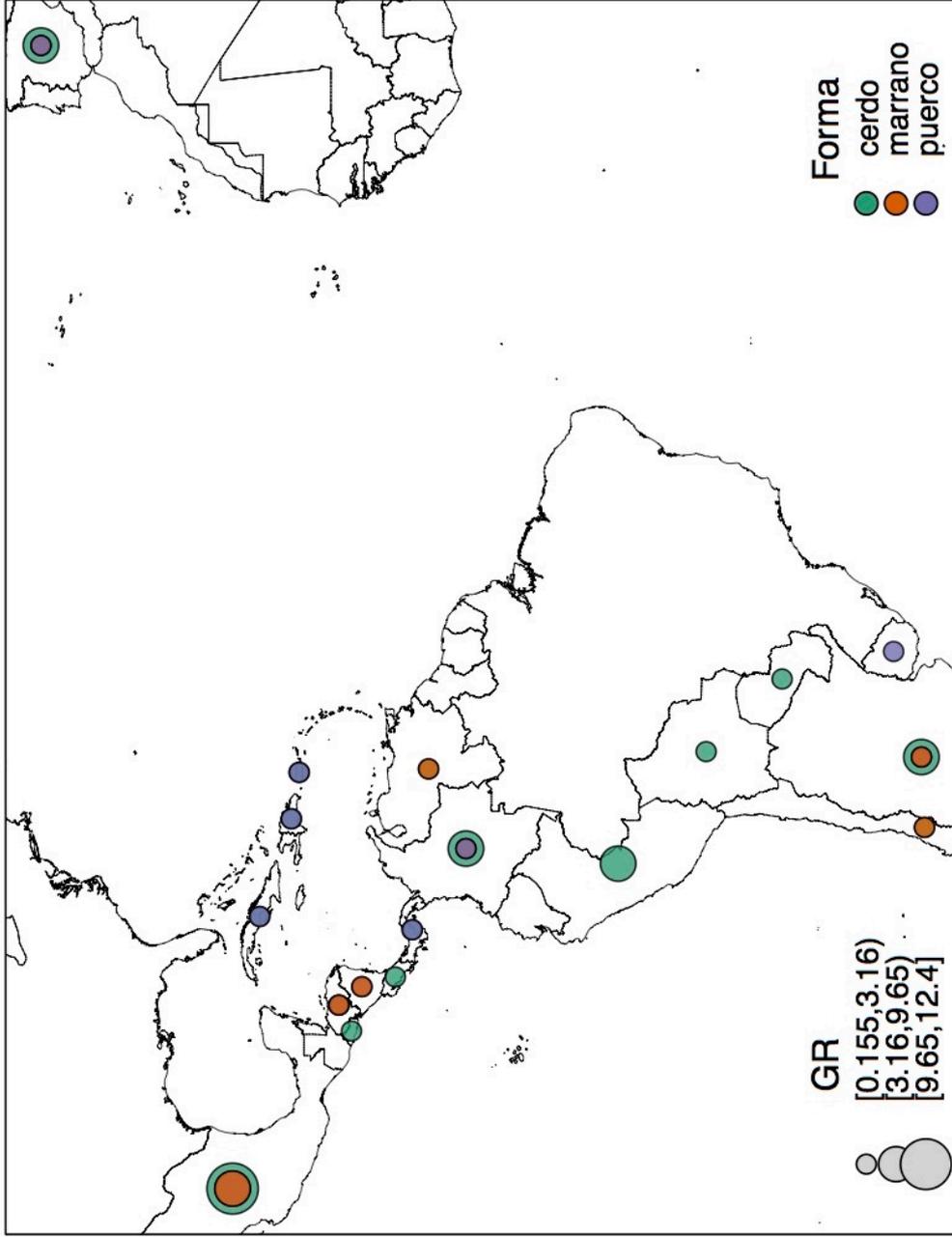
PHOTOCOPIER



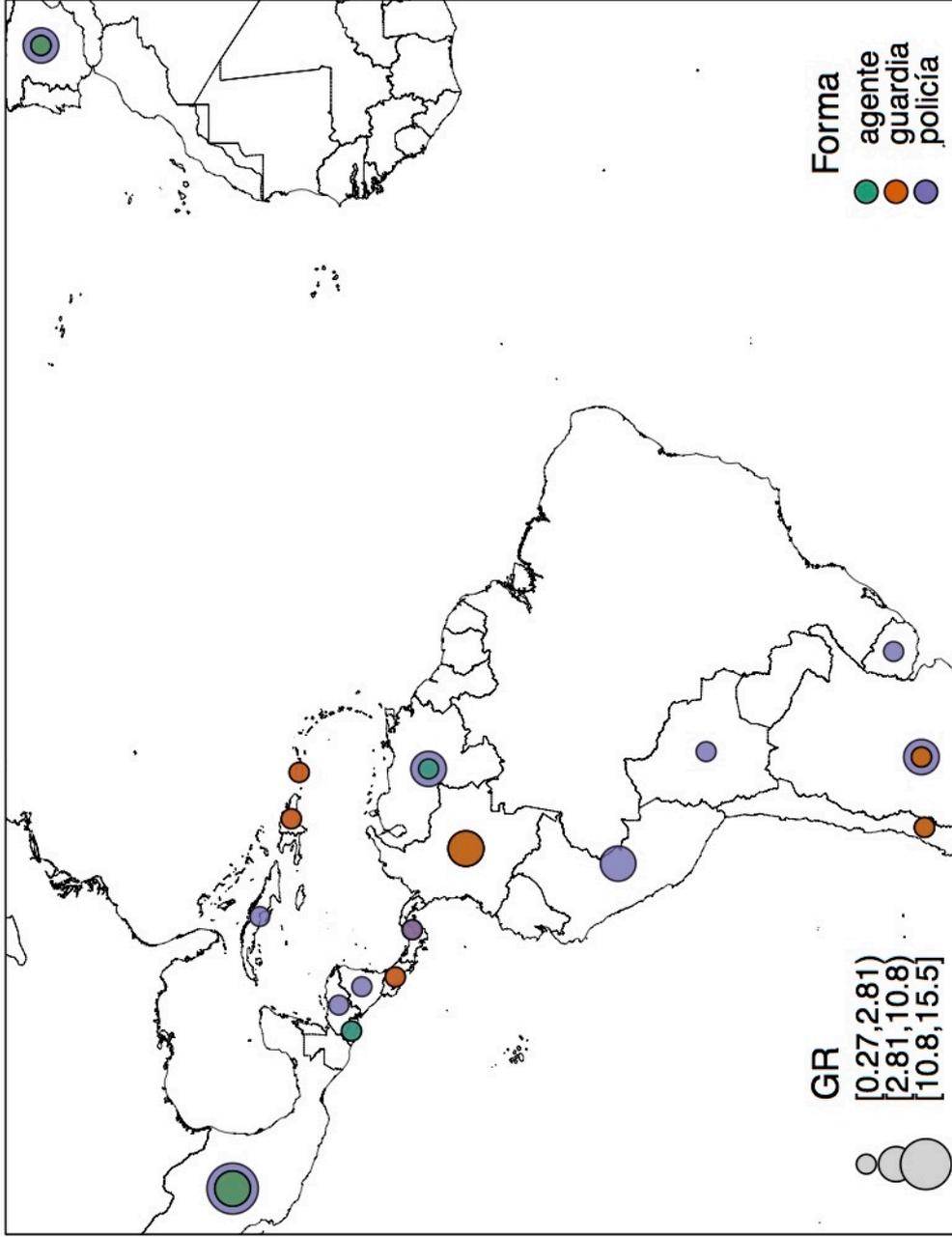
PICK A GIRL UP (COL.)



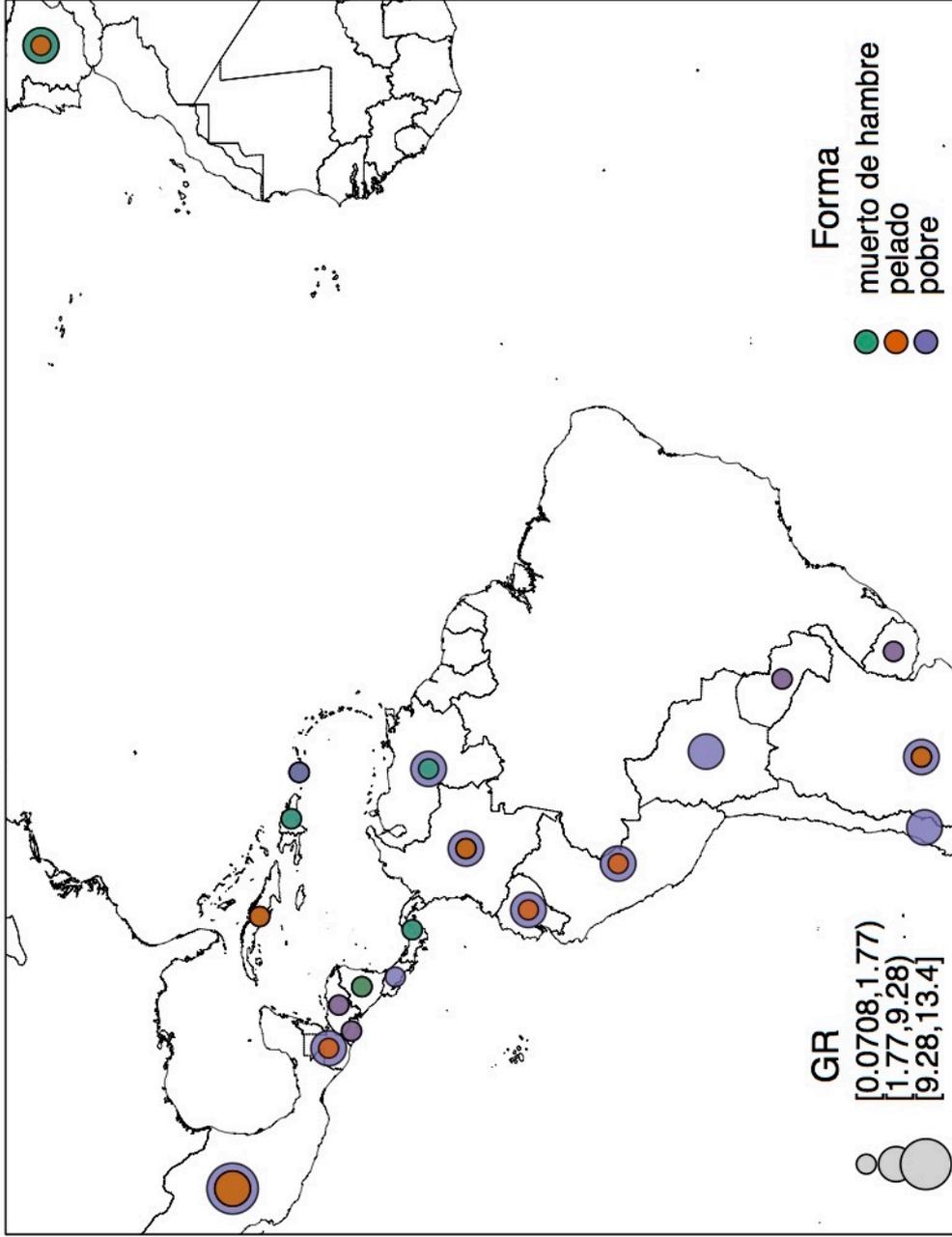
PIG



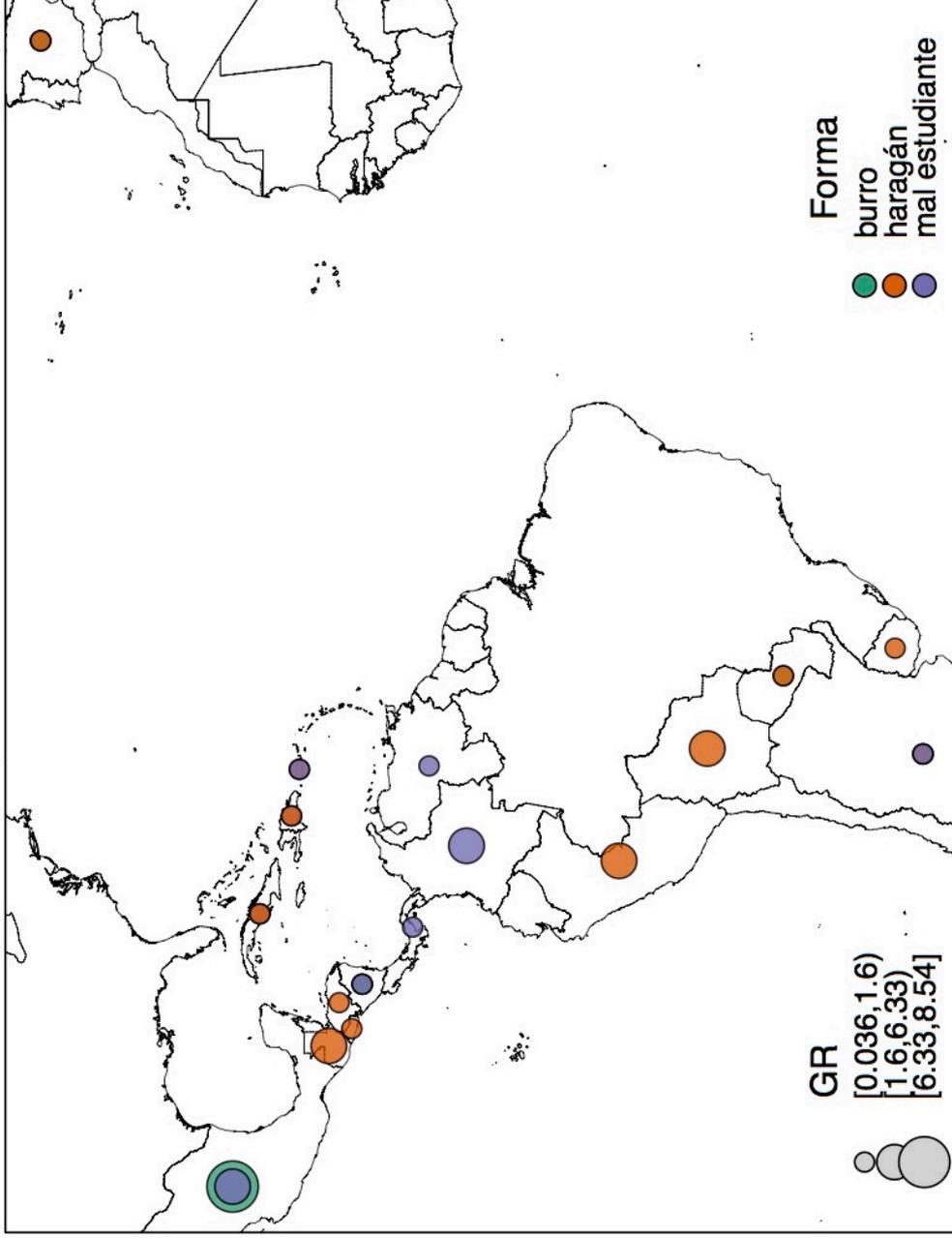
POLICEMAN



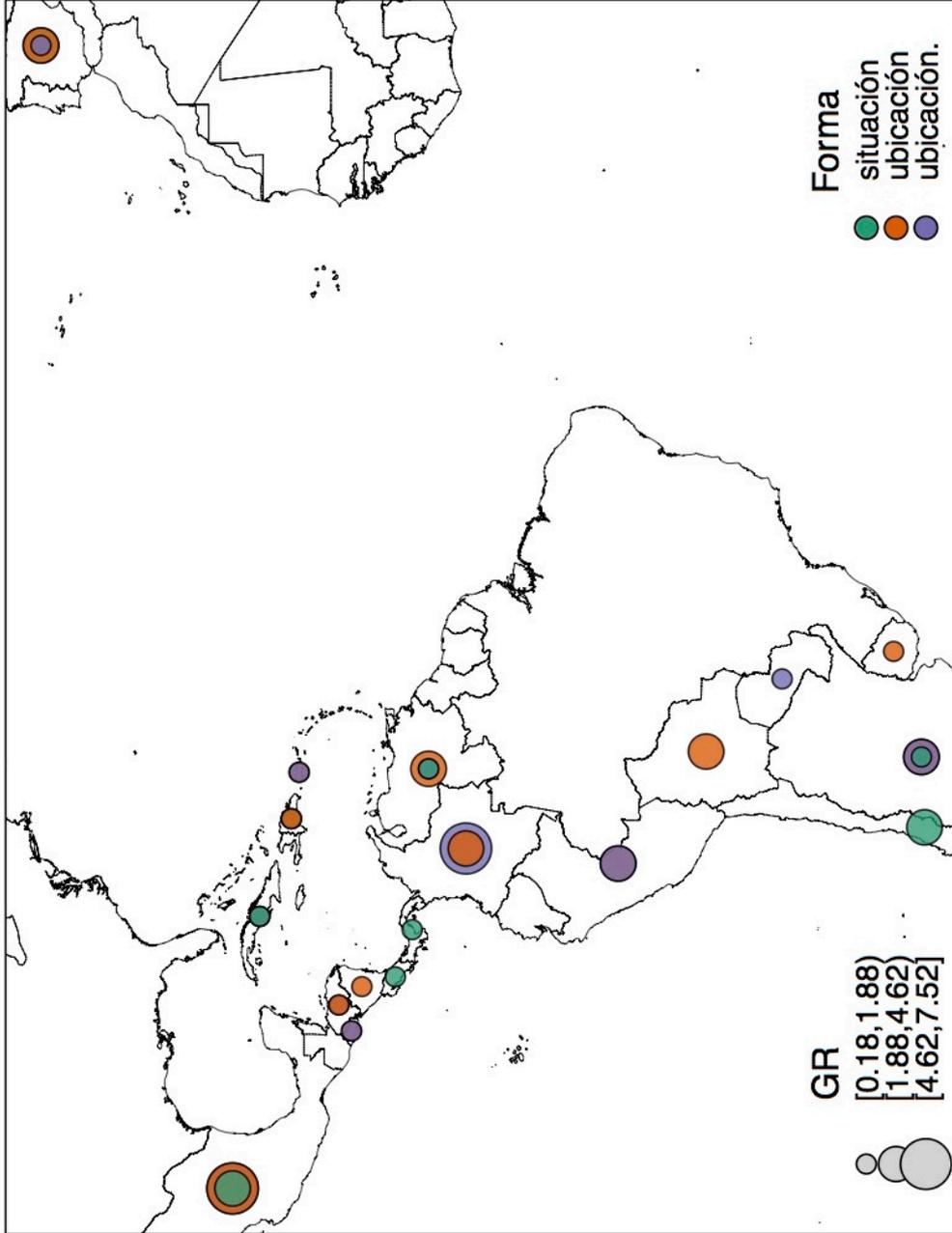
POOR



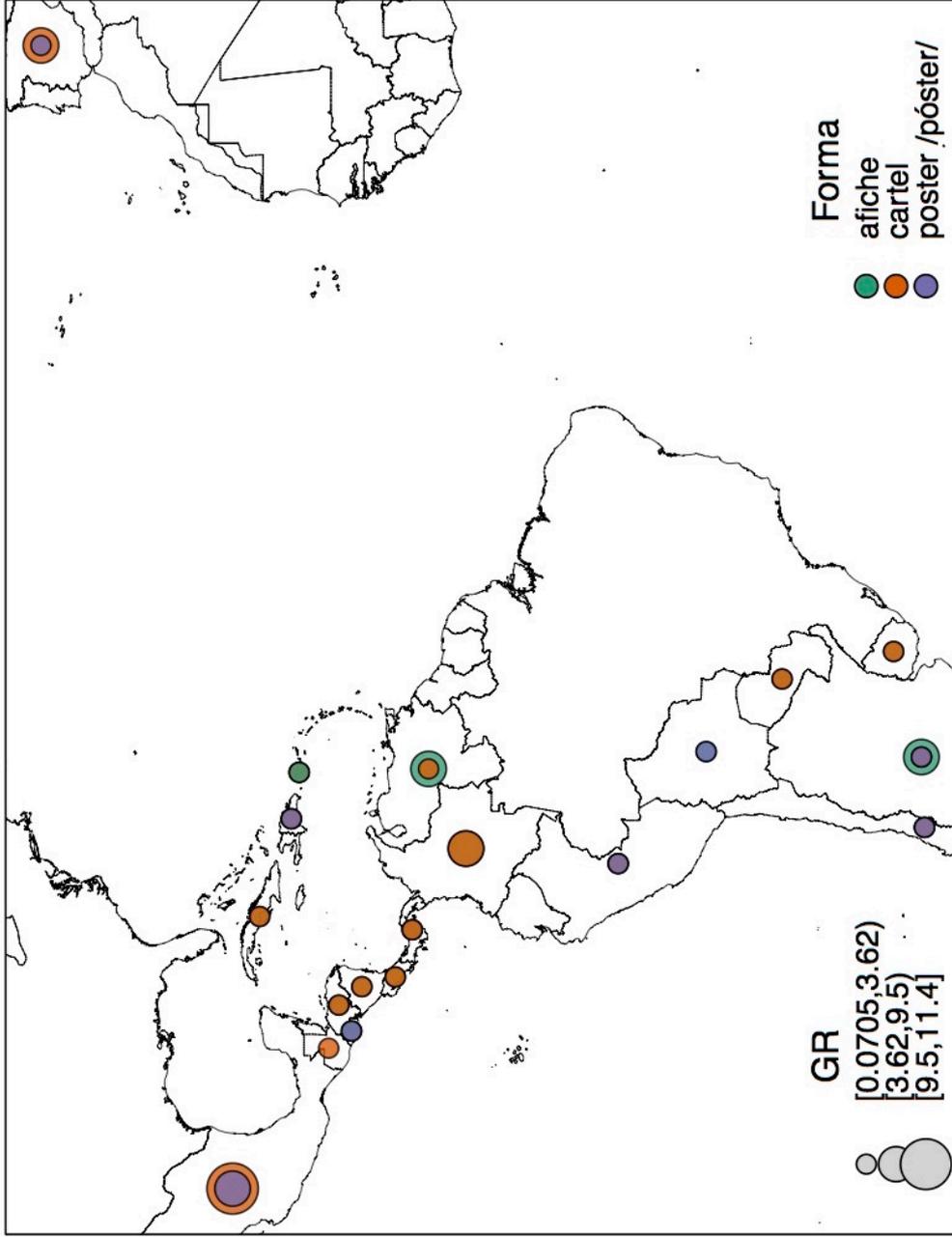
POOR STUDENT



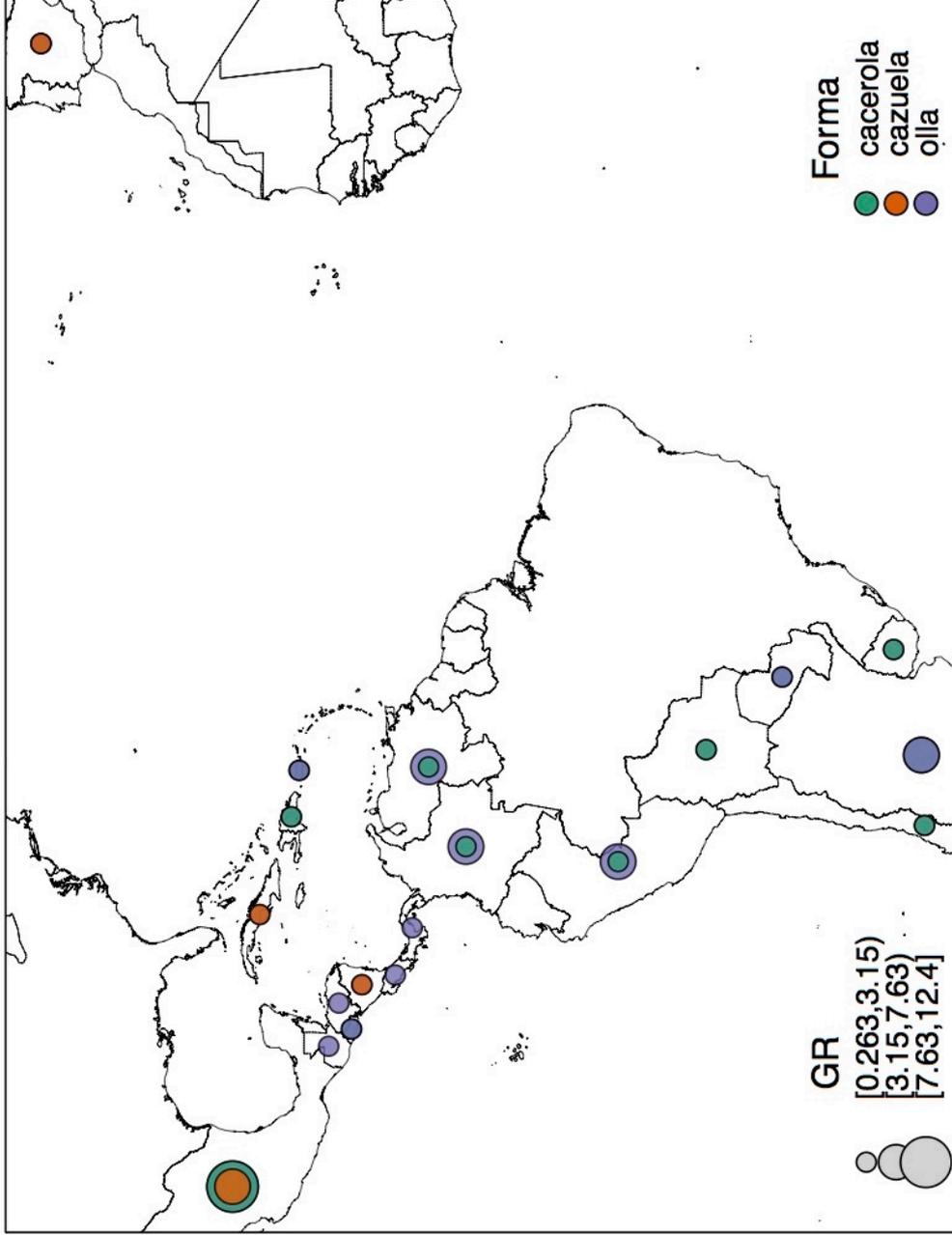
POSITION



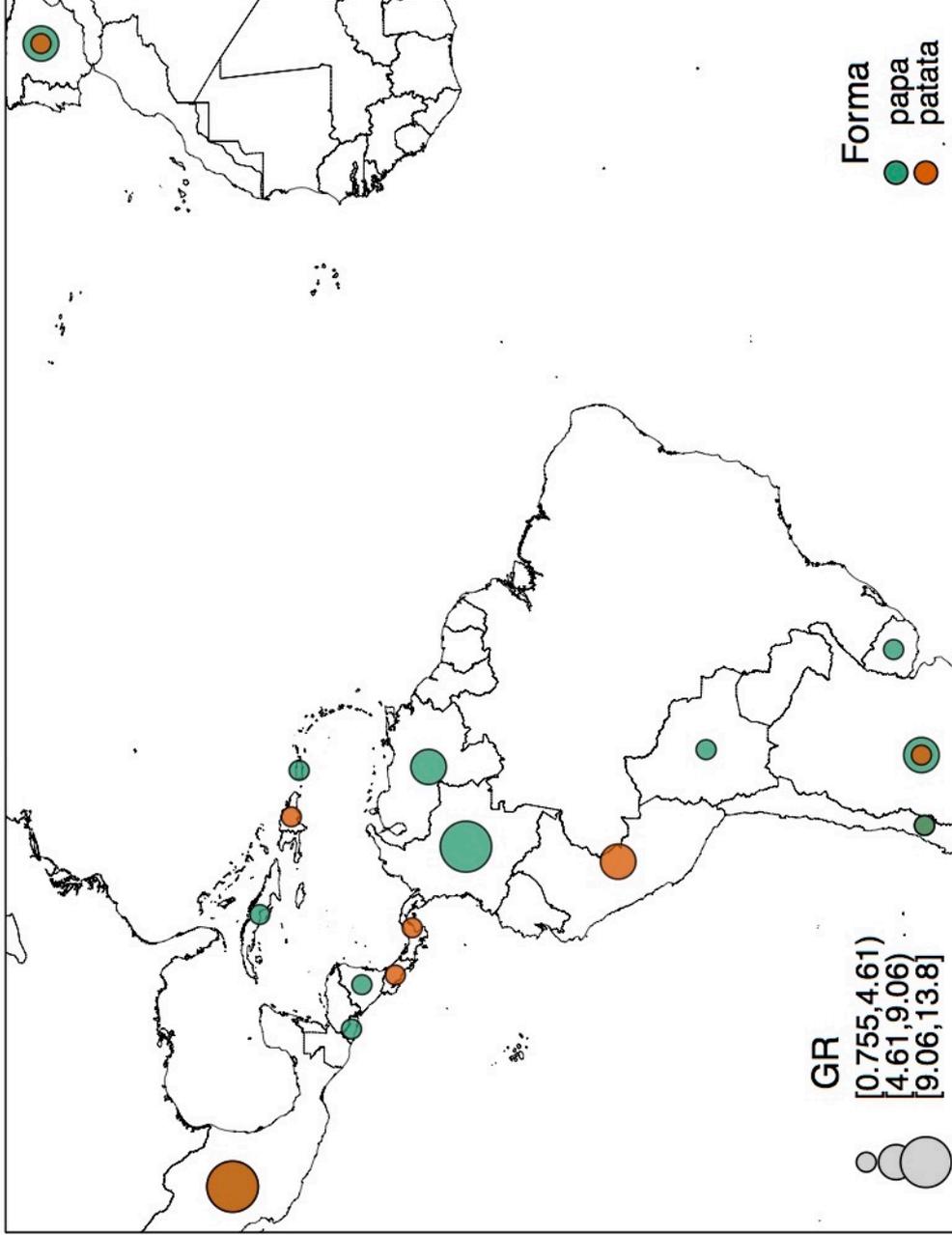
POSTER



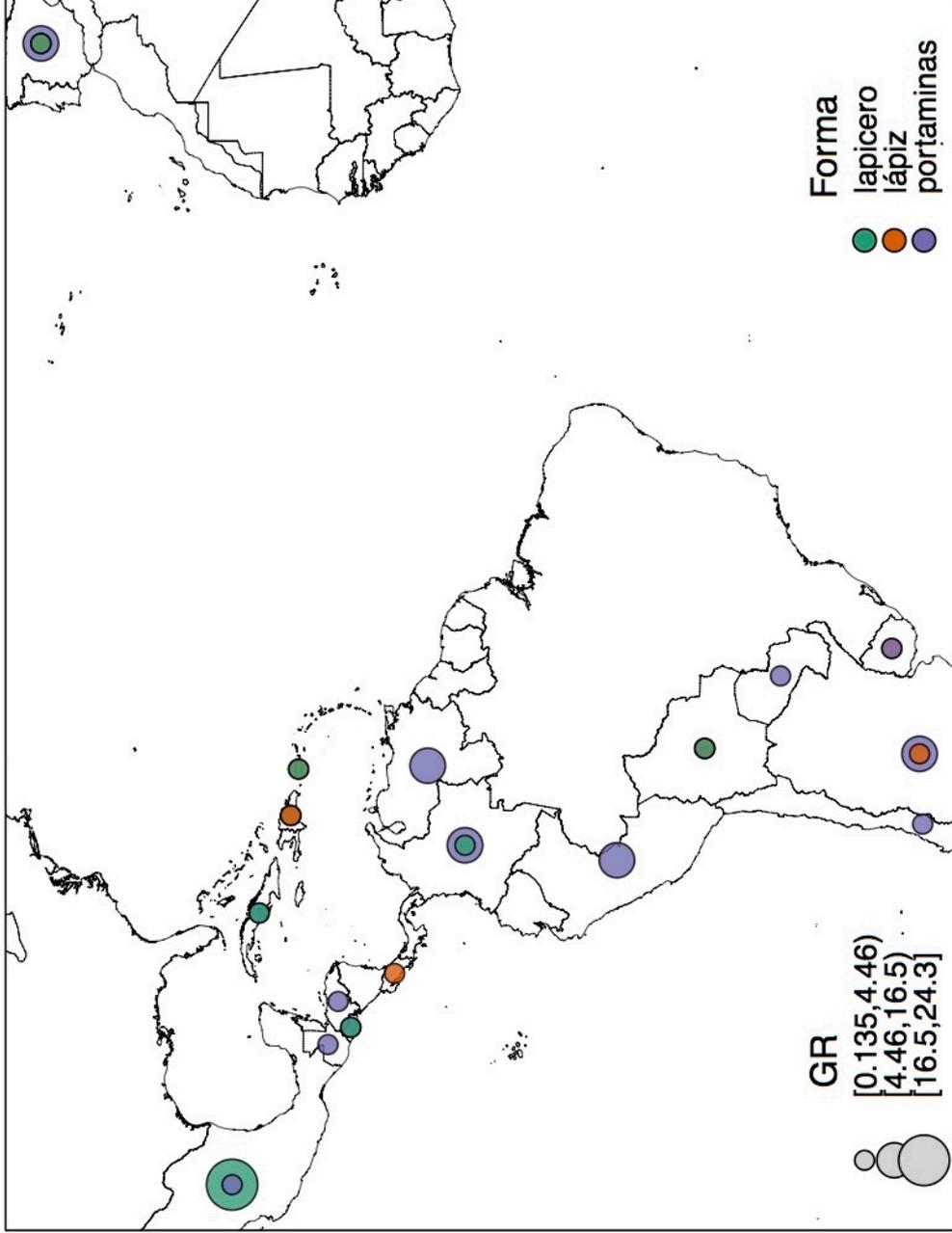
POT



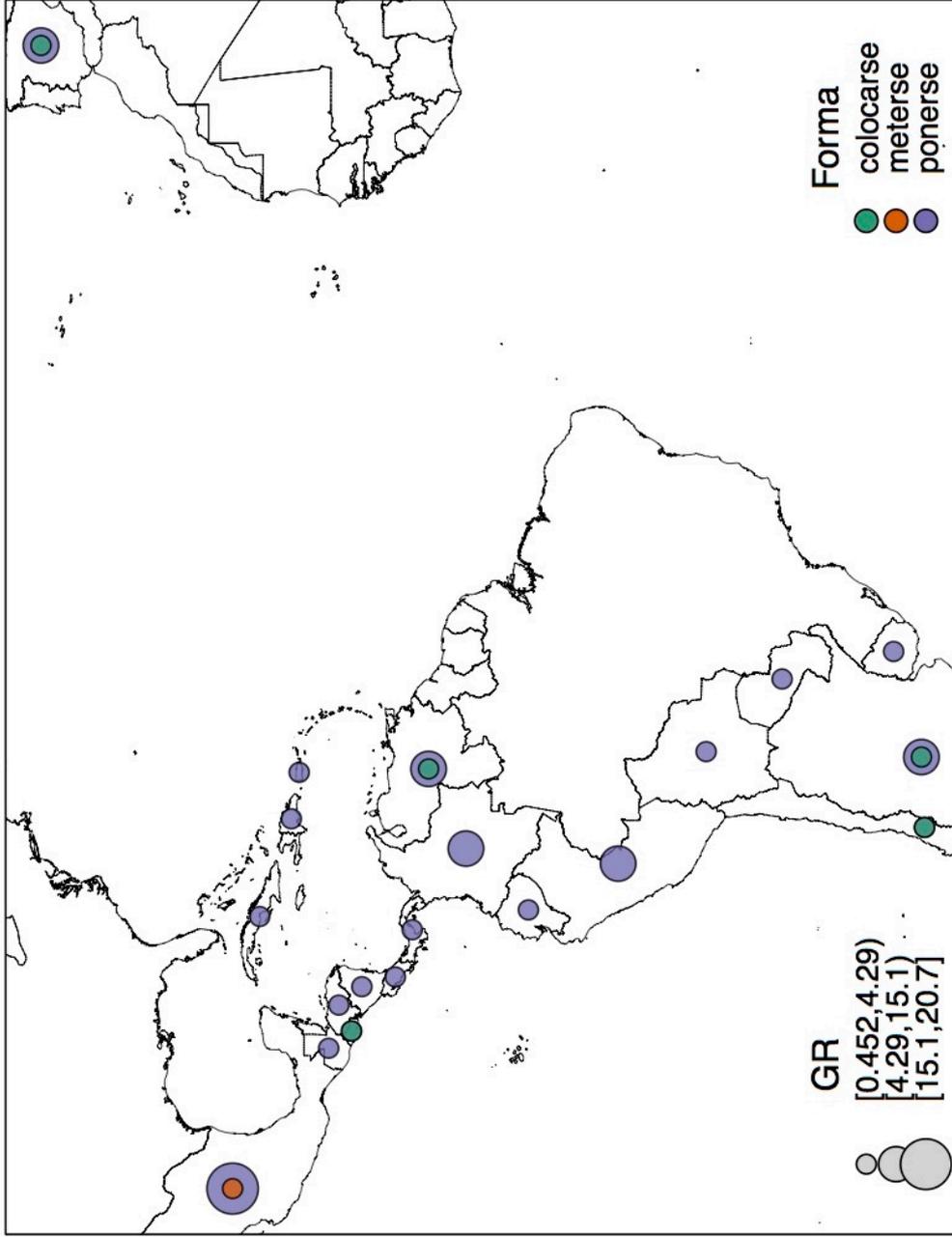
POTATO



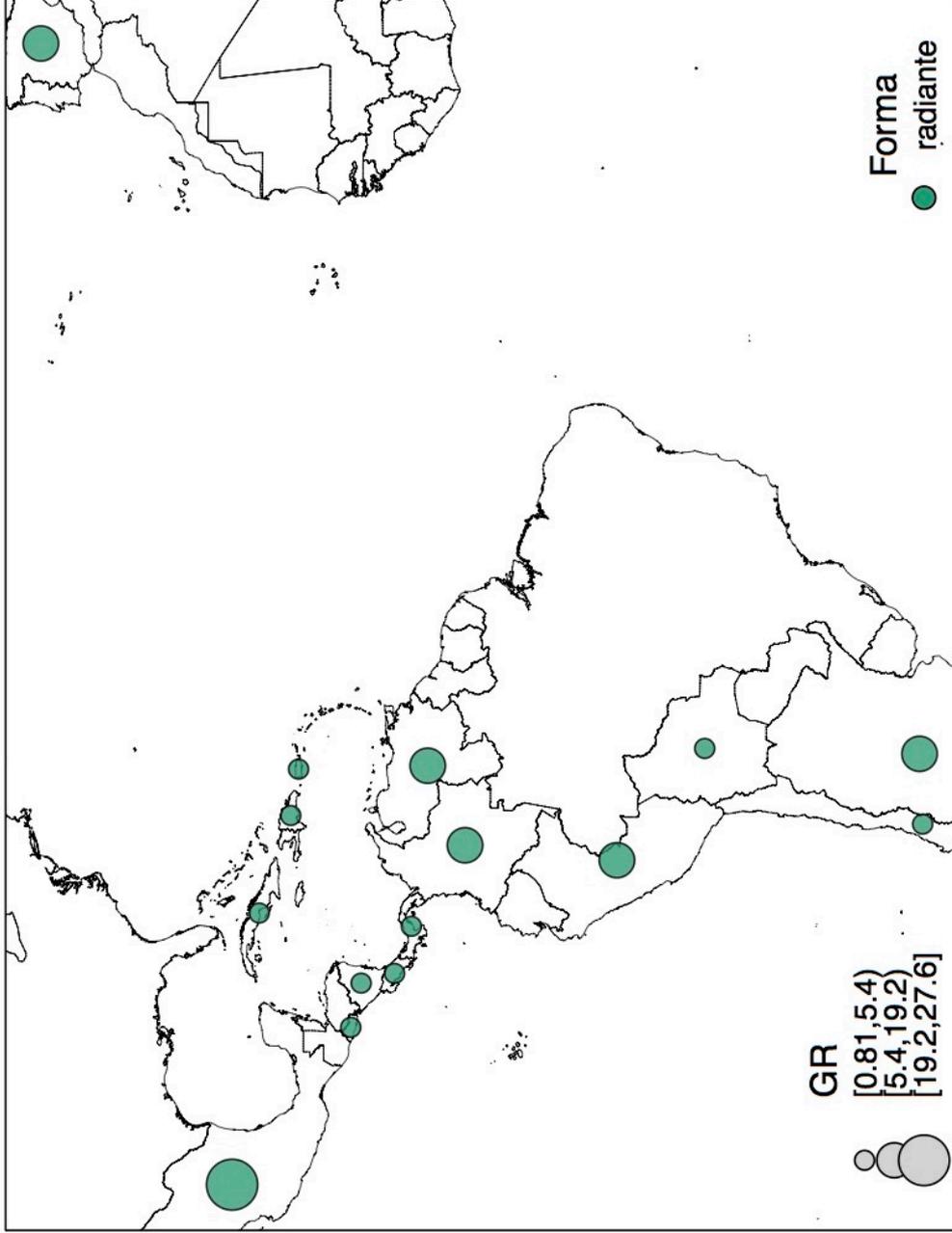
PROPELLING PENCIL (US: MECHANICAL PENCIL)



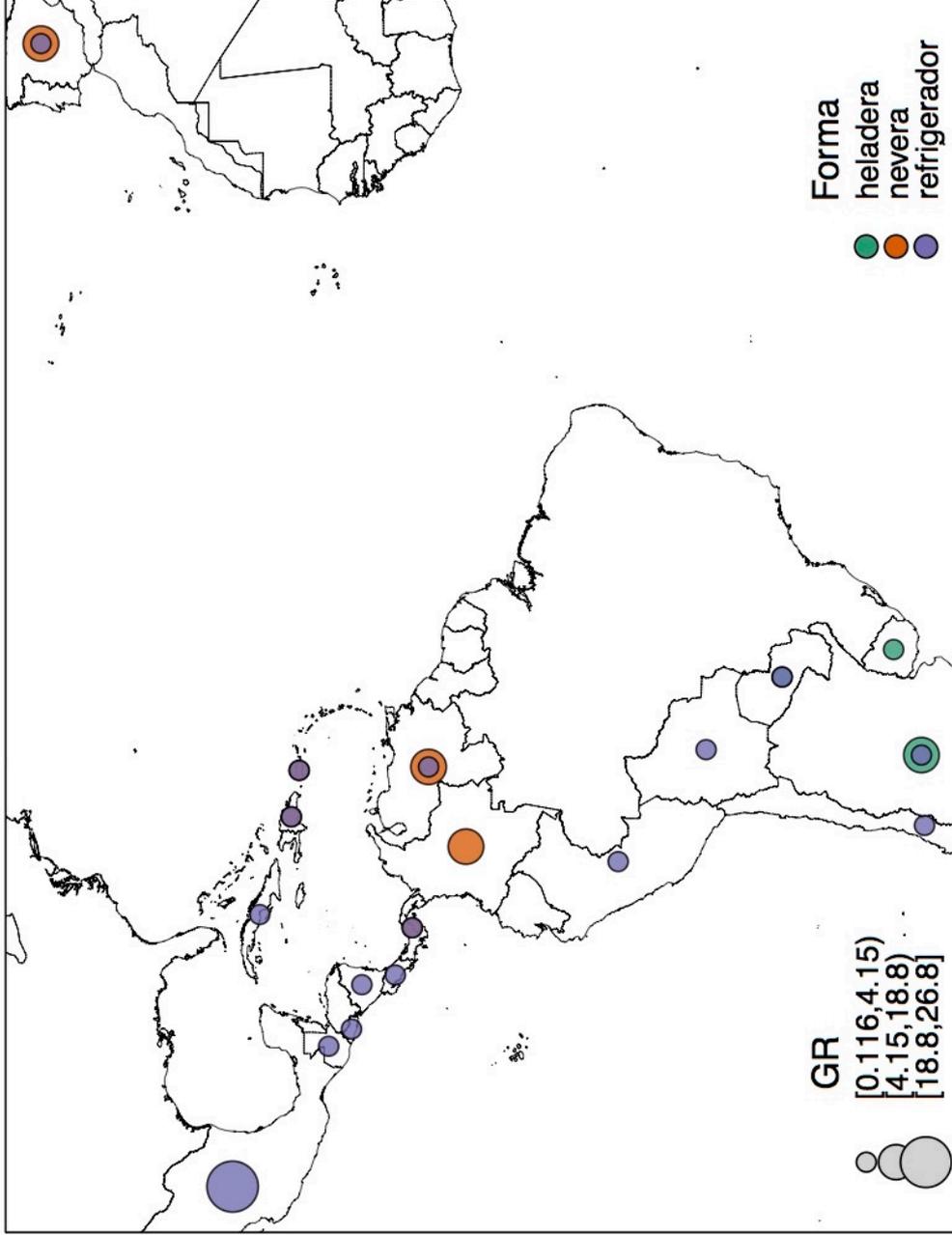
PUT ON



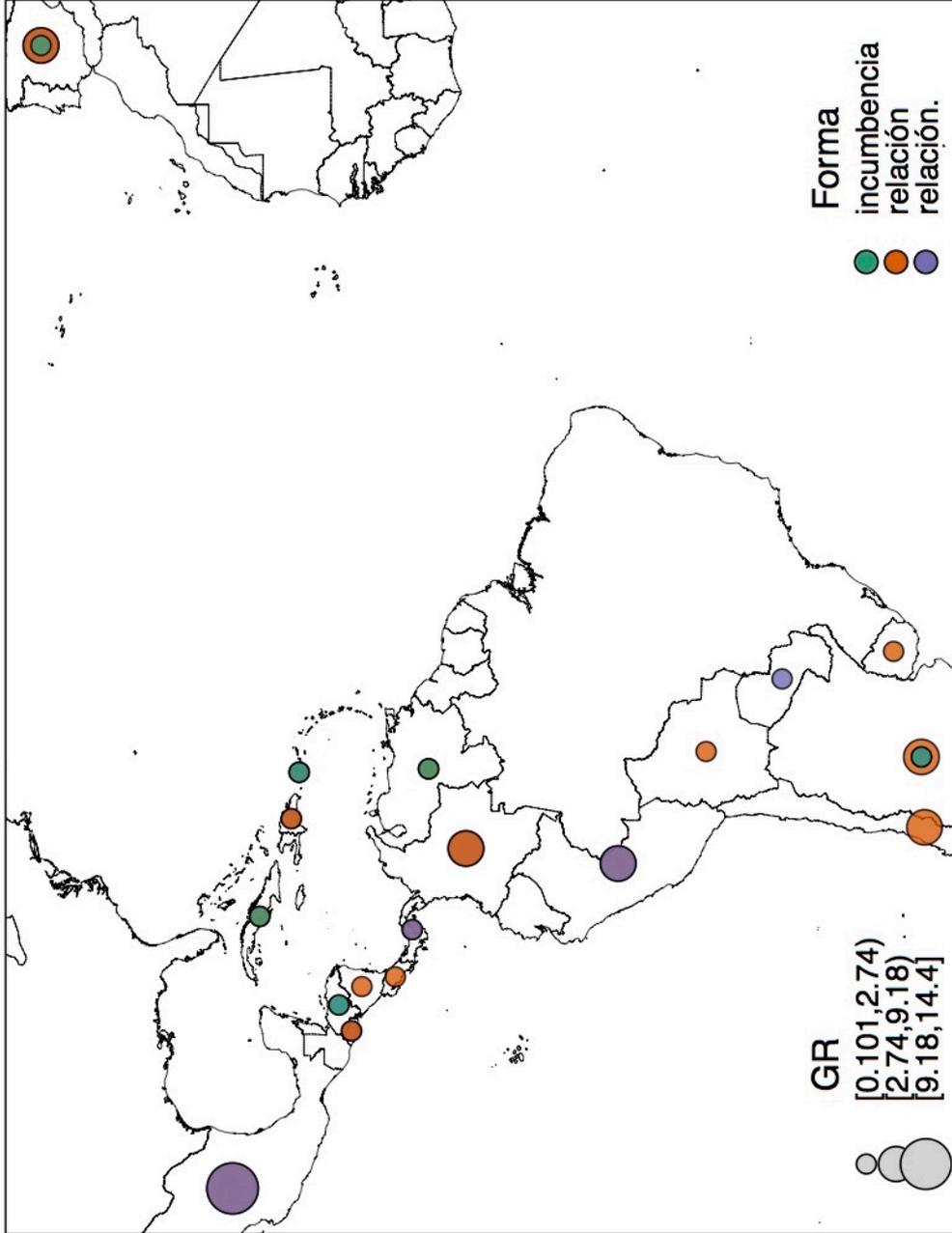
RADIANT



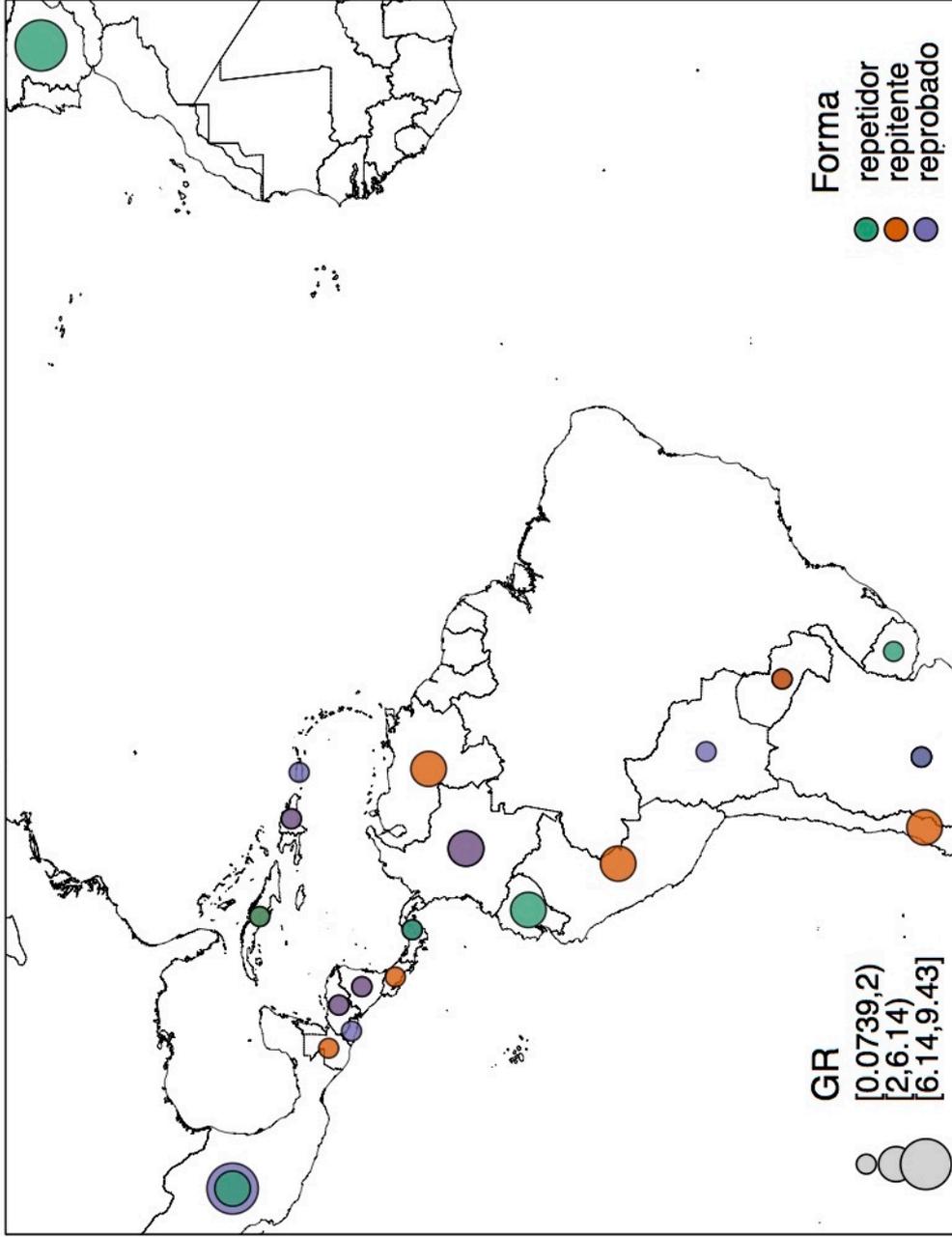
REFRIGERATOR



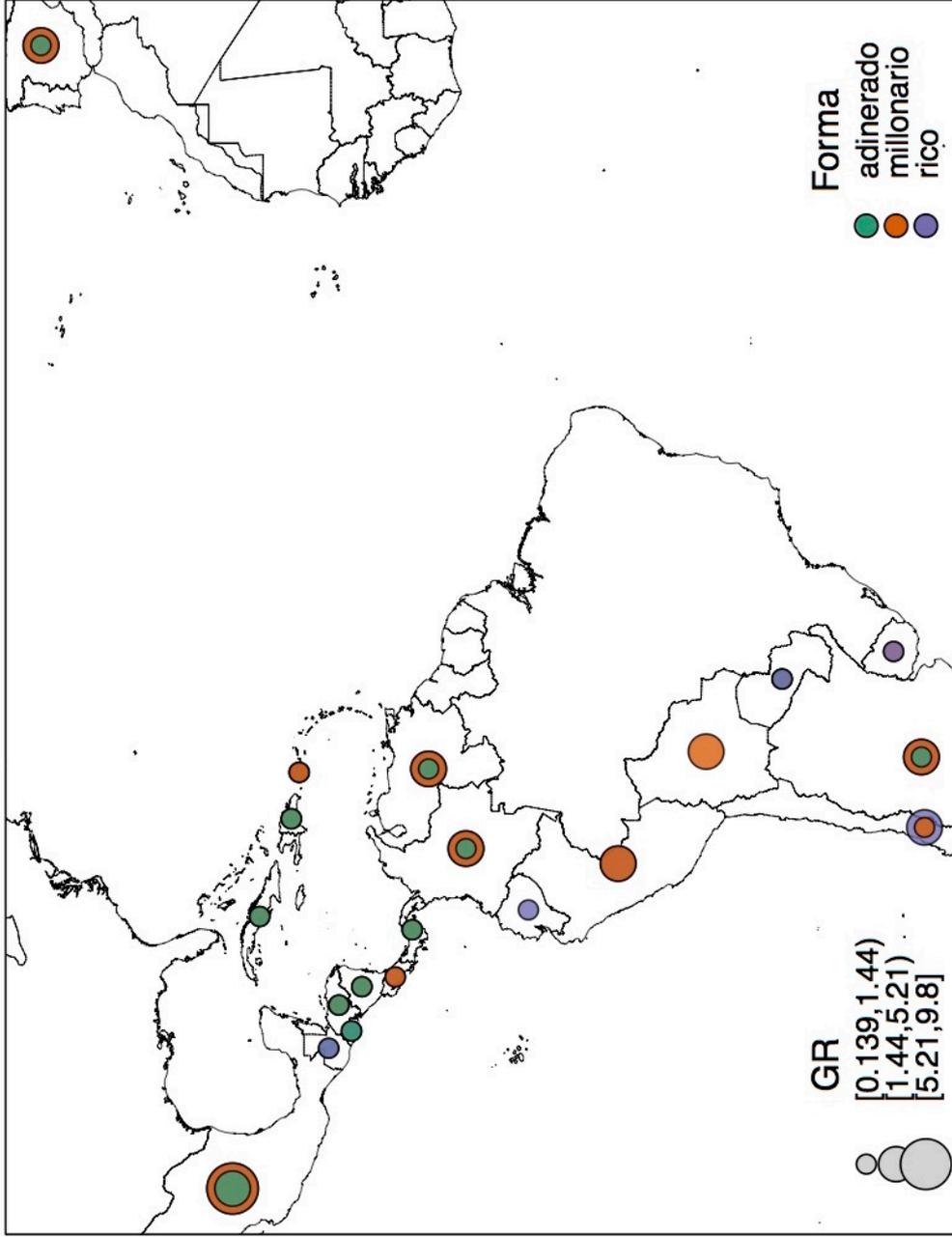
RELATION



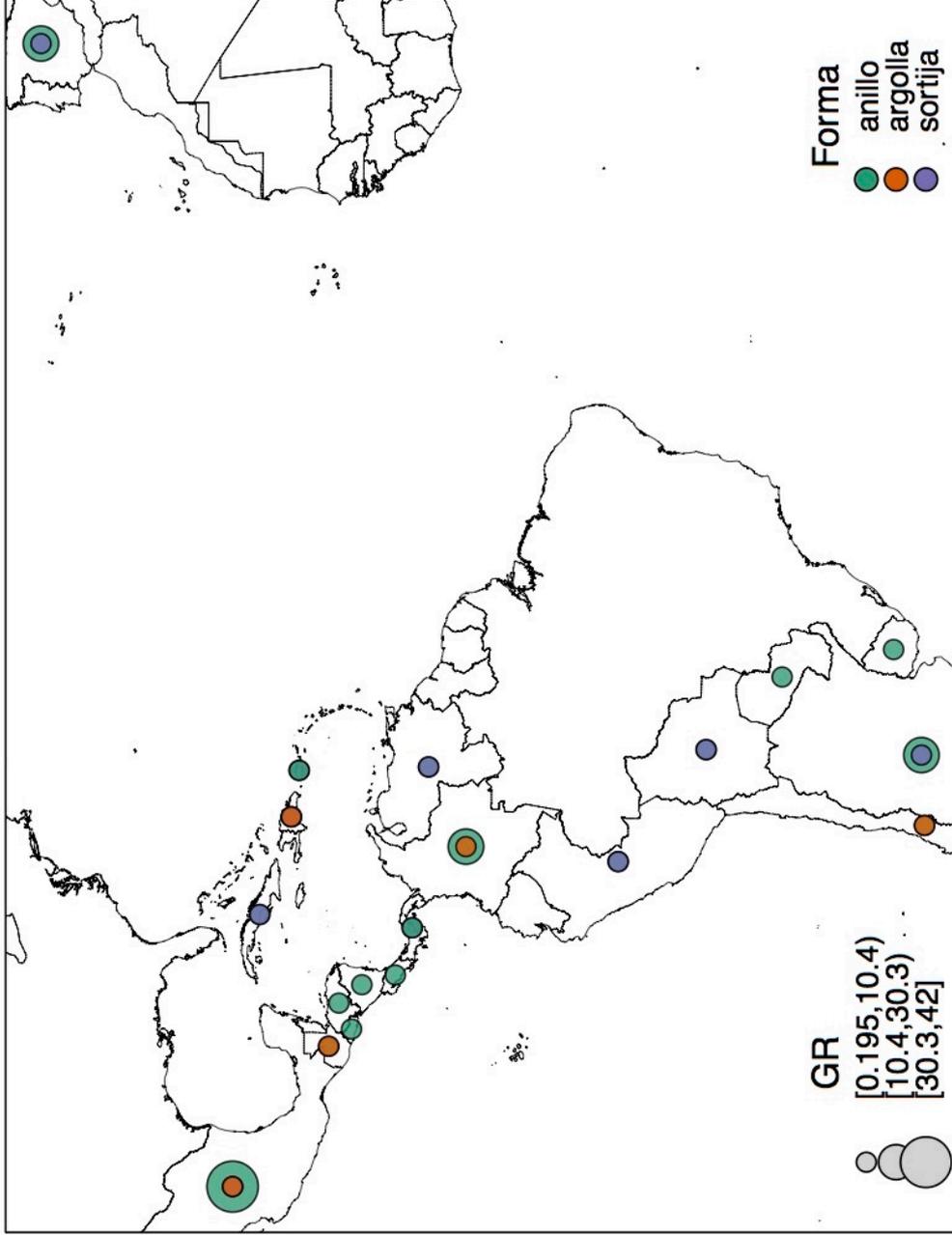
REPEATER



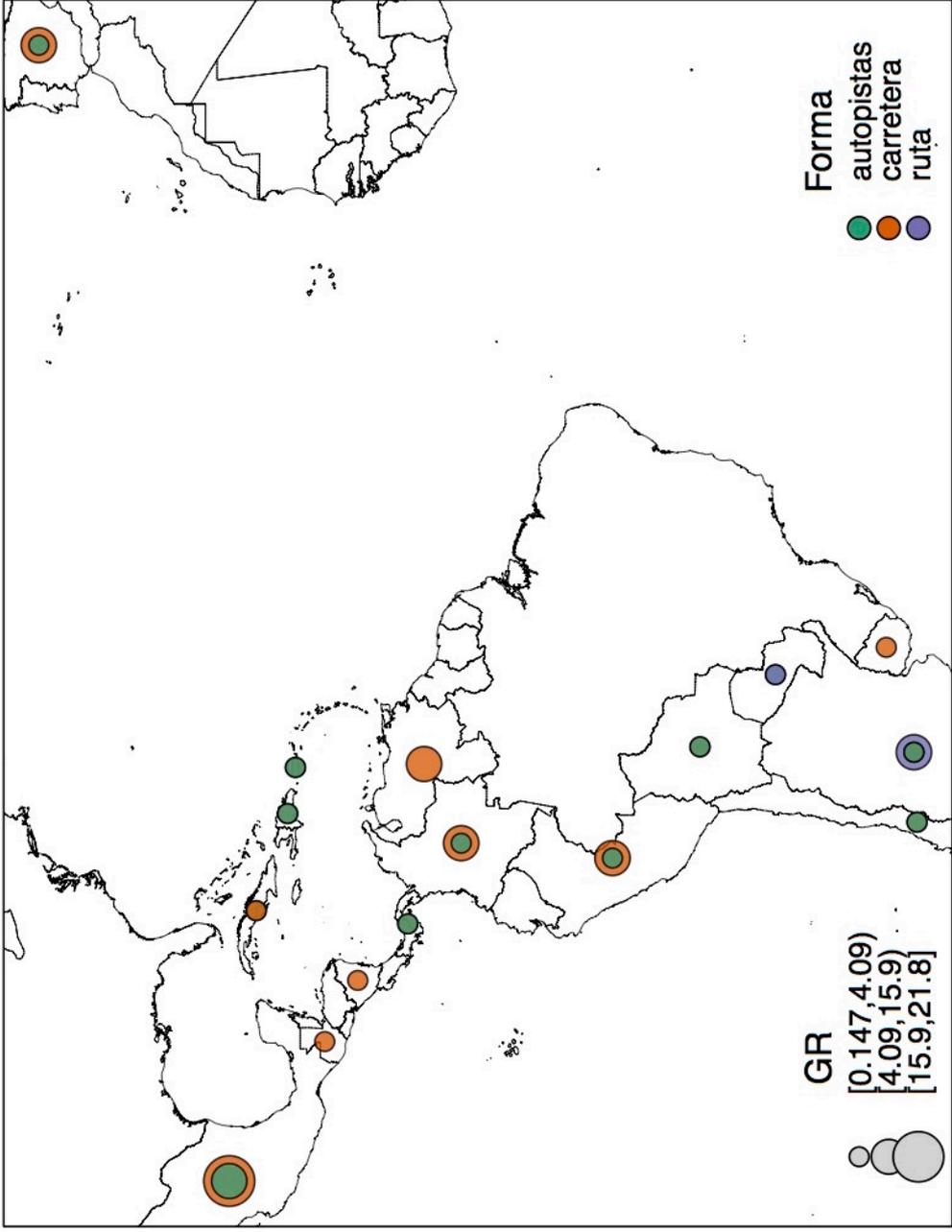
RICH



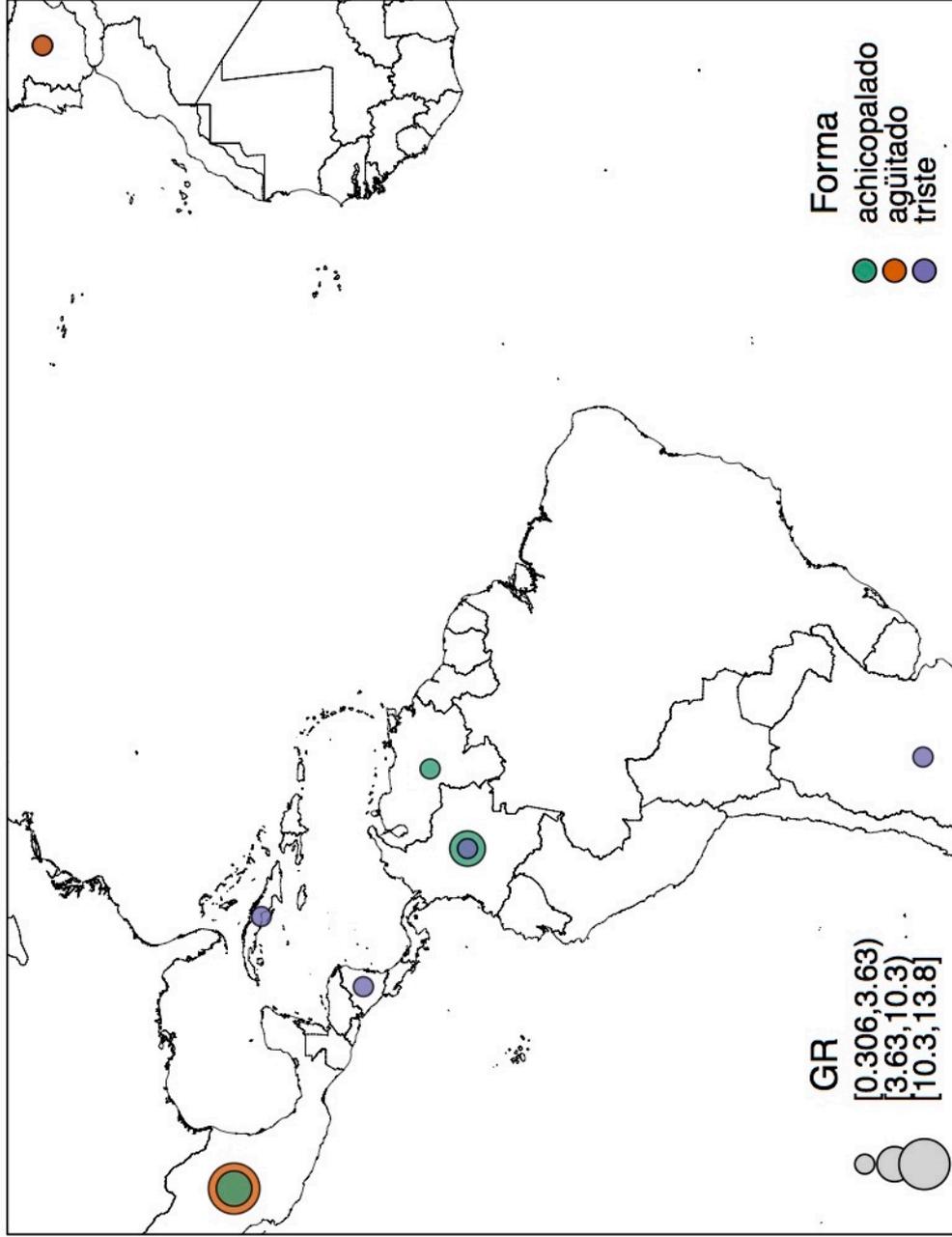
RING



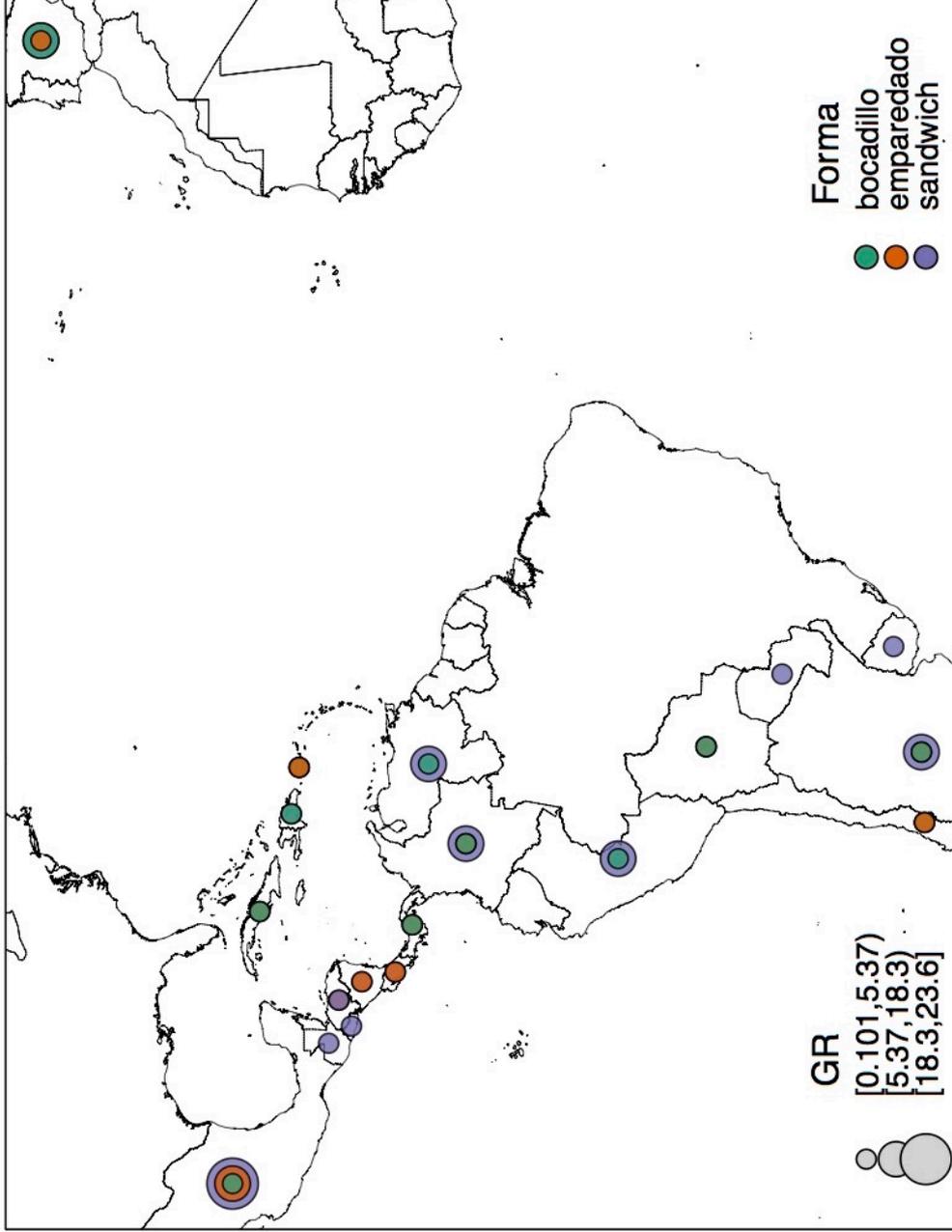
ROAD



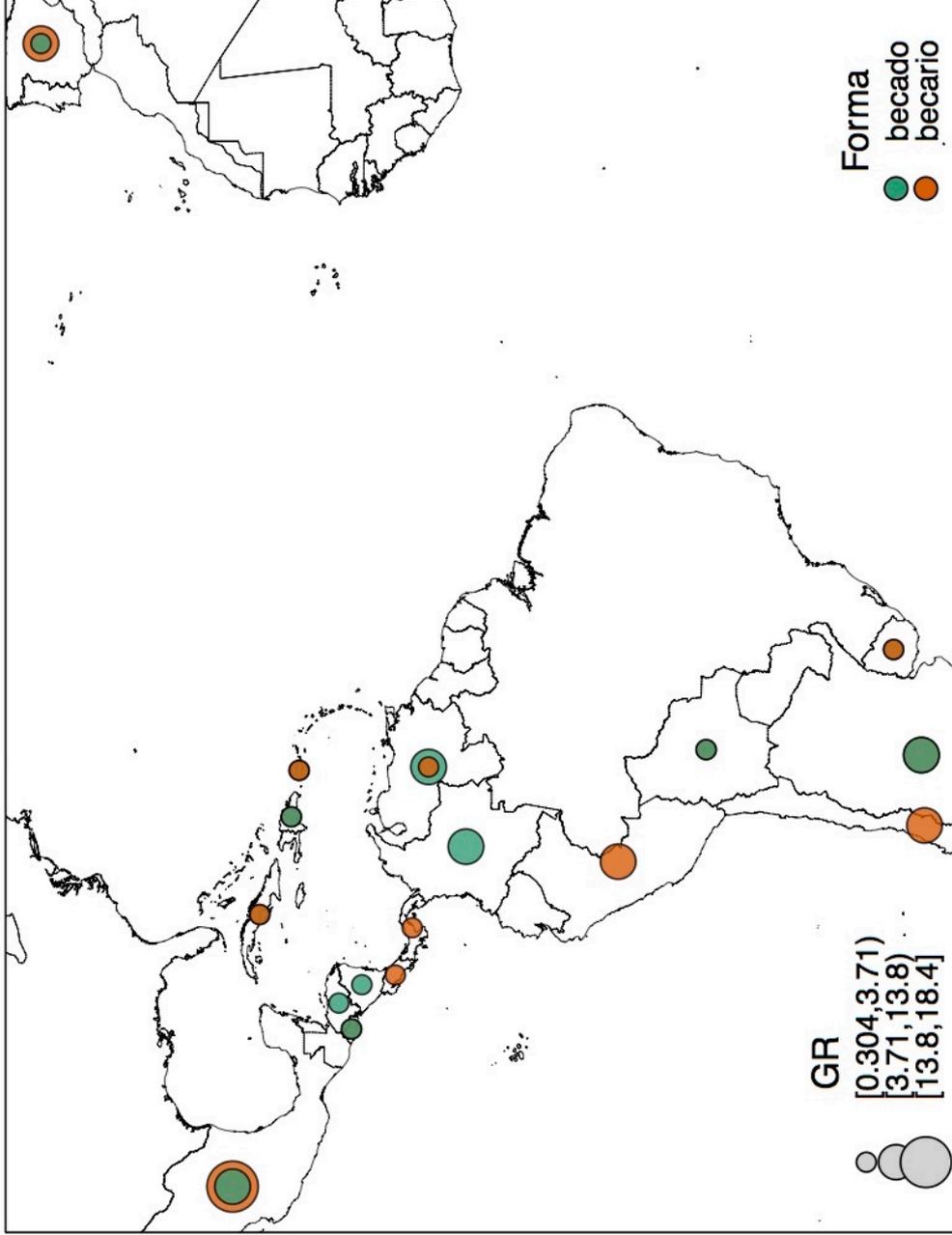
SAD



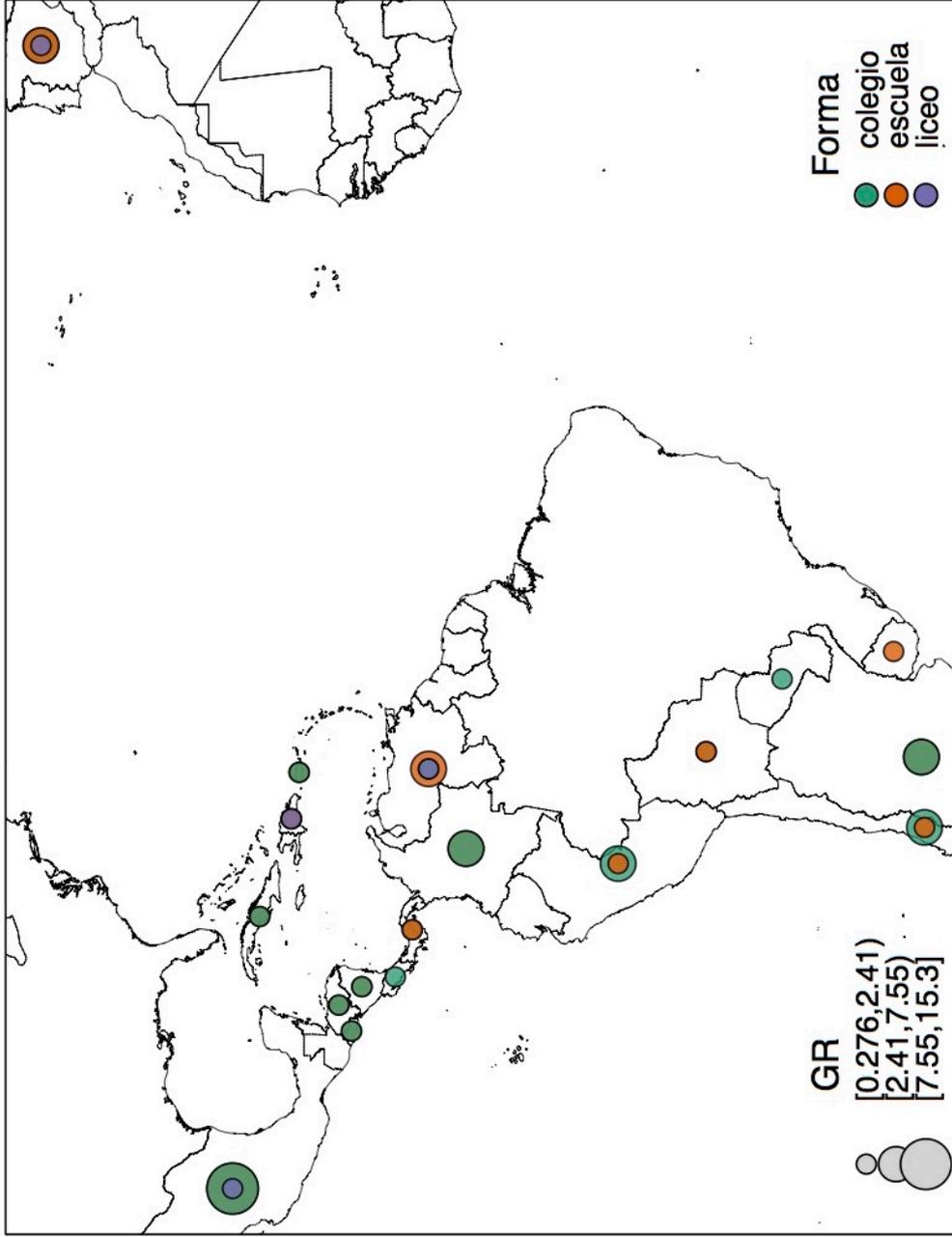
SANDWICH



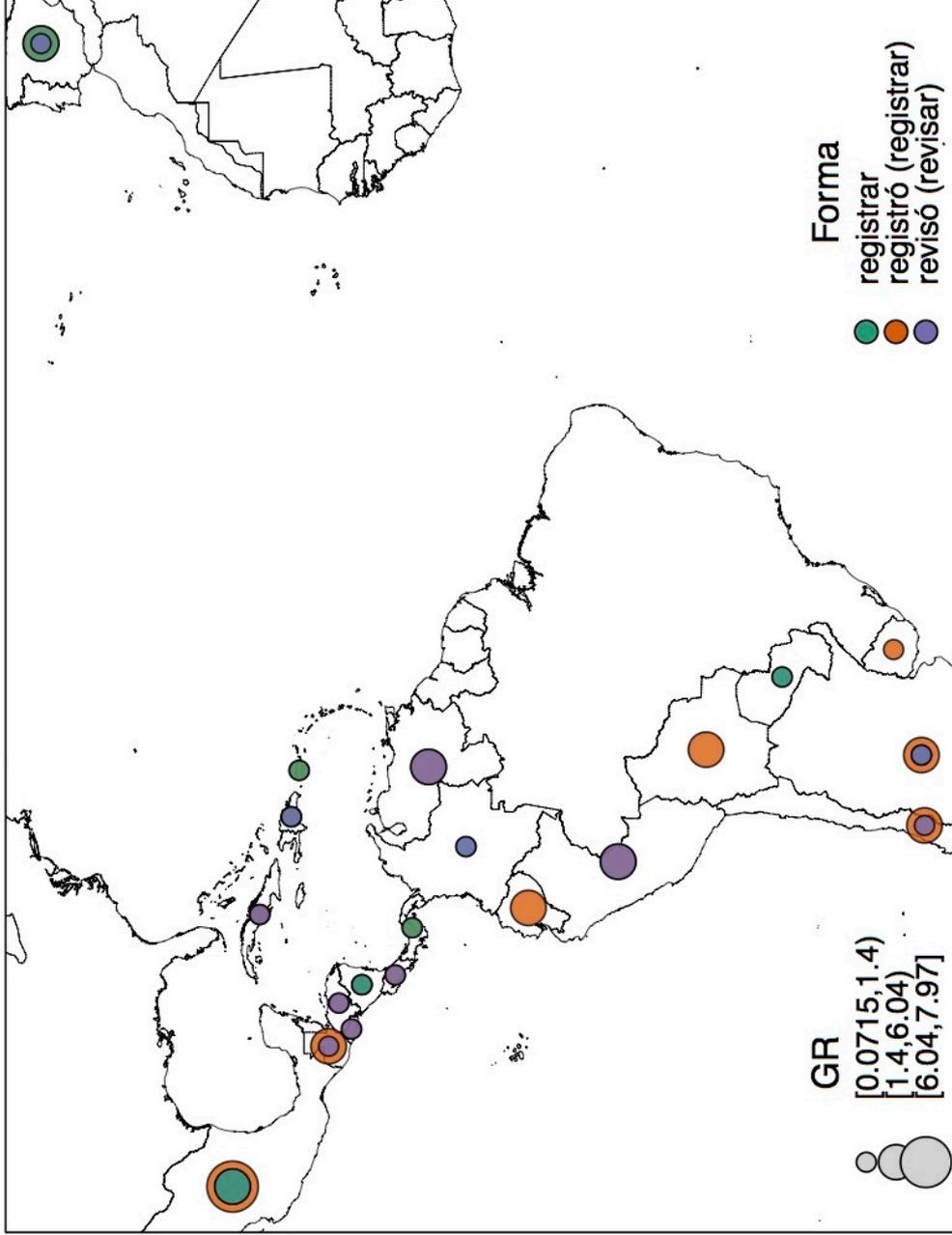
SCHOLARSHIP HOLDER



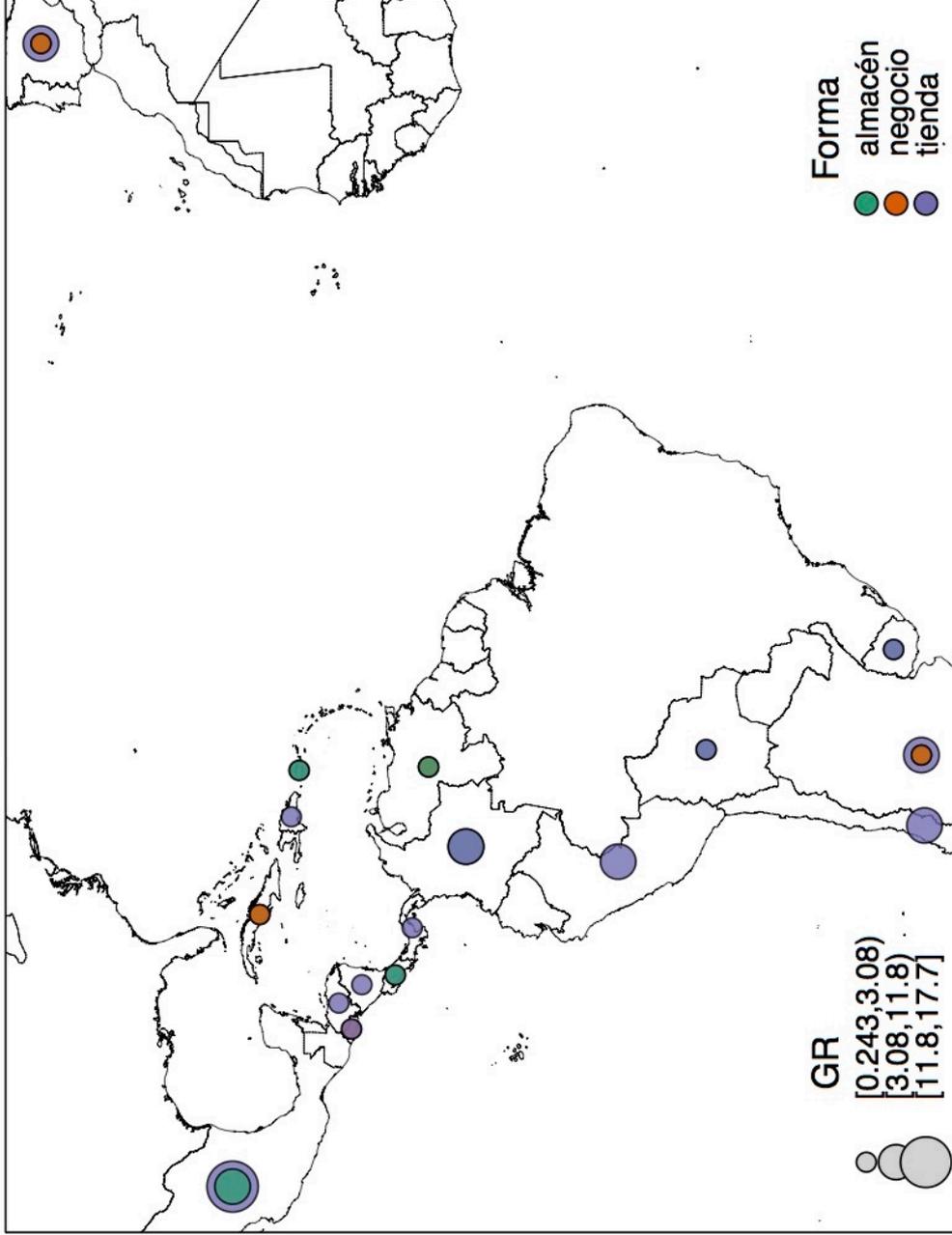
SCHOOL



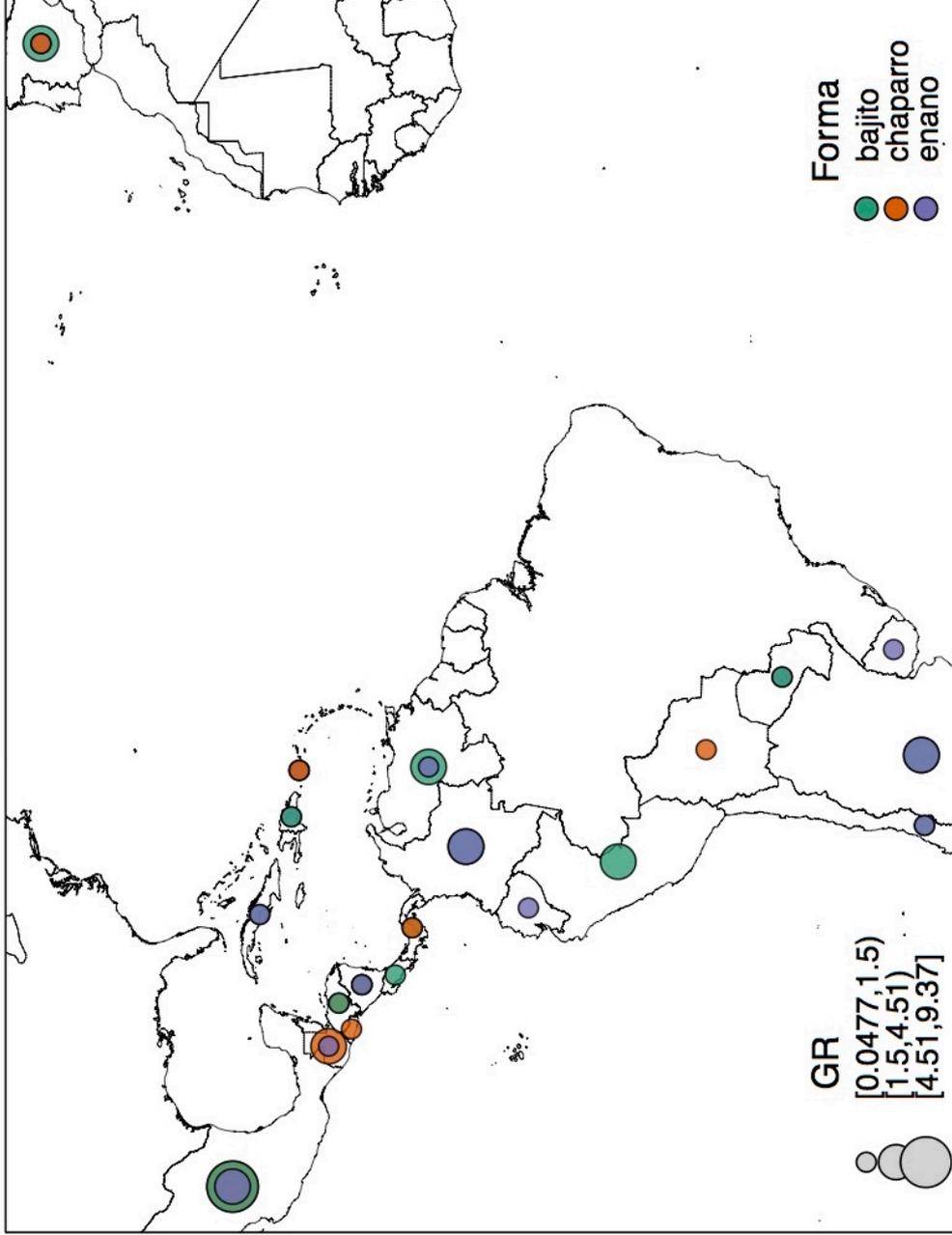
SEARCH



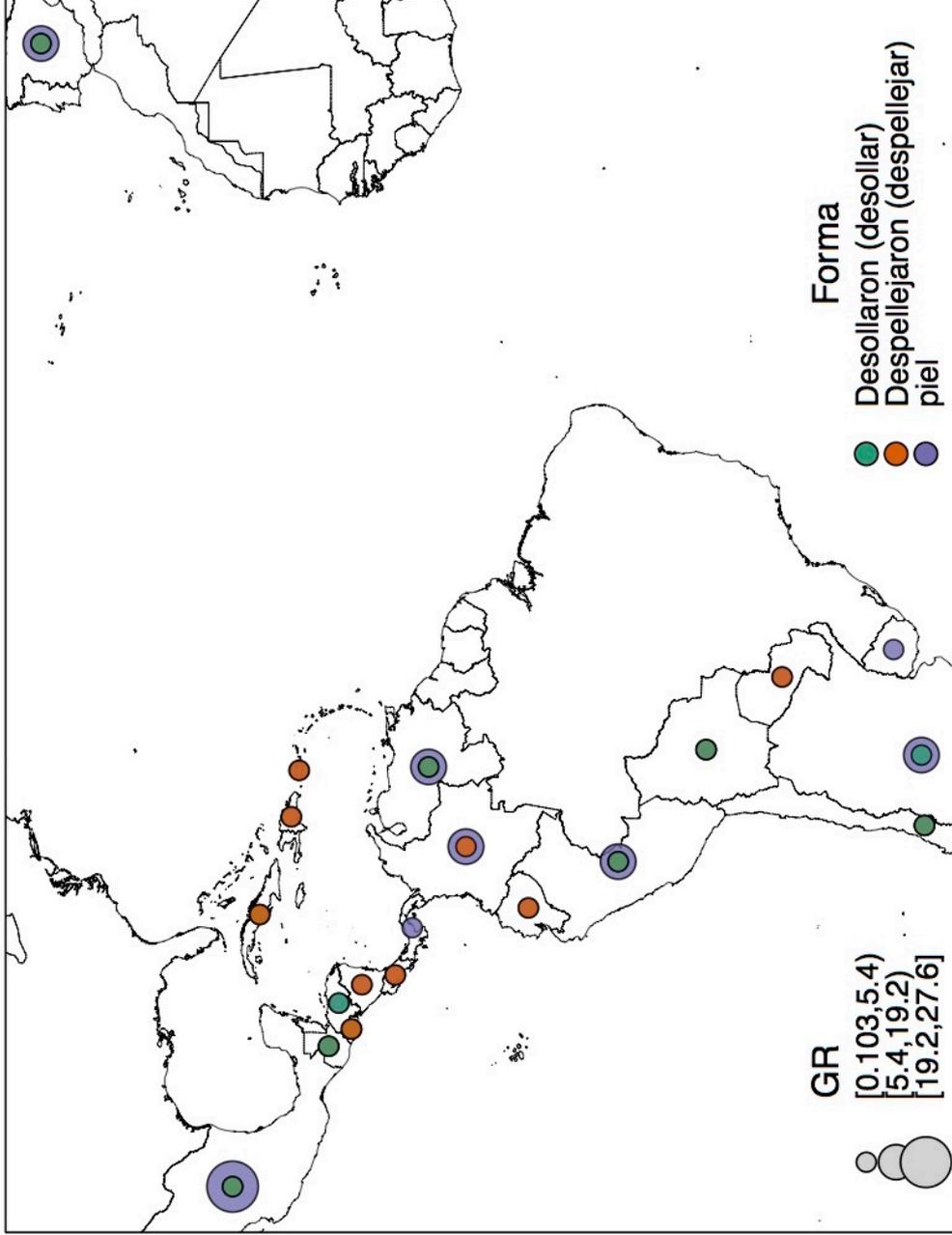
SHOP (US: STORE)



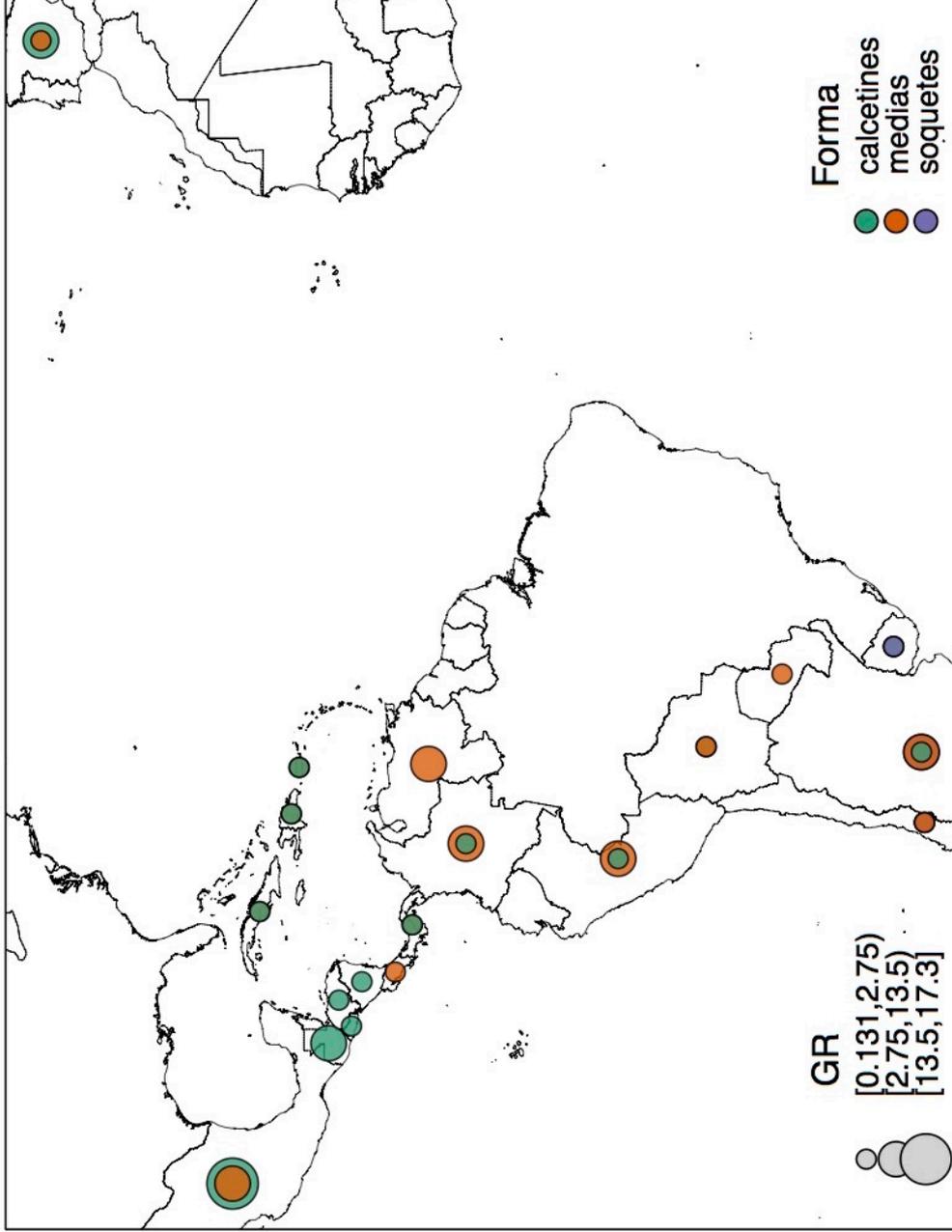
SHORT (PERSON)



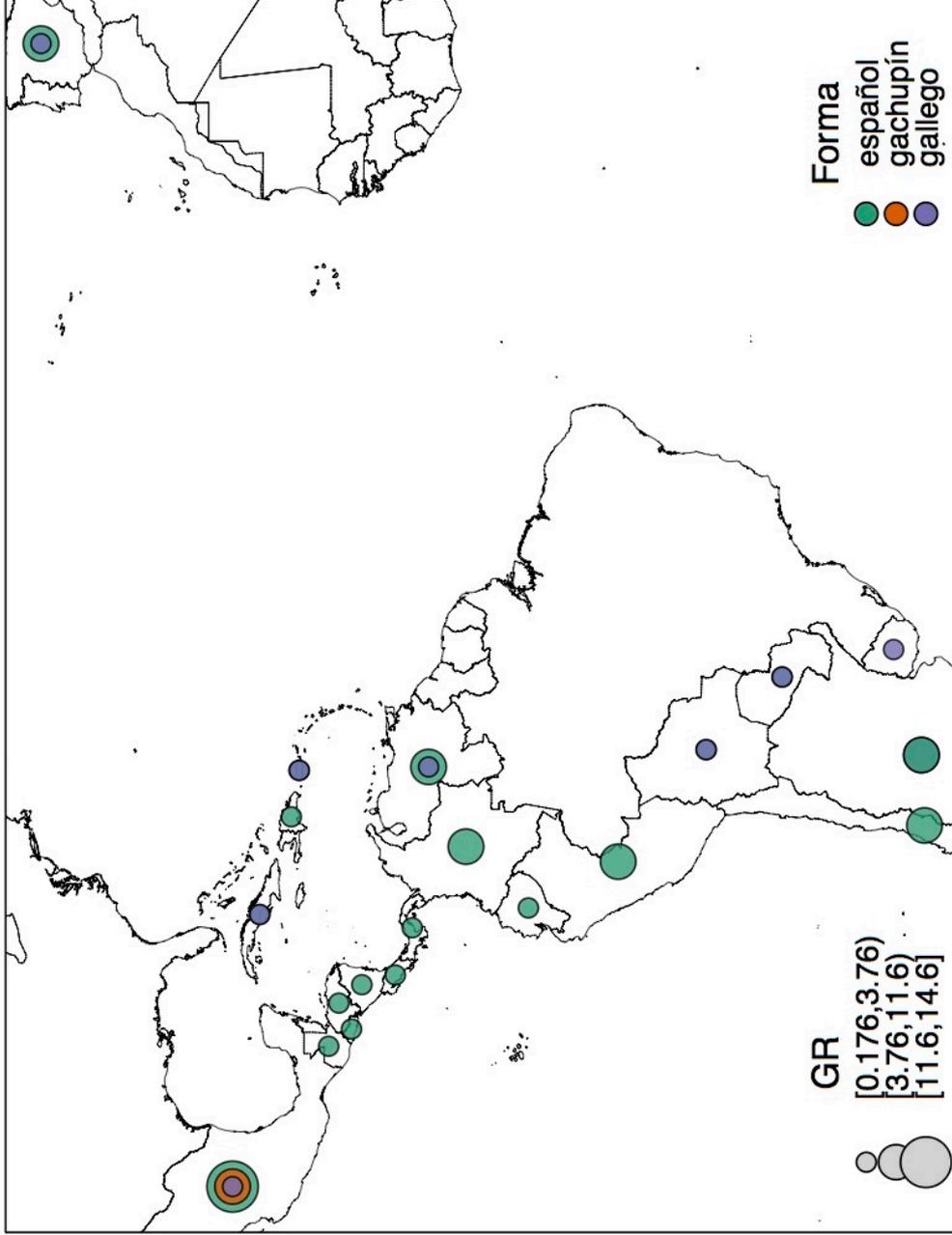
SKIN



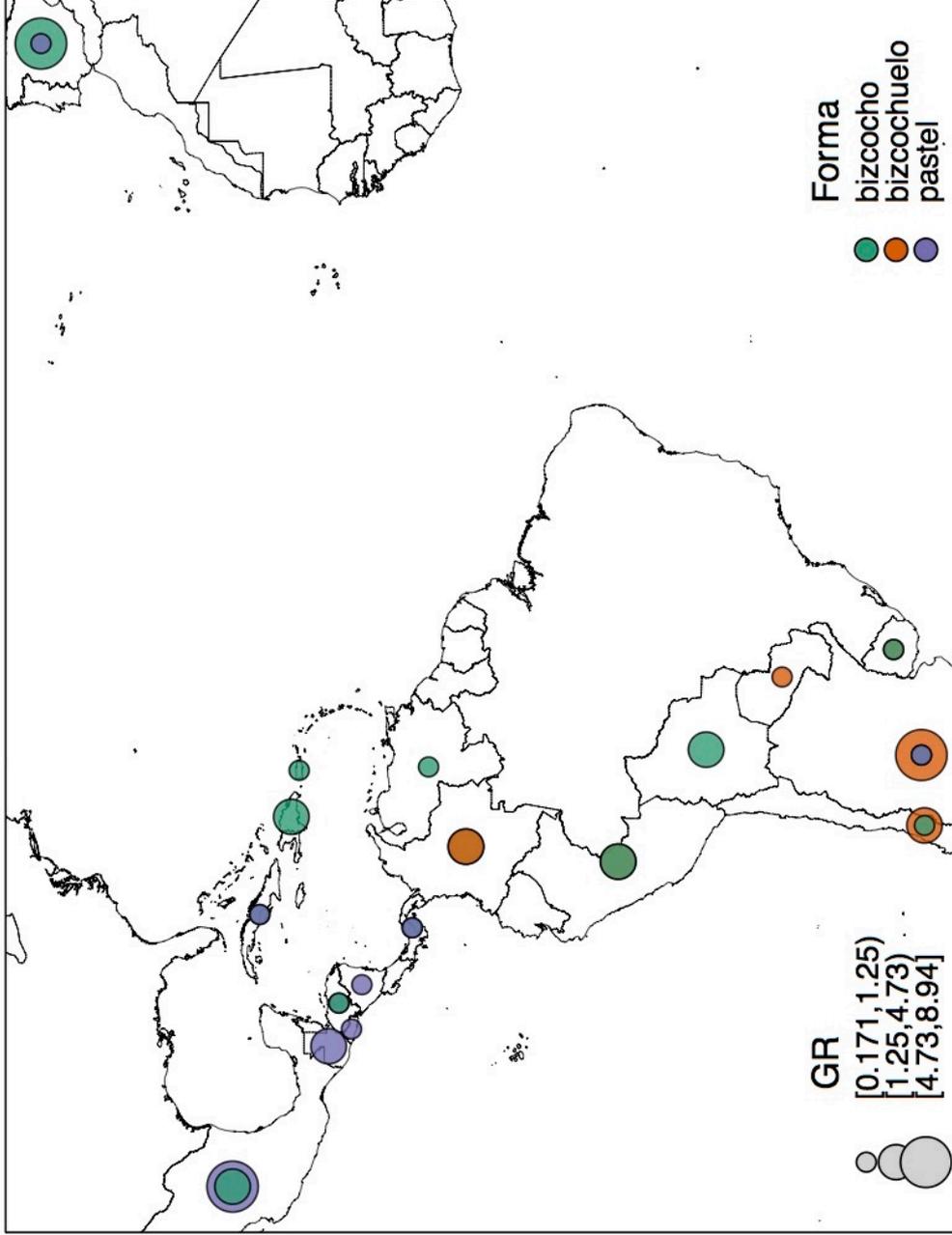
SOCKS



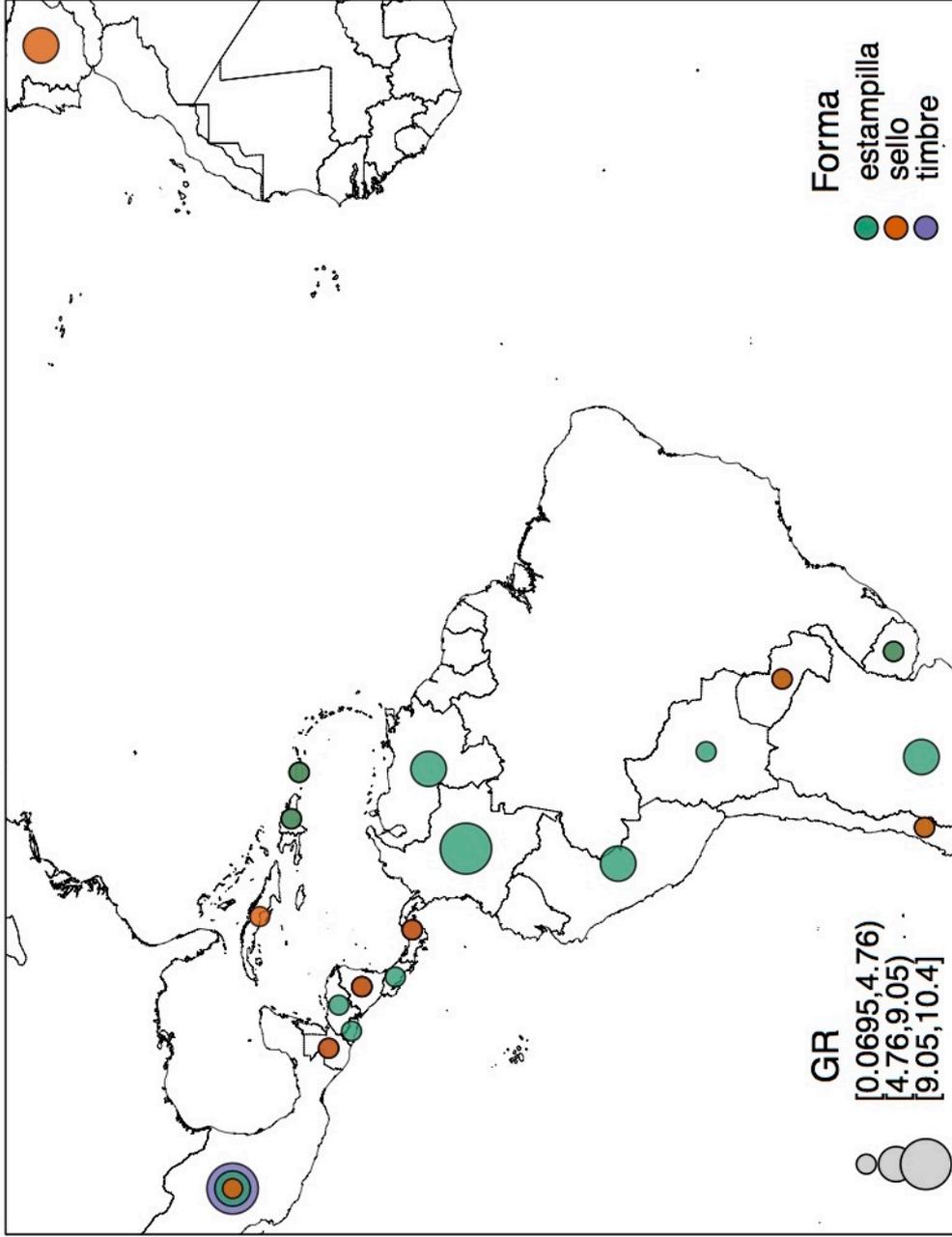
SPANIARD



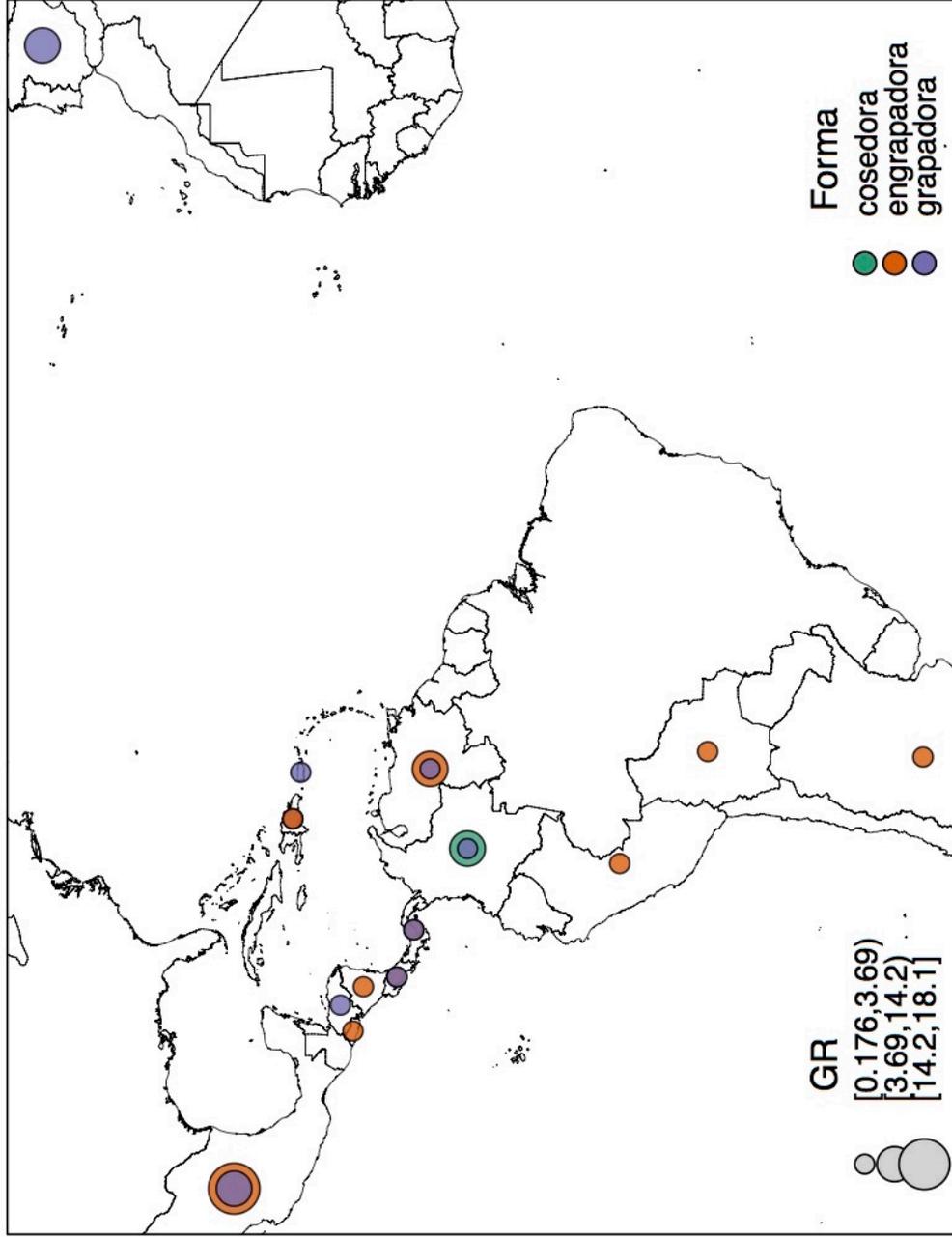
SPONGE CAKE



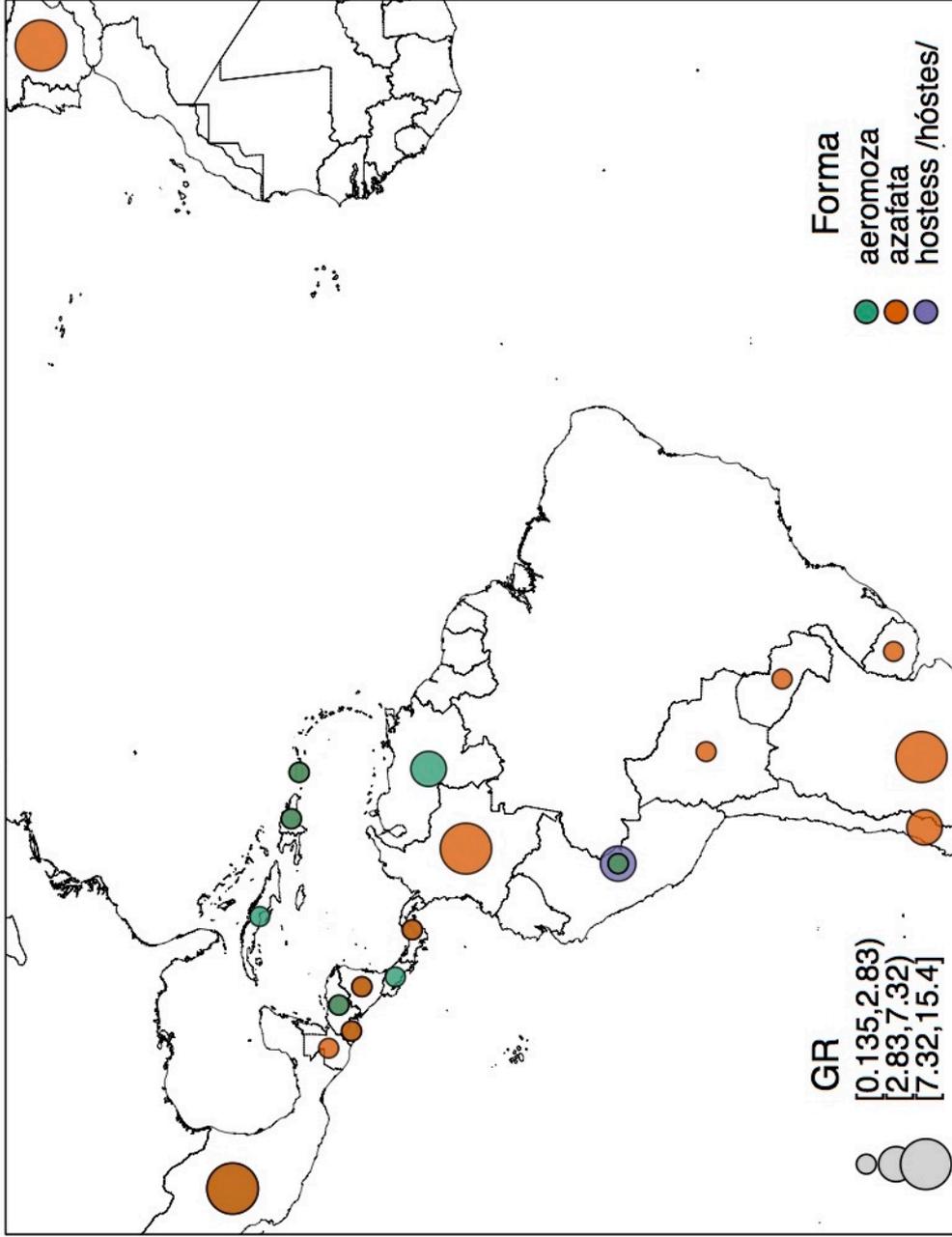
STAMP



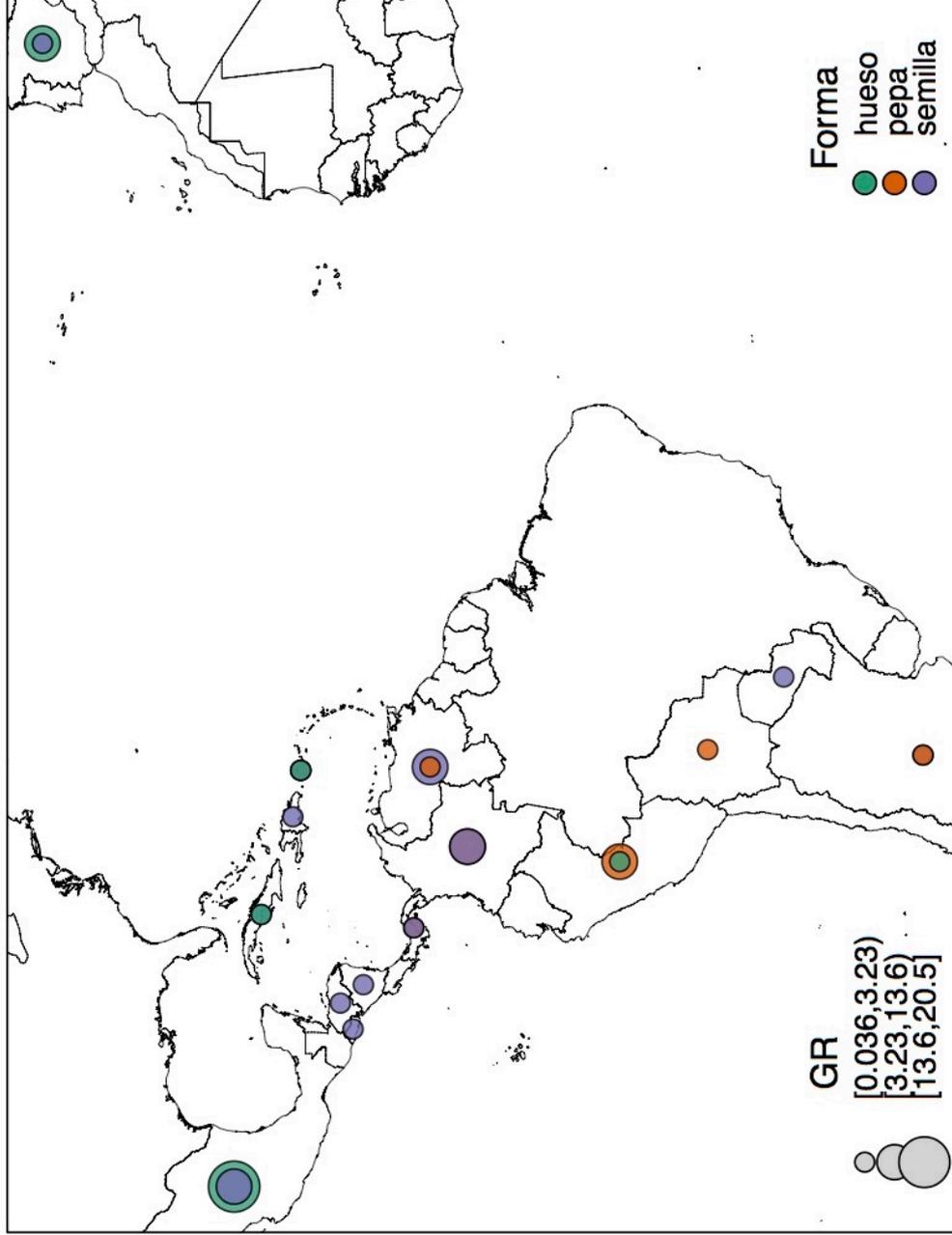
STAPLER



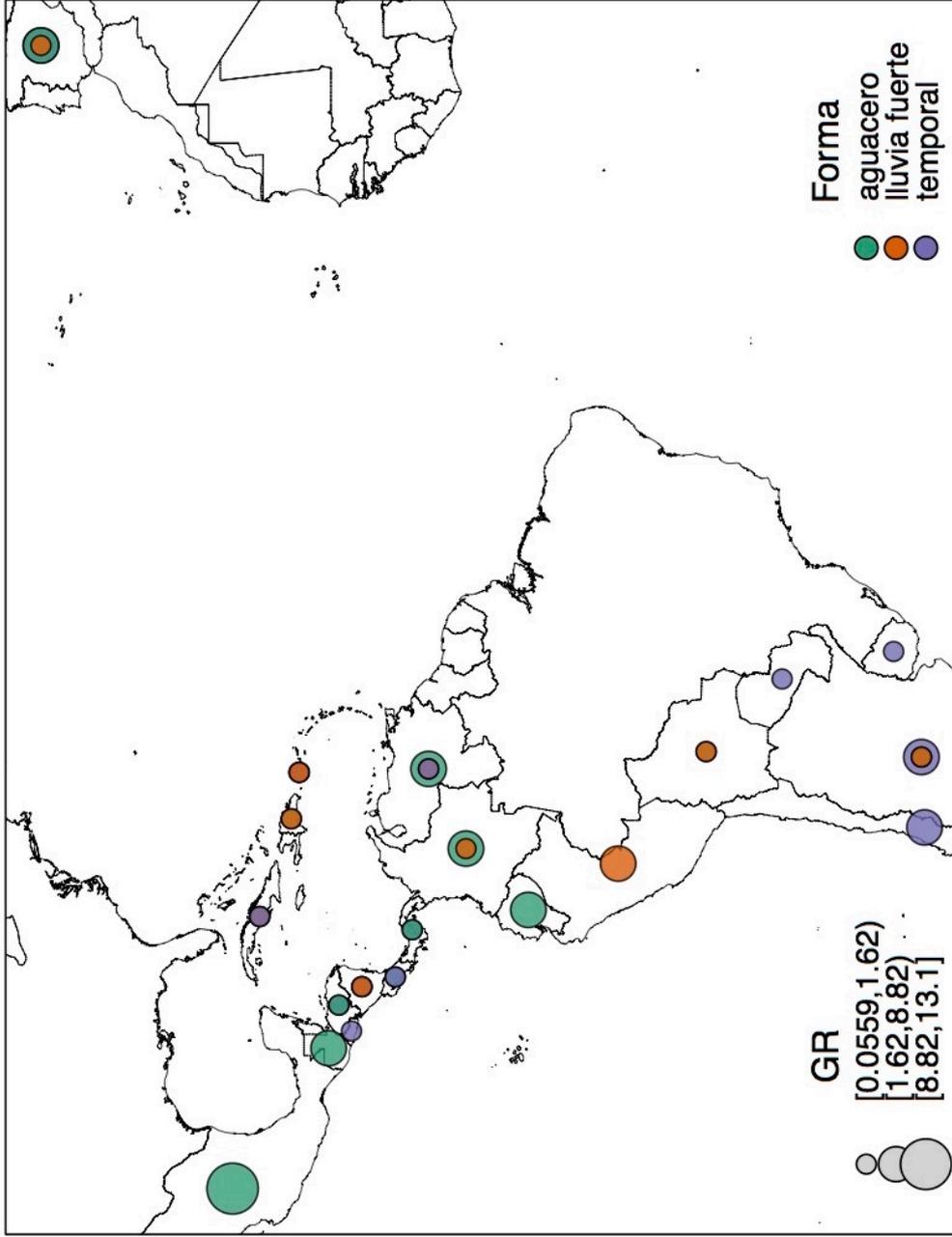
STEWARDESS



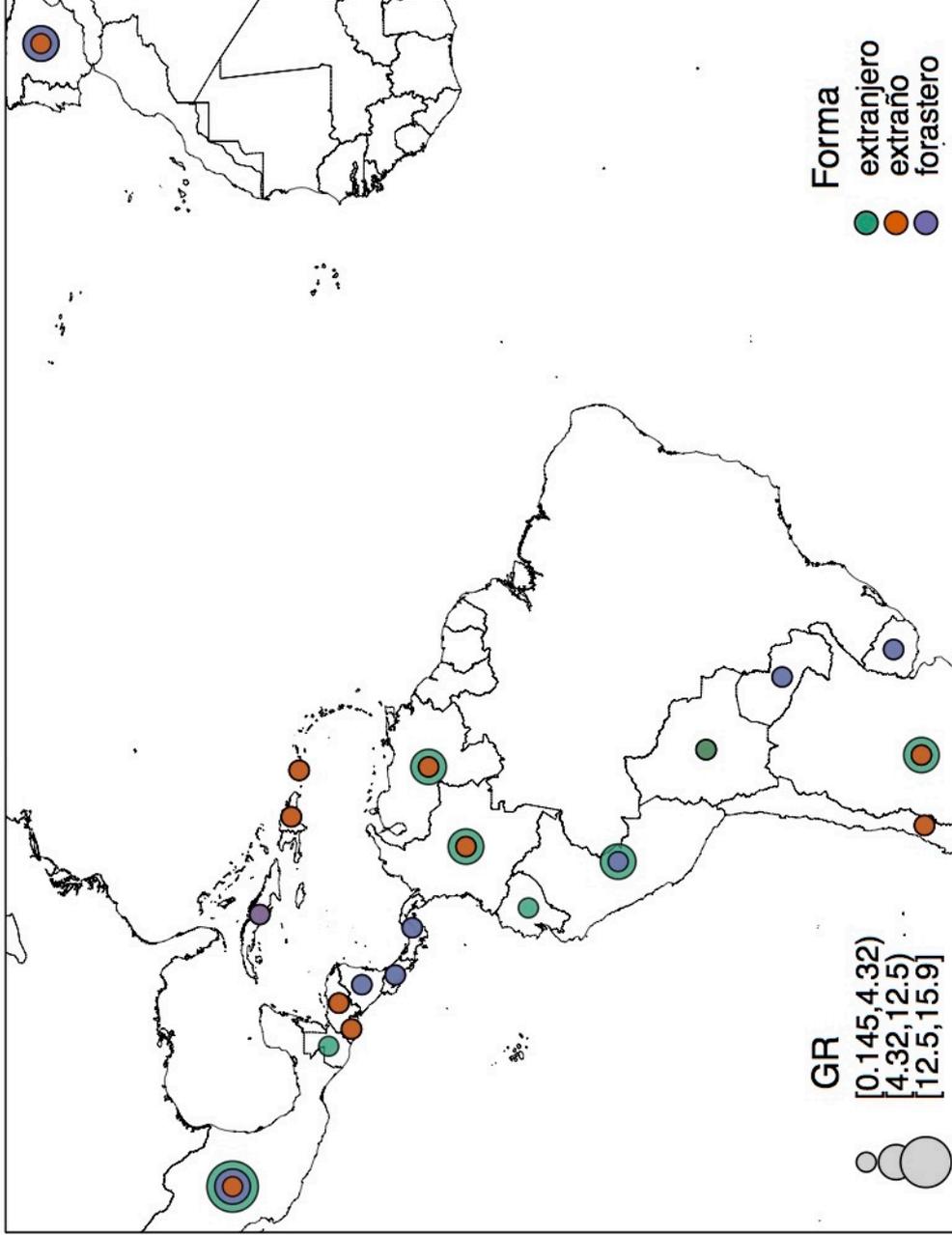
STONE (US: PIT)



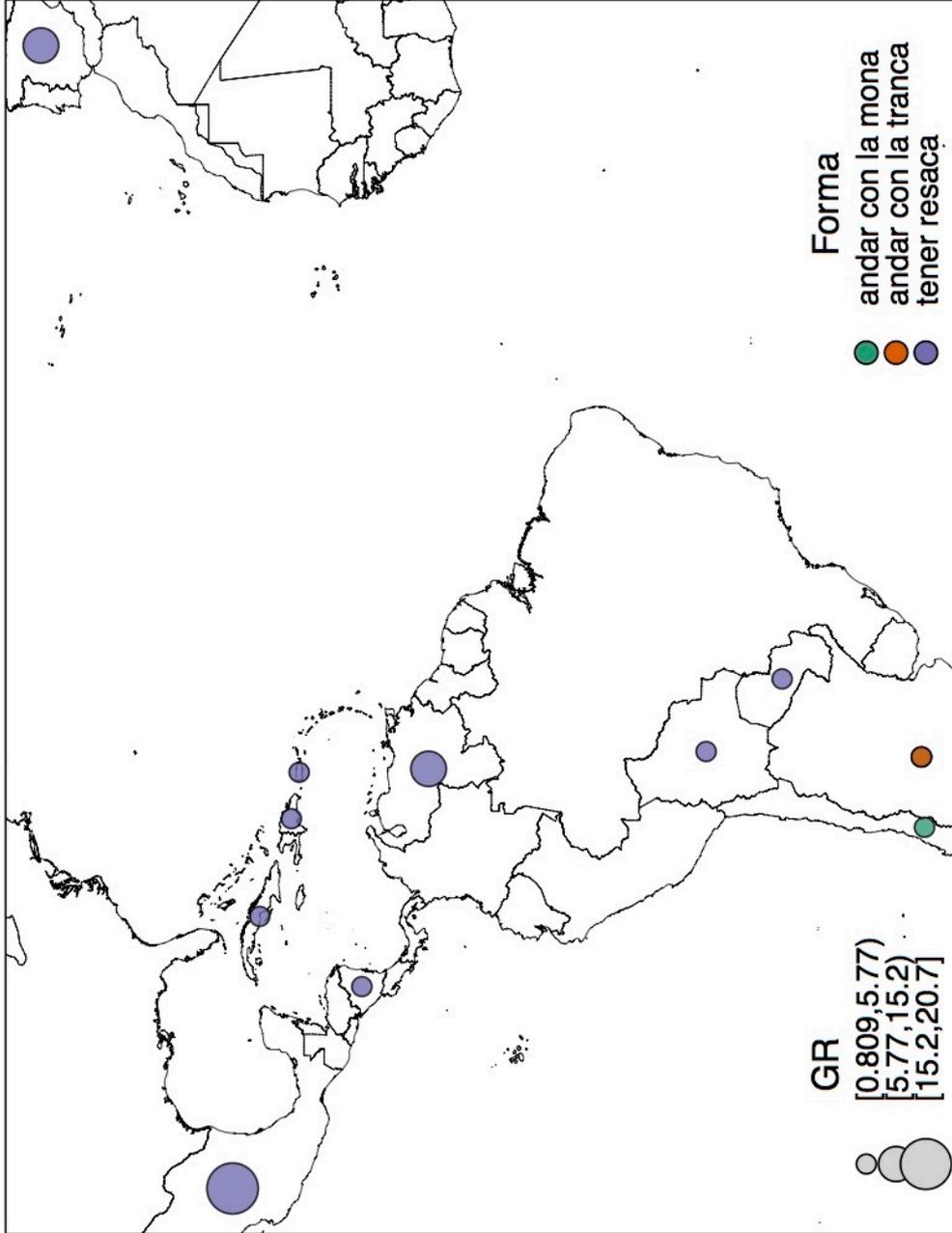
STORM



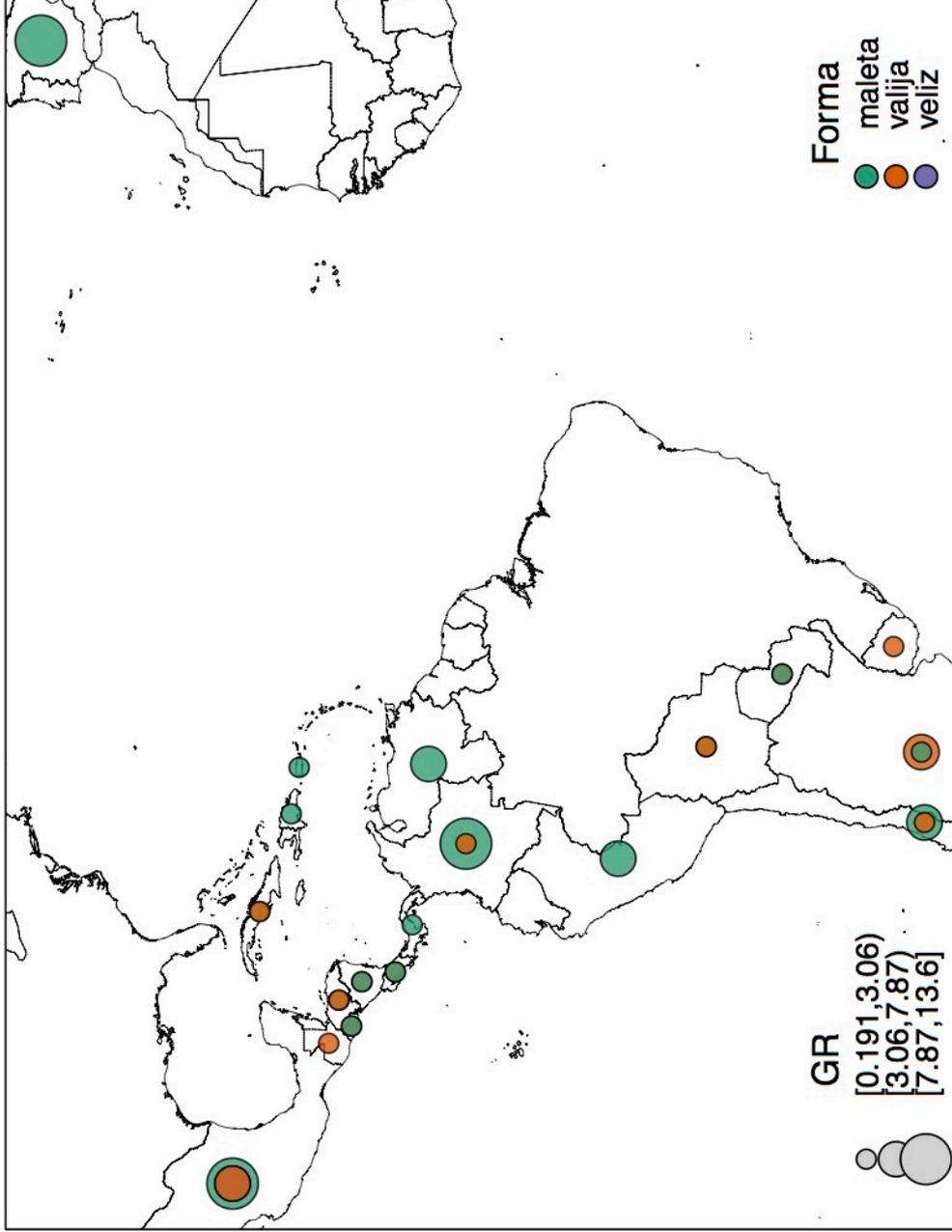
STRANGER



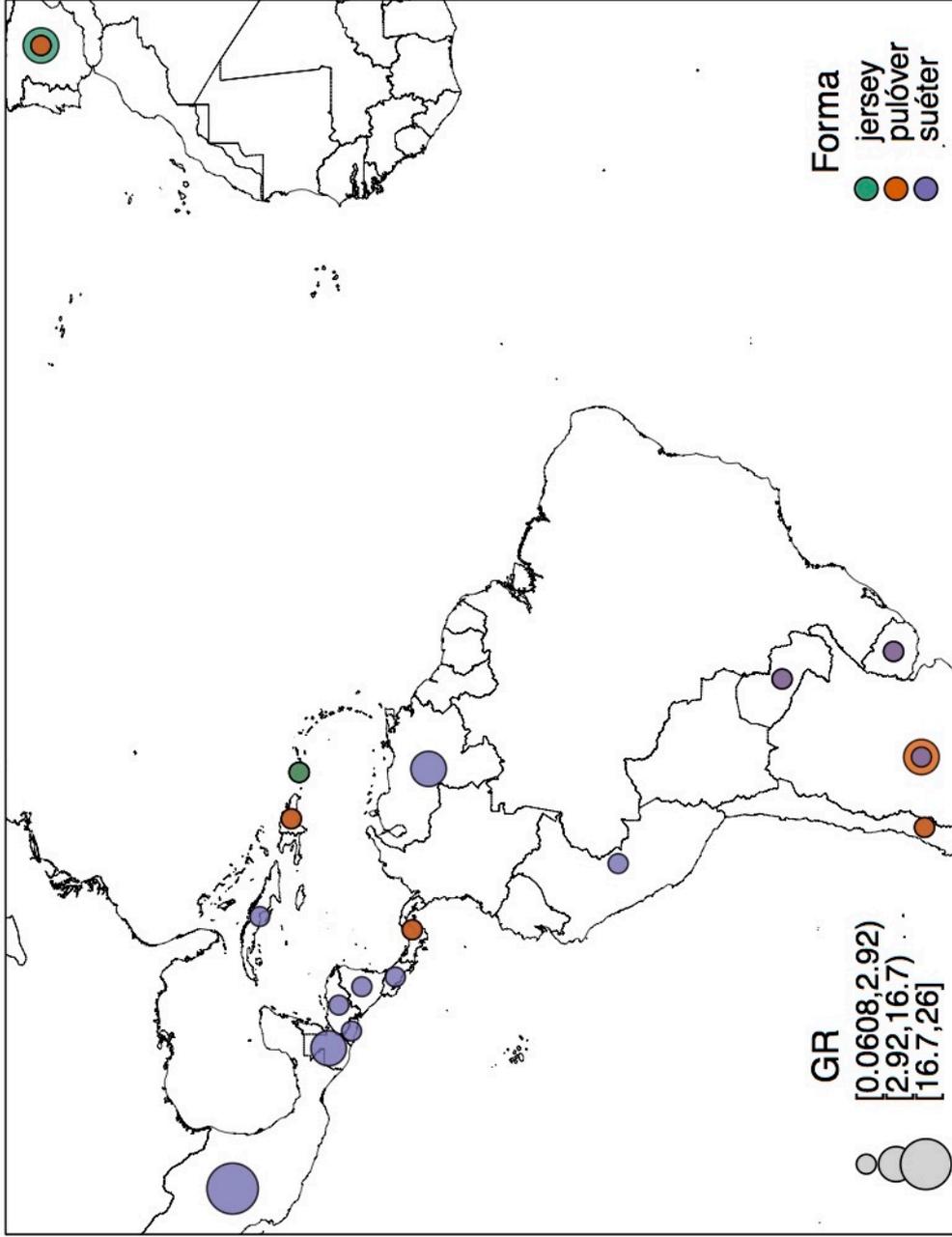
SUFFER FROM A HANGOVER



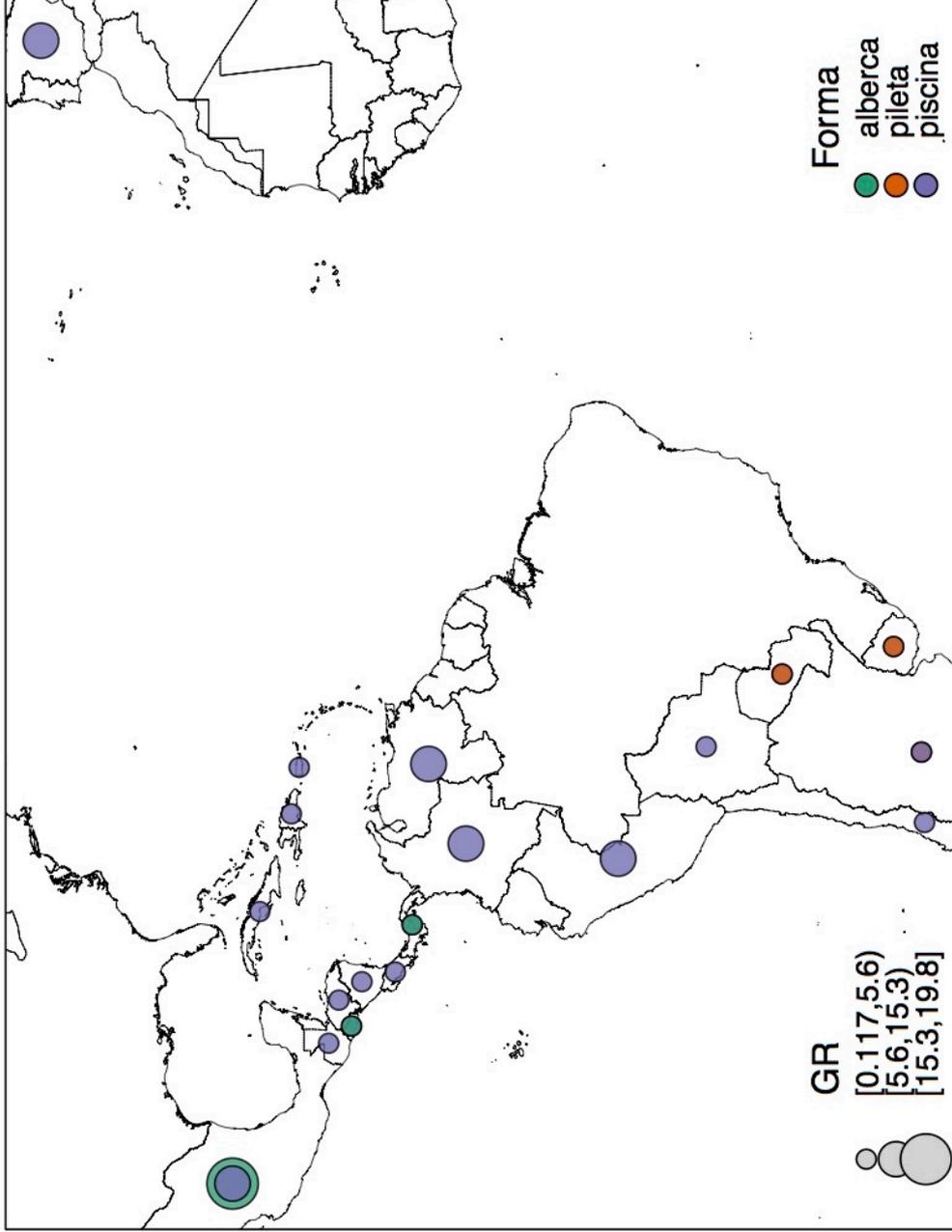
SUITCASE



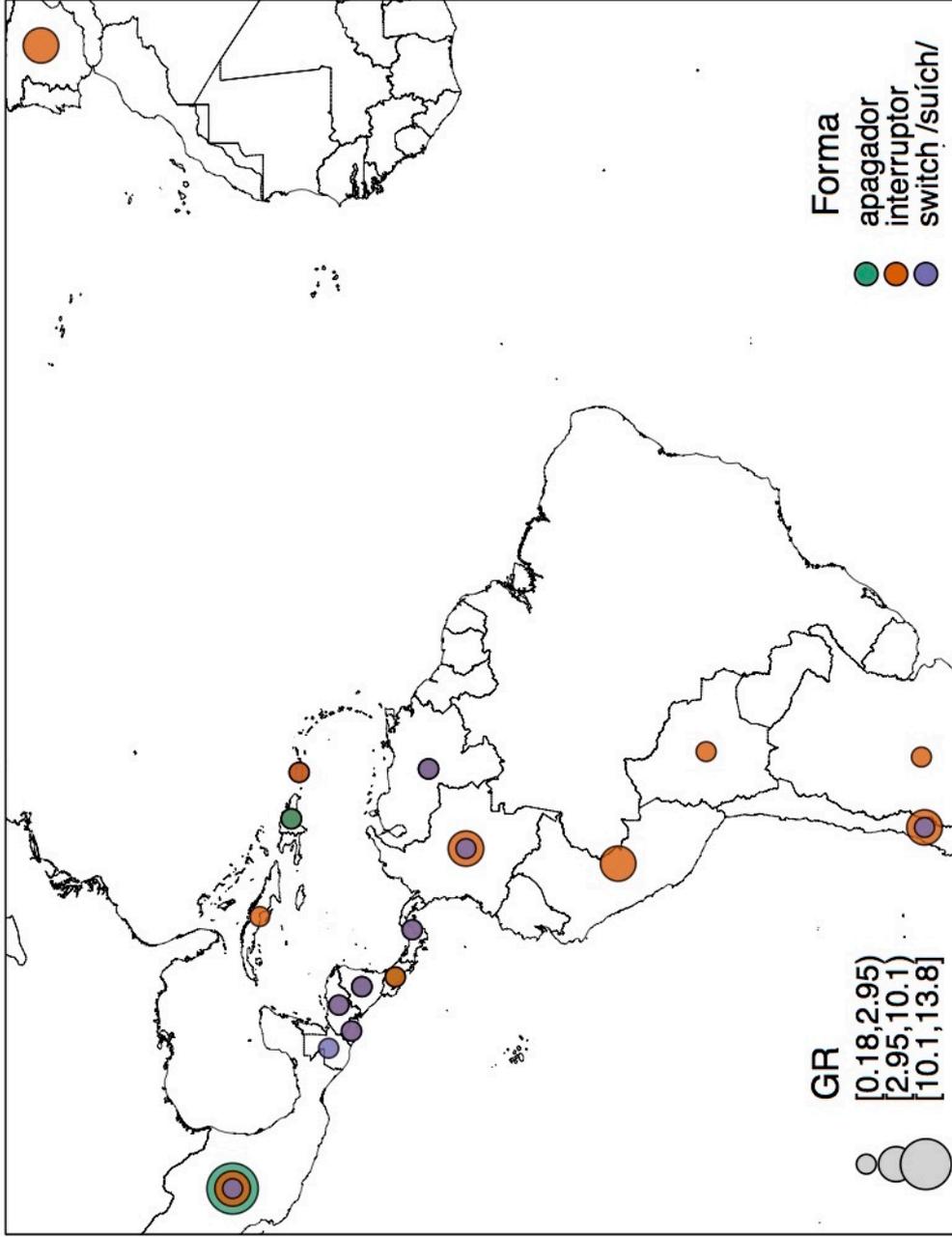
SWEATER



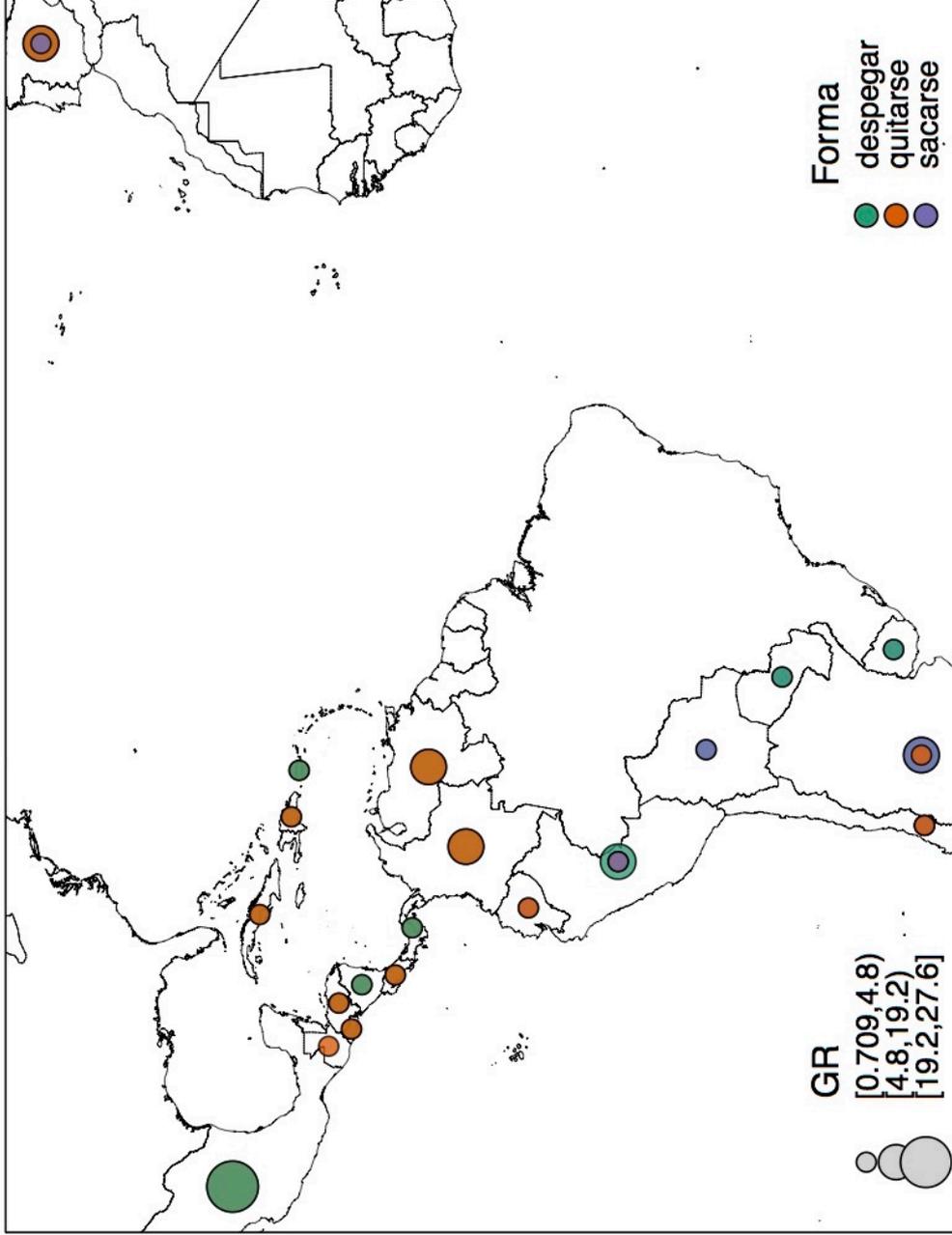
SWIMMING POOL



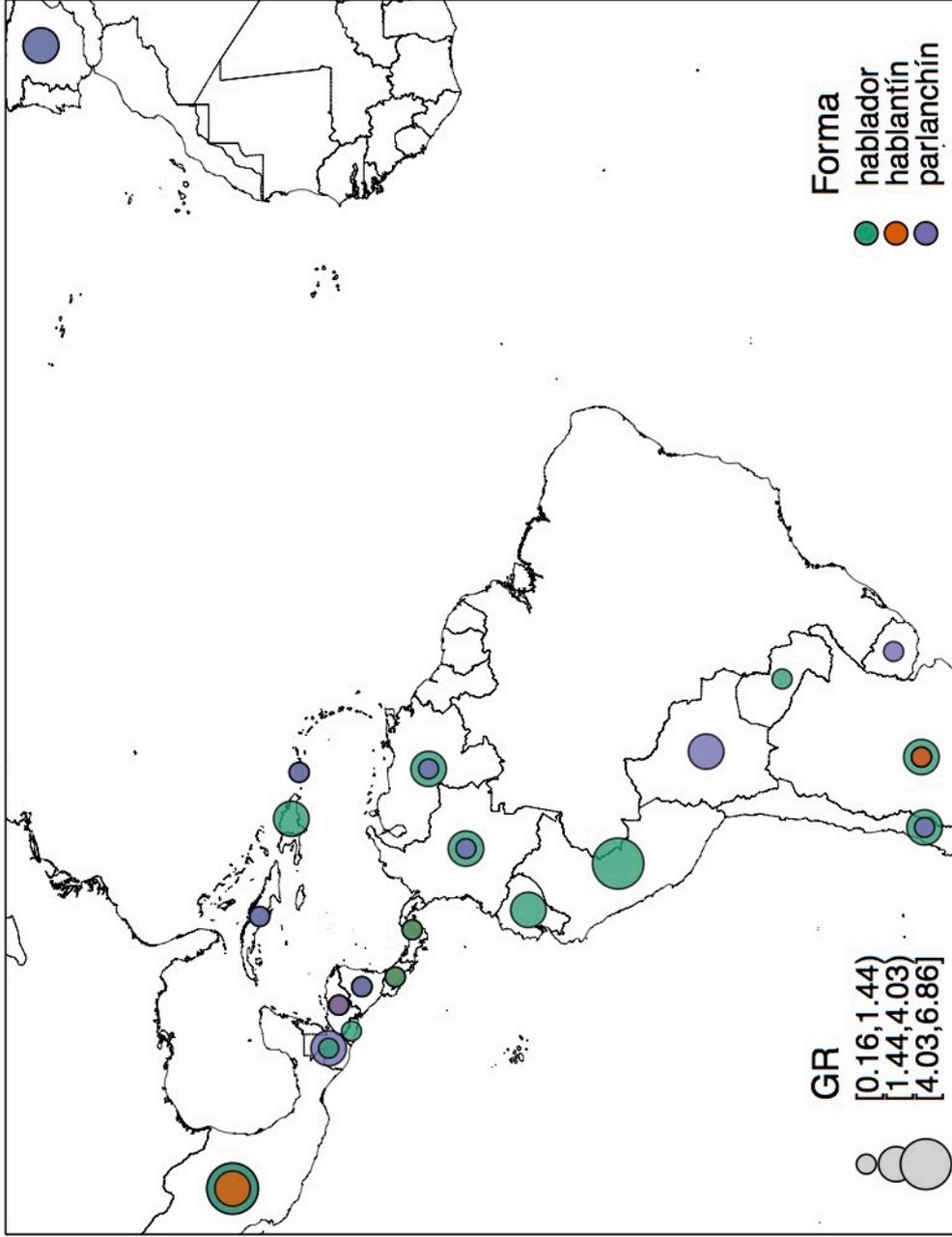
SWITCH



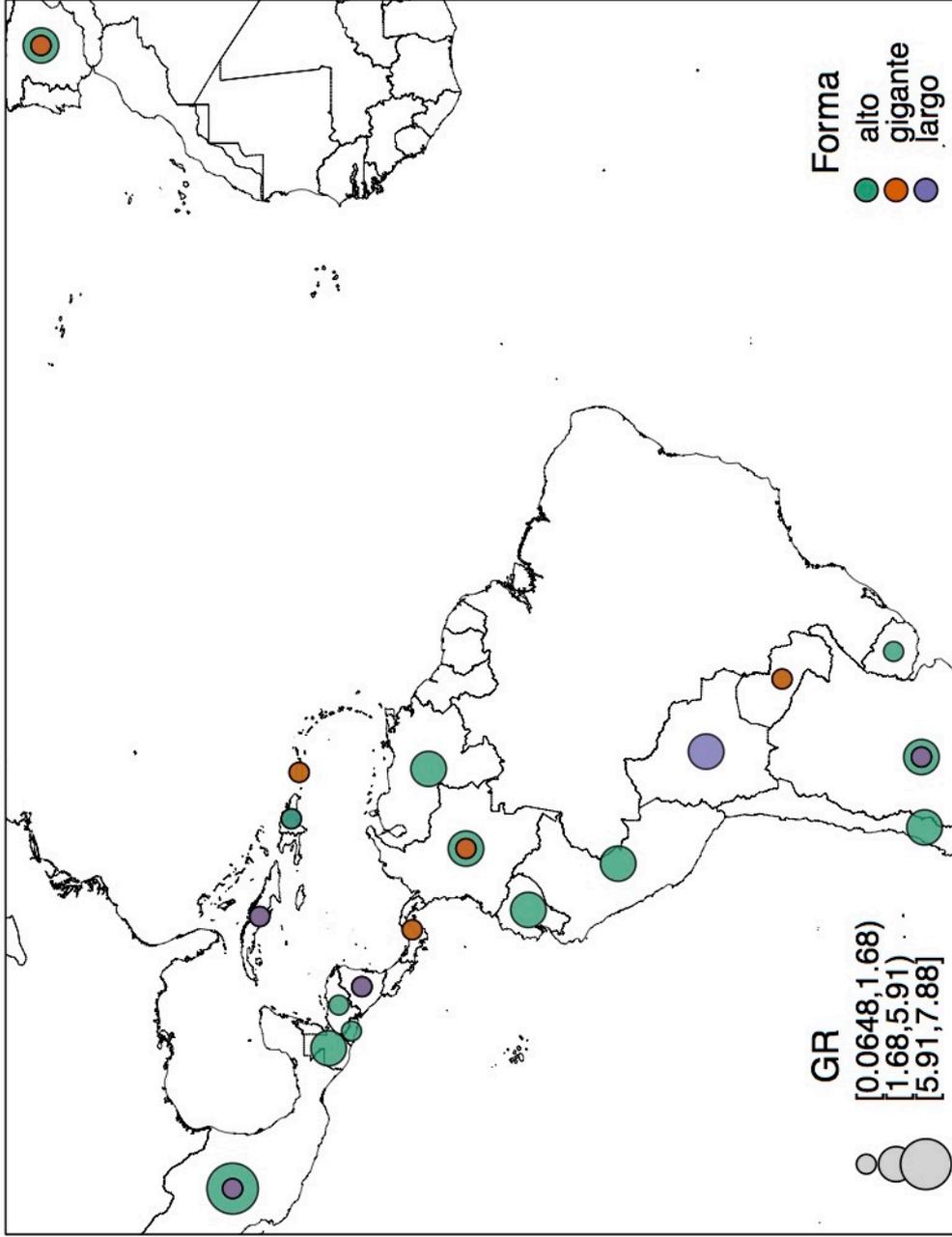
TAKE OFF



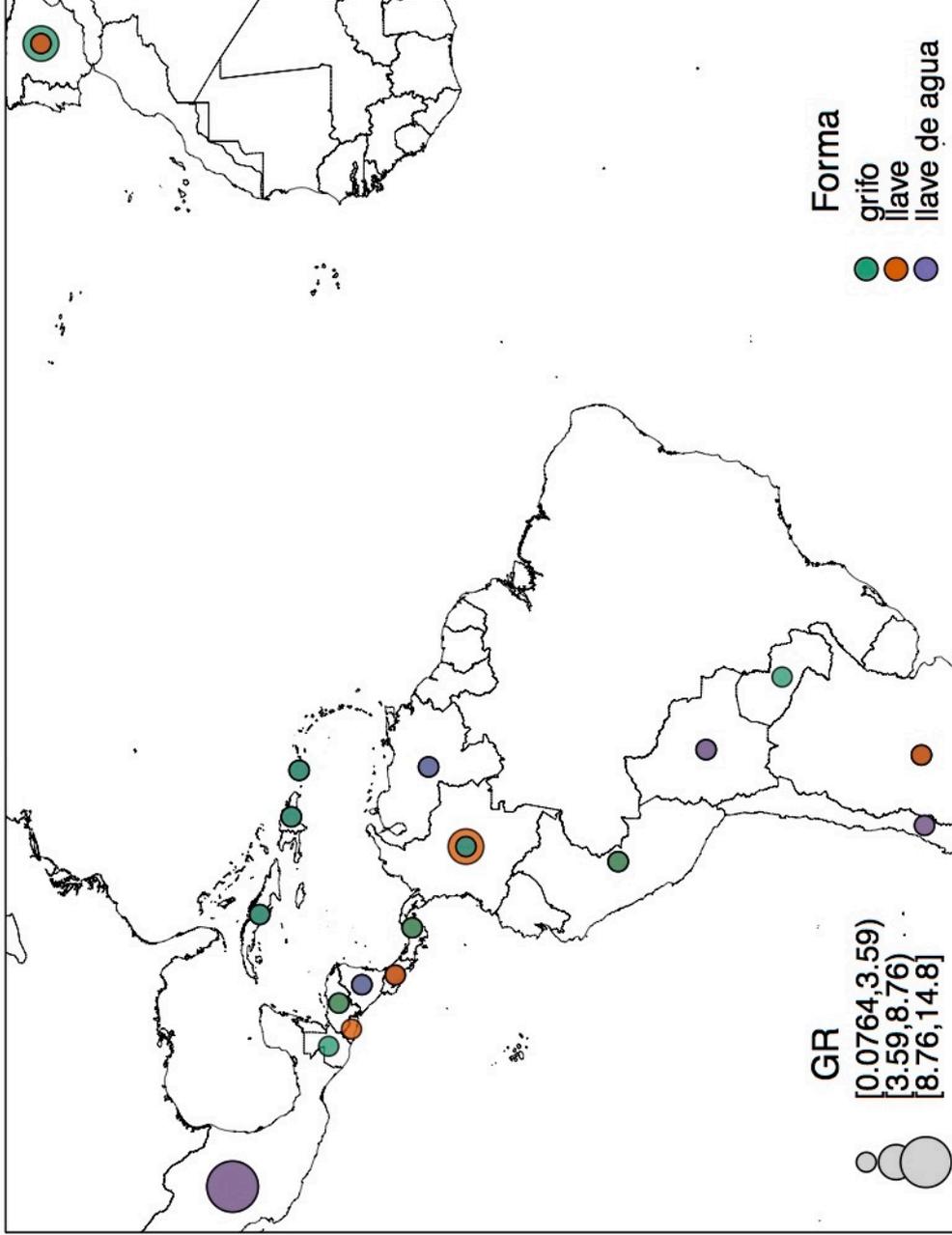
TALKATIVE



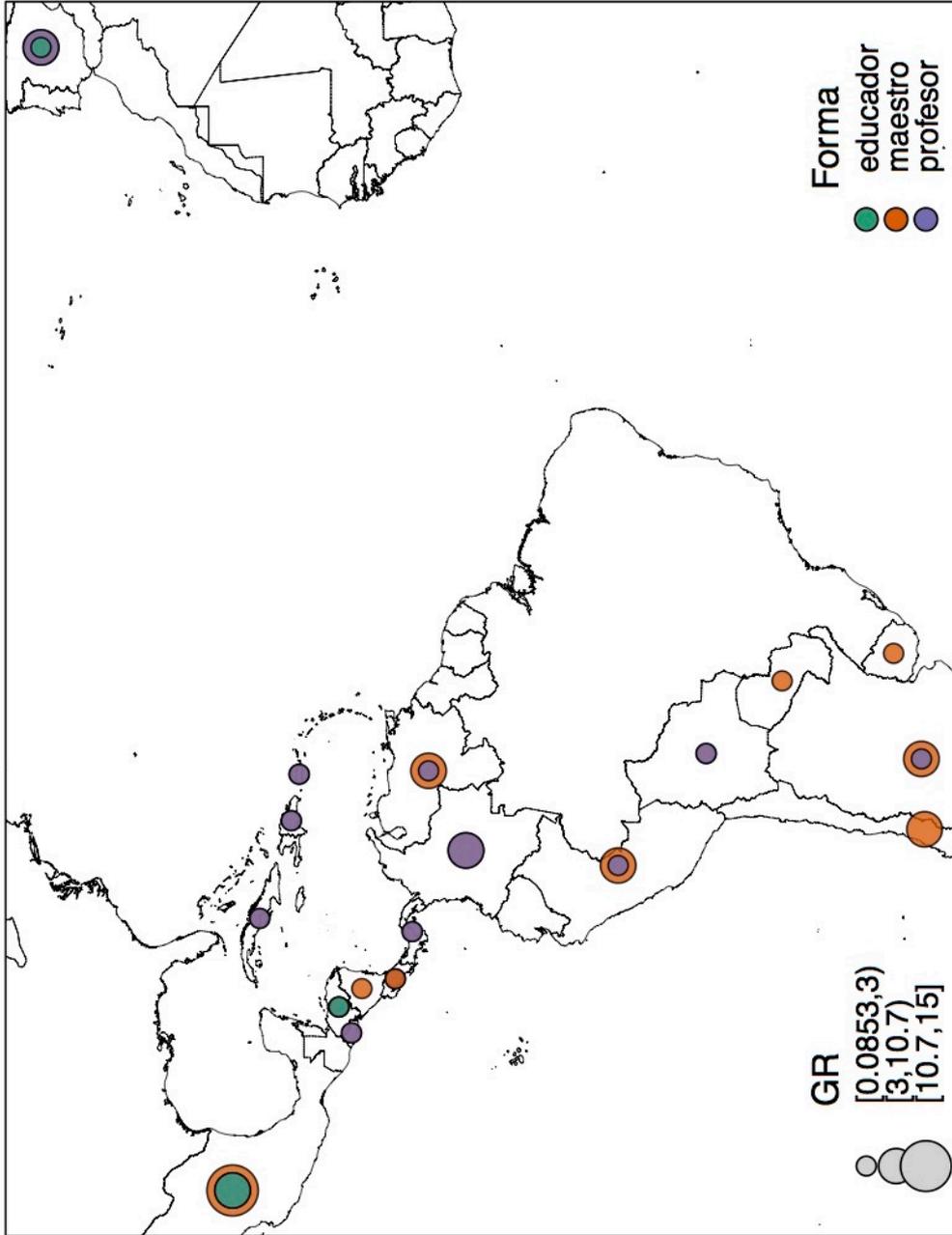
TALL (PERSON)



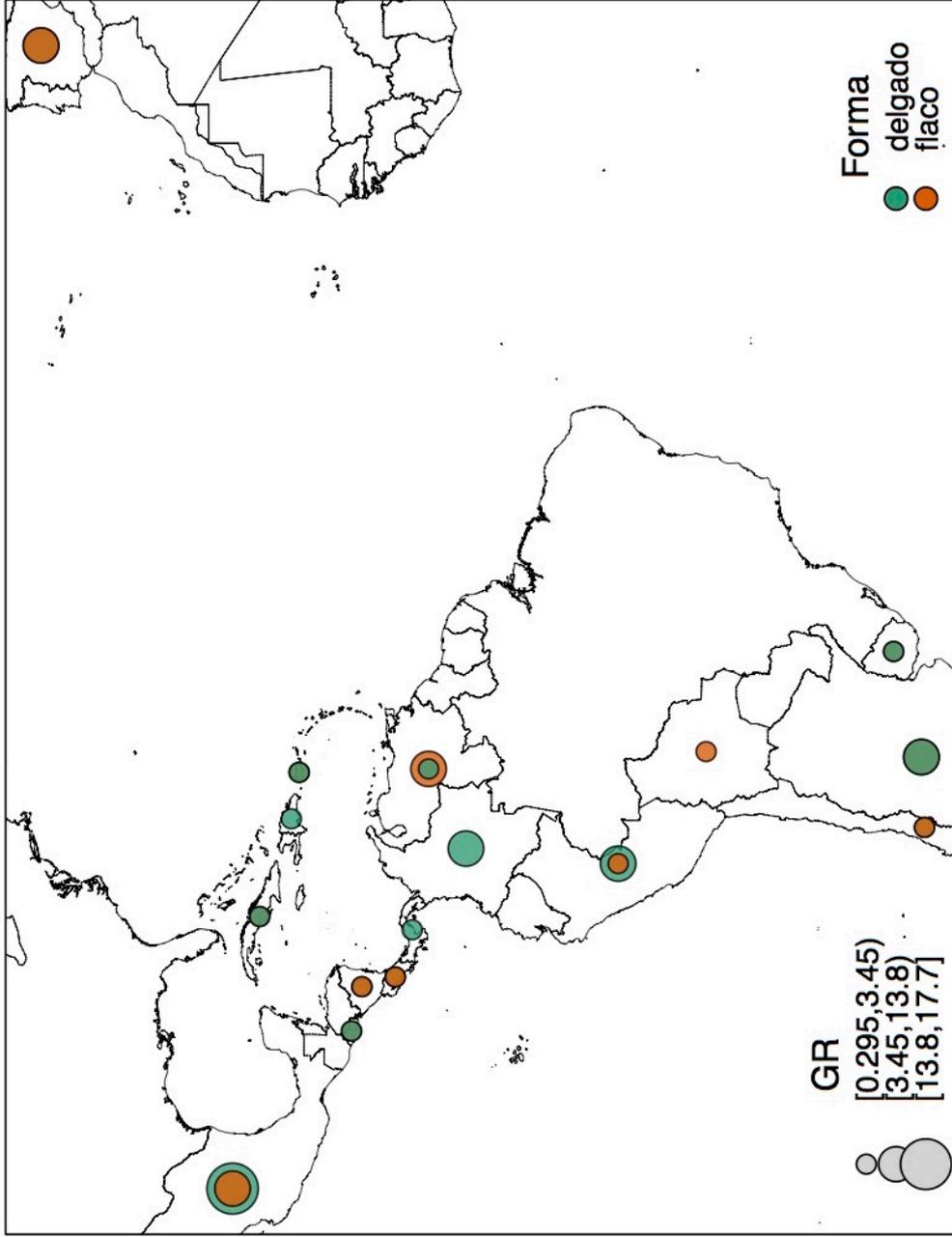
TAP (US: FAUCET)



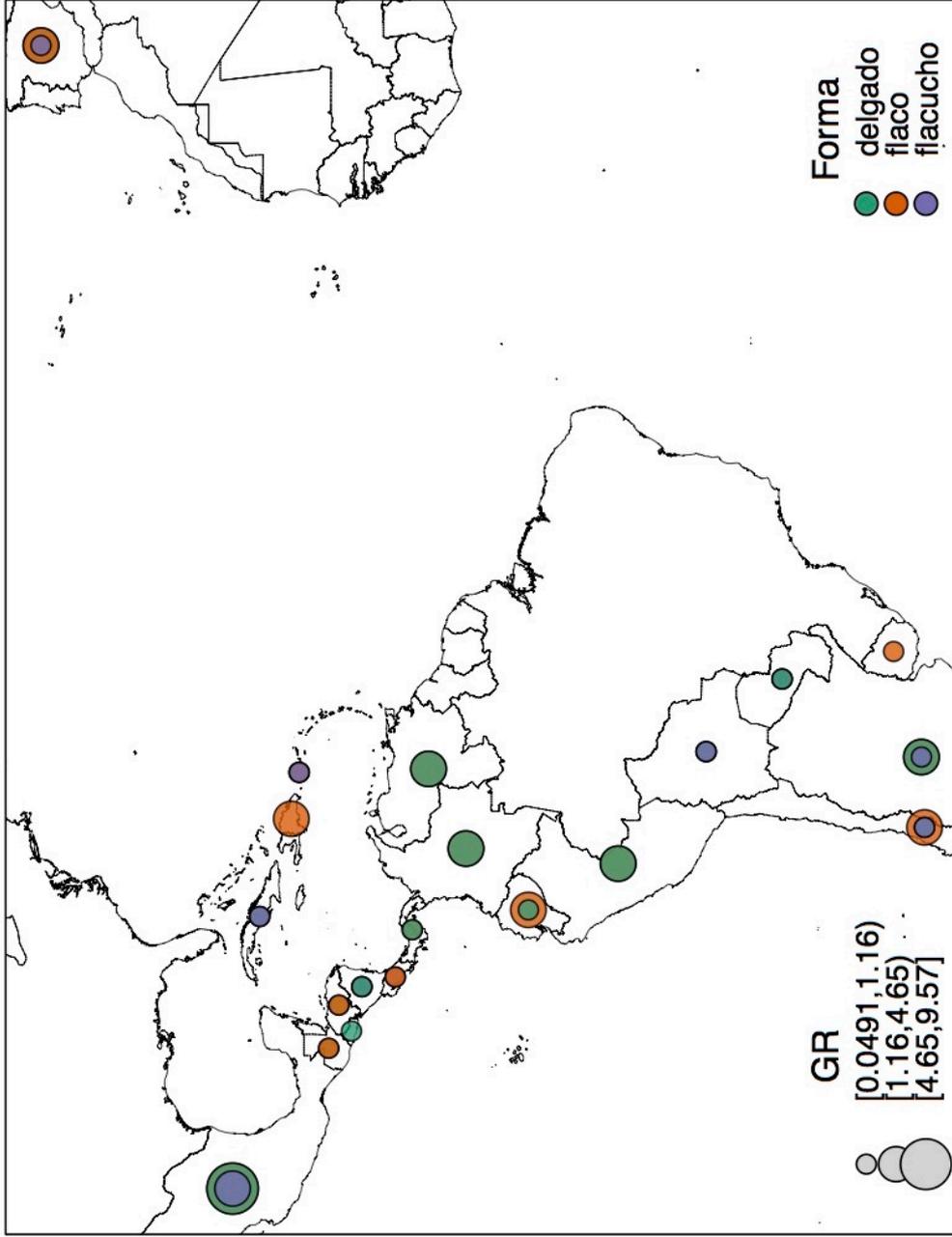
TEACHER



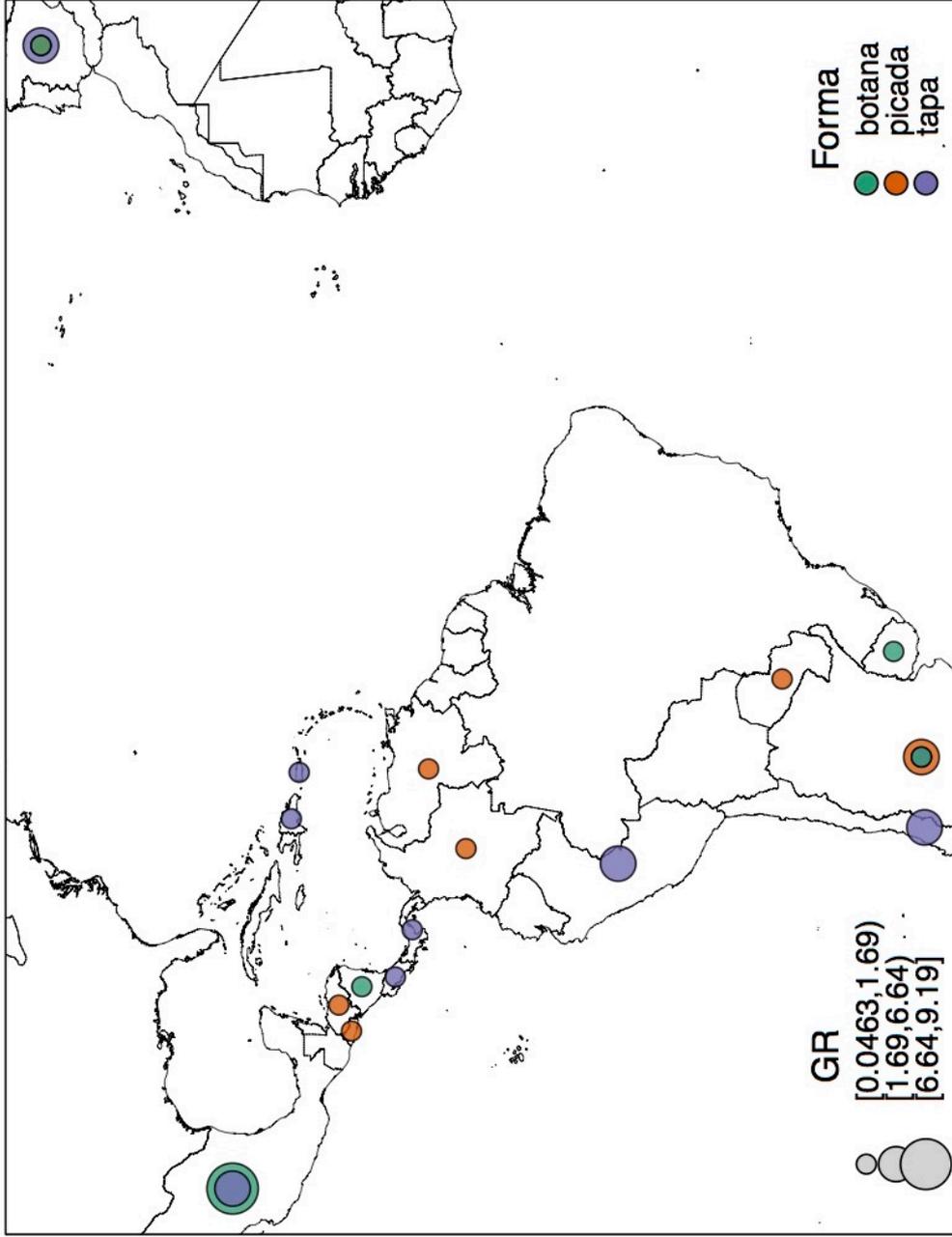
THIN



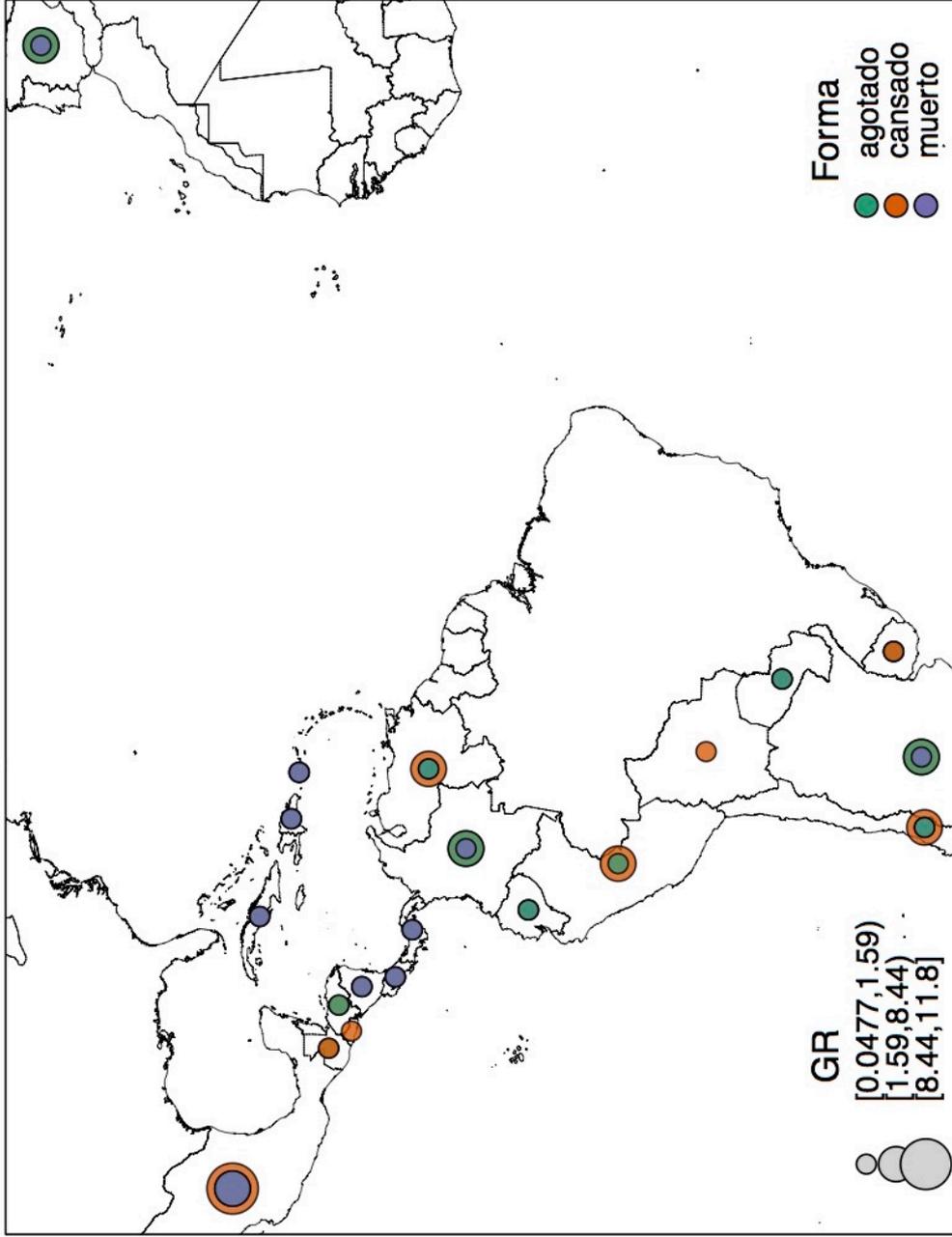
THIN (SKINNY)



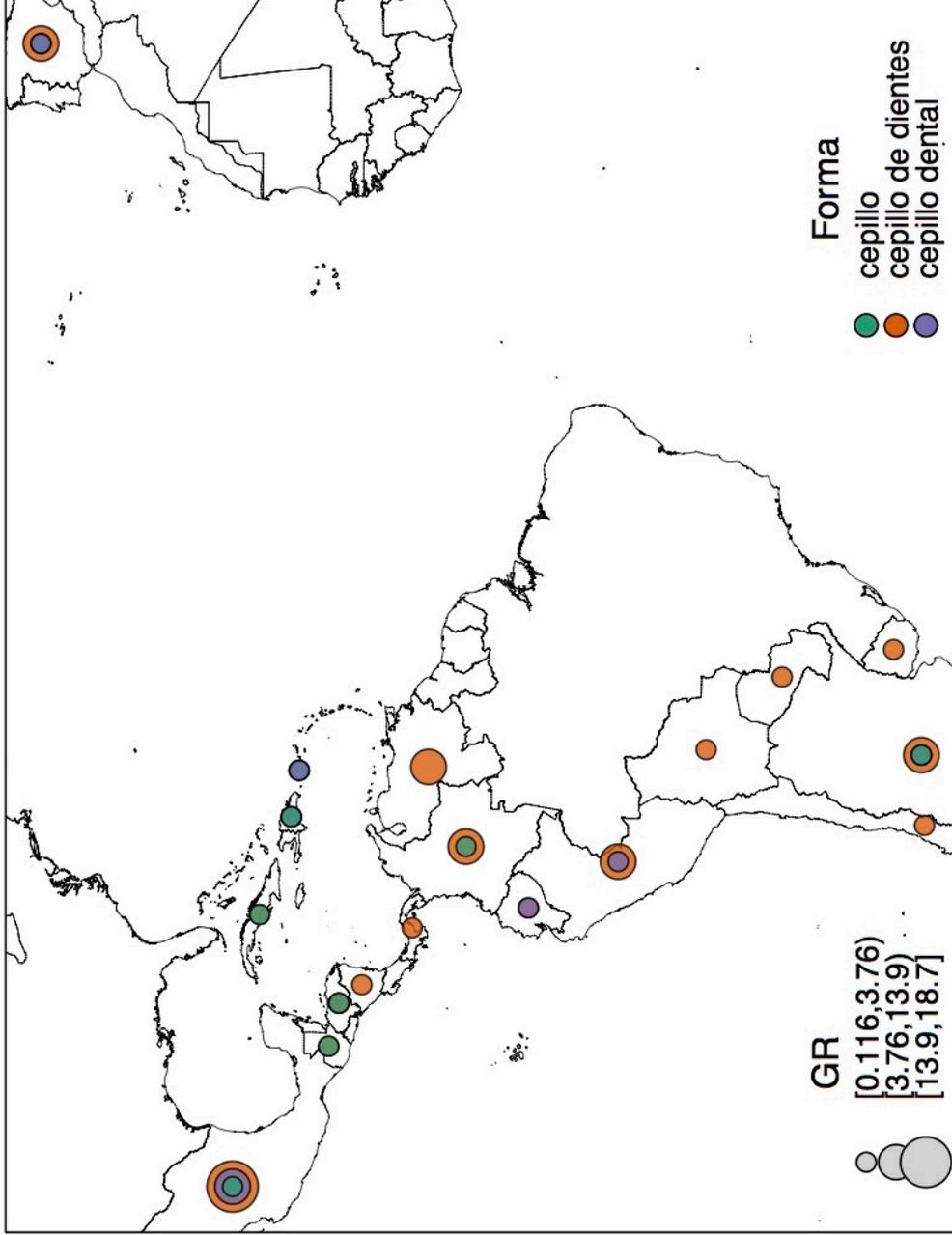
TIDBITS



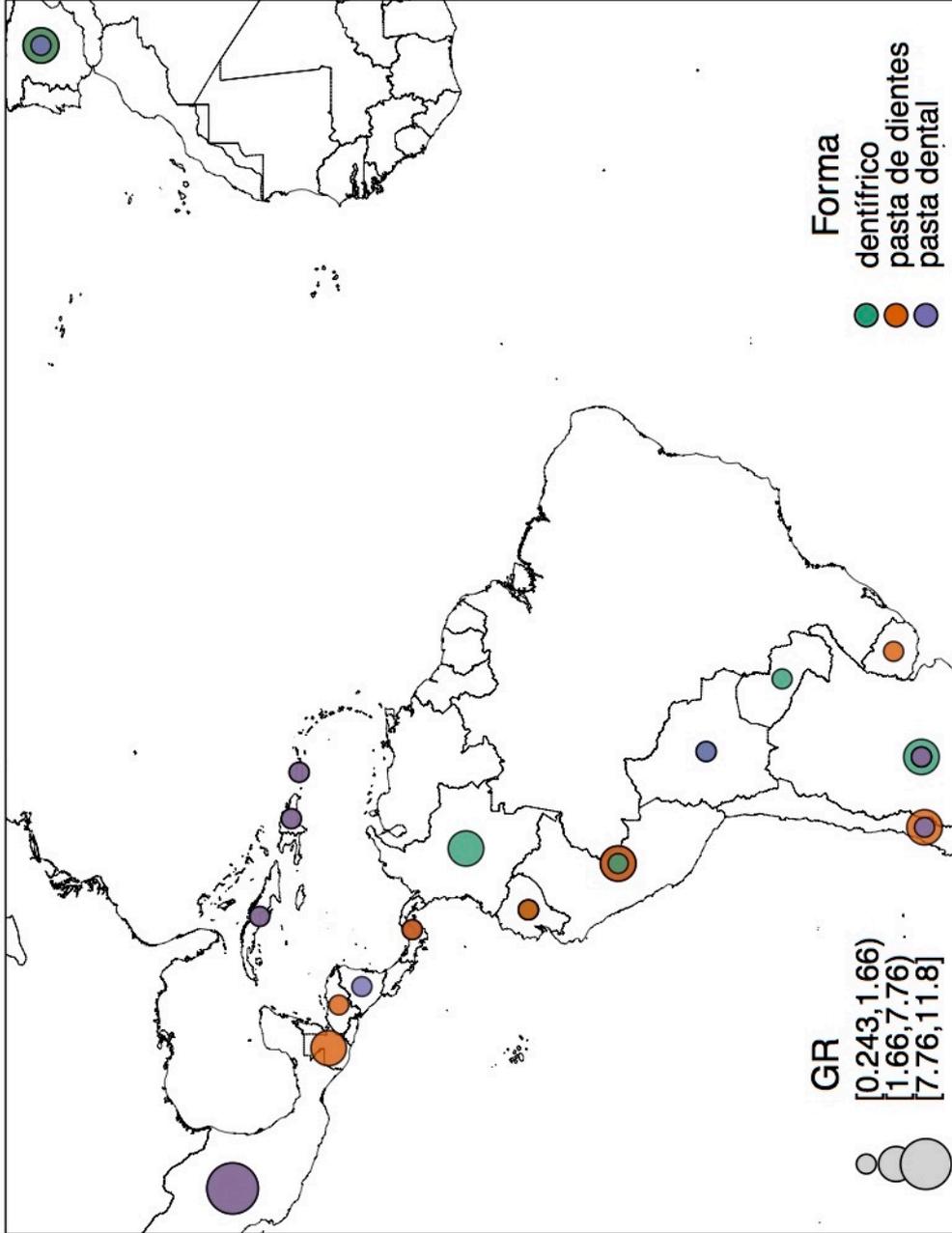
TIRED



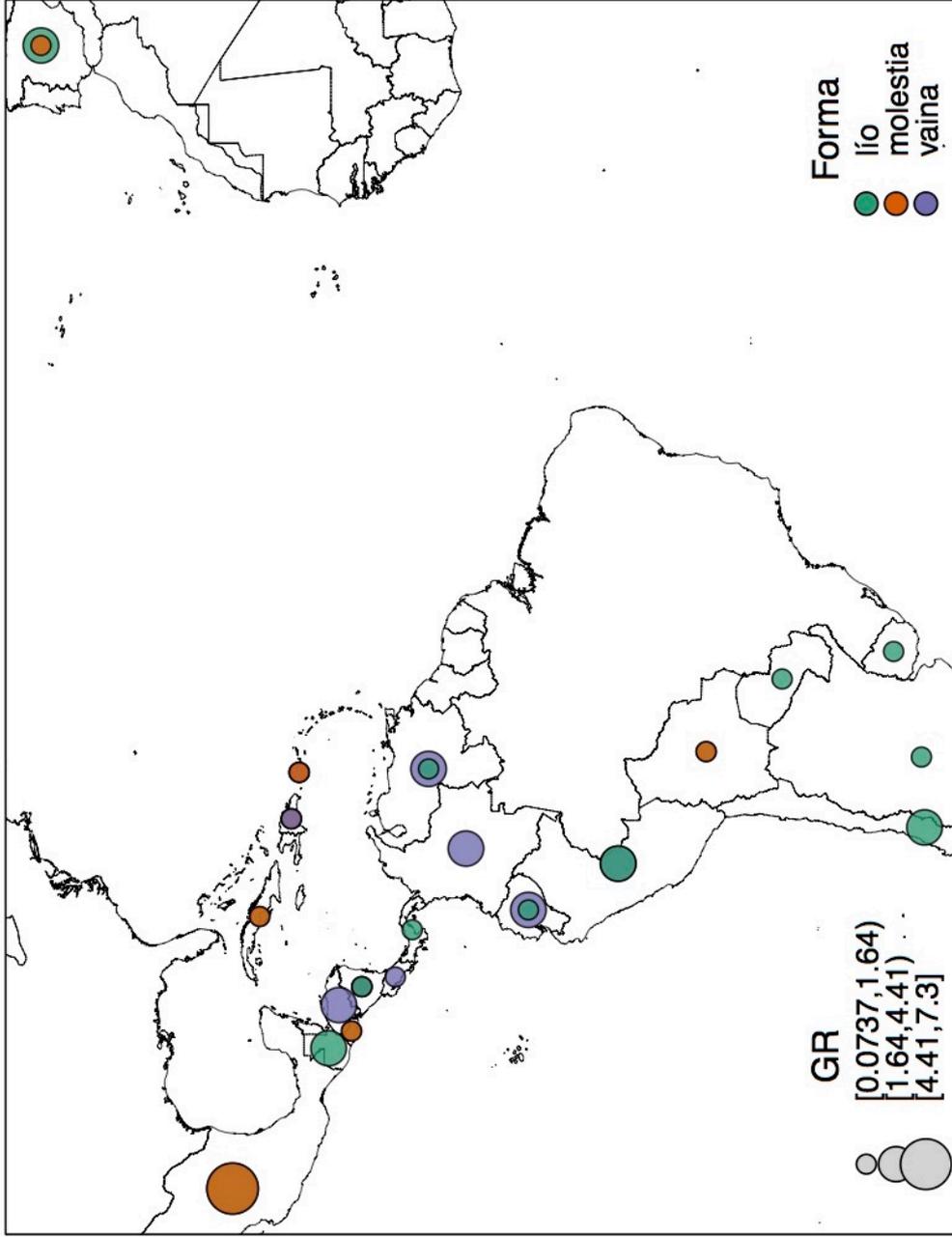
TOOTHBRUSH



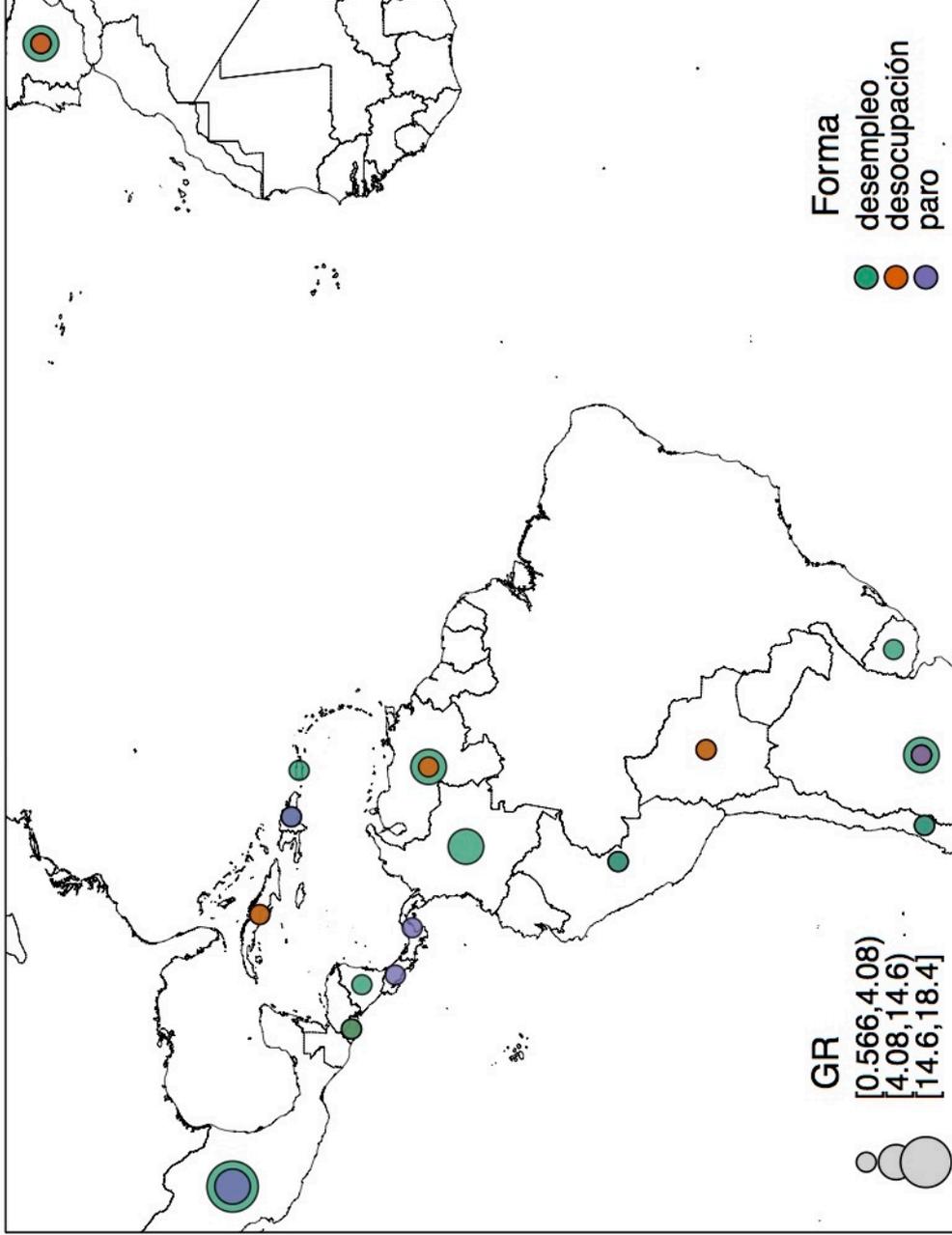
TOOTHPASTE



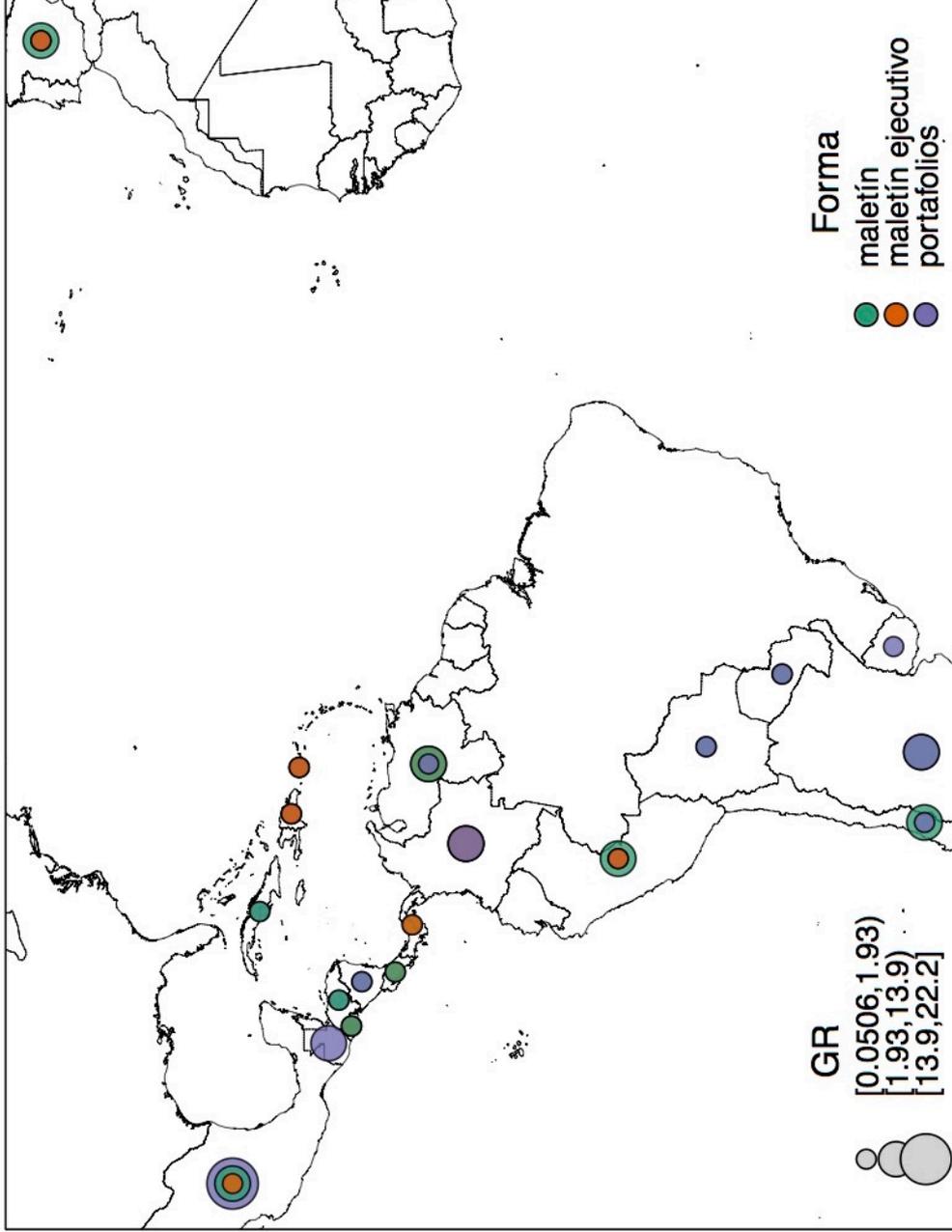
TROUBLE



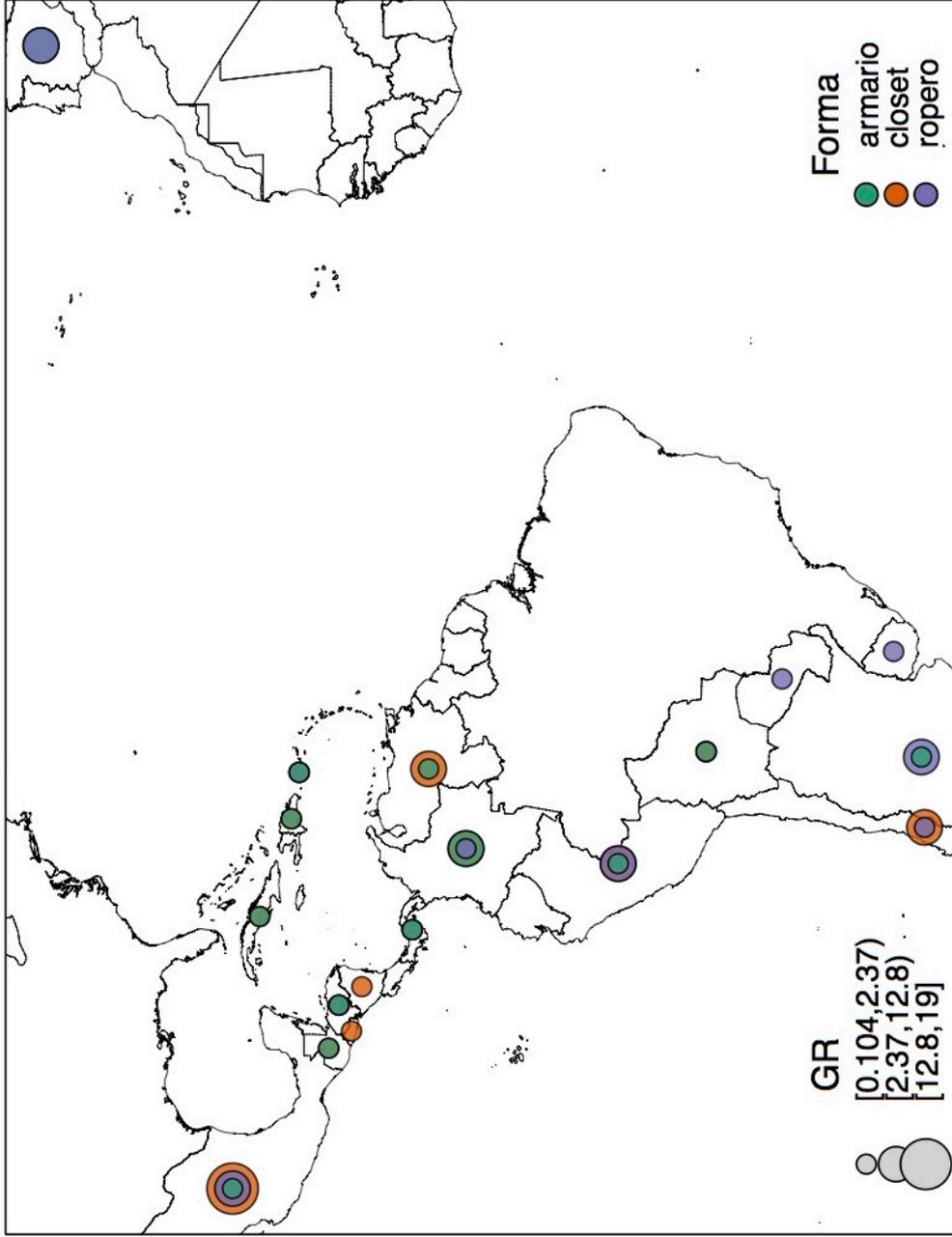
UNEMPLOYMENT



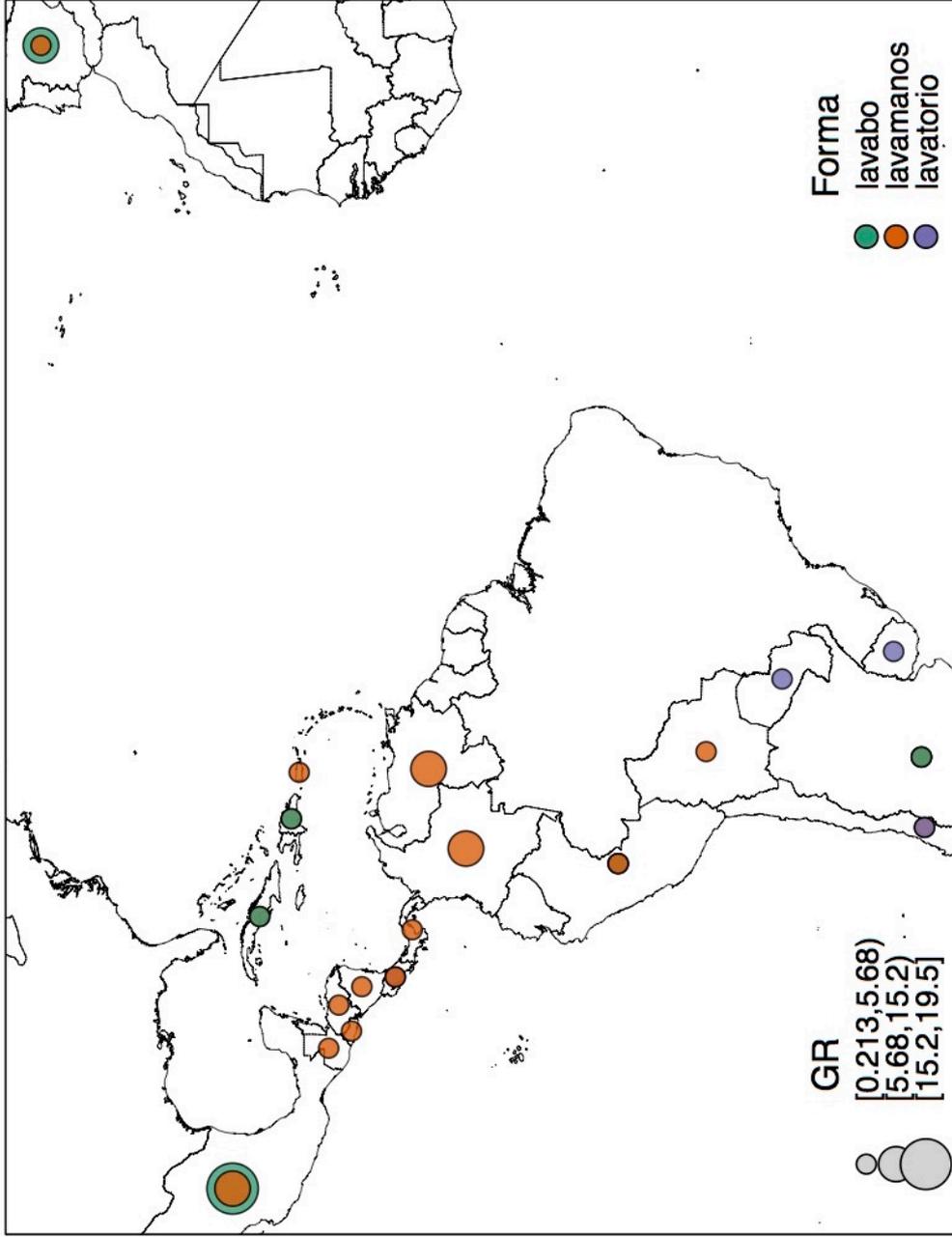
VALISE



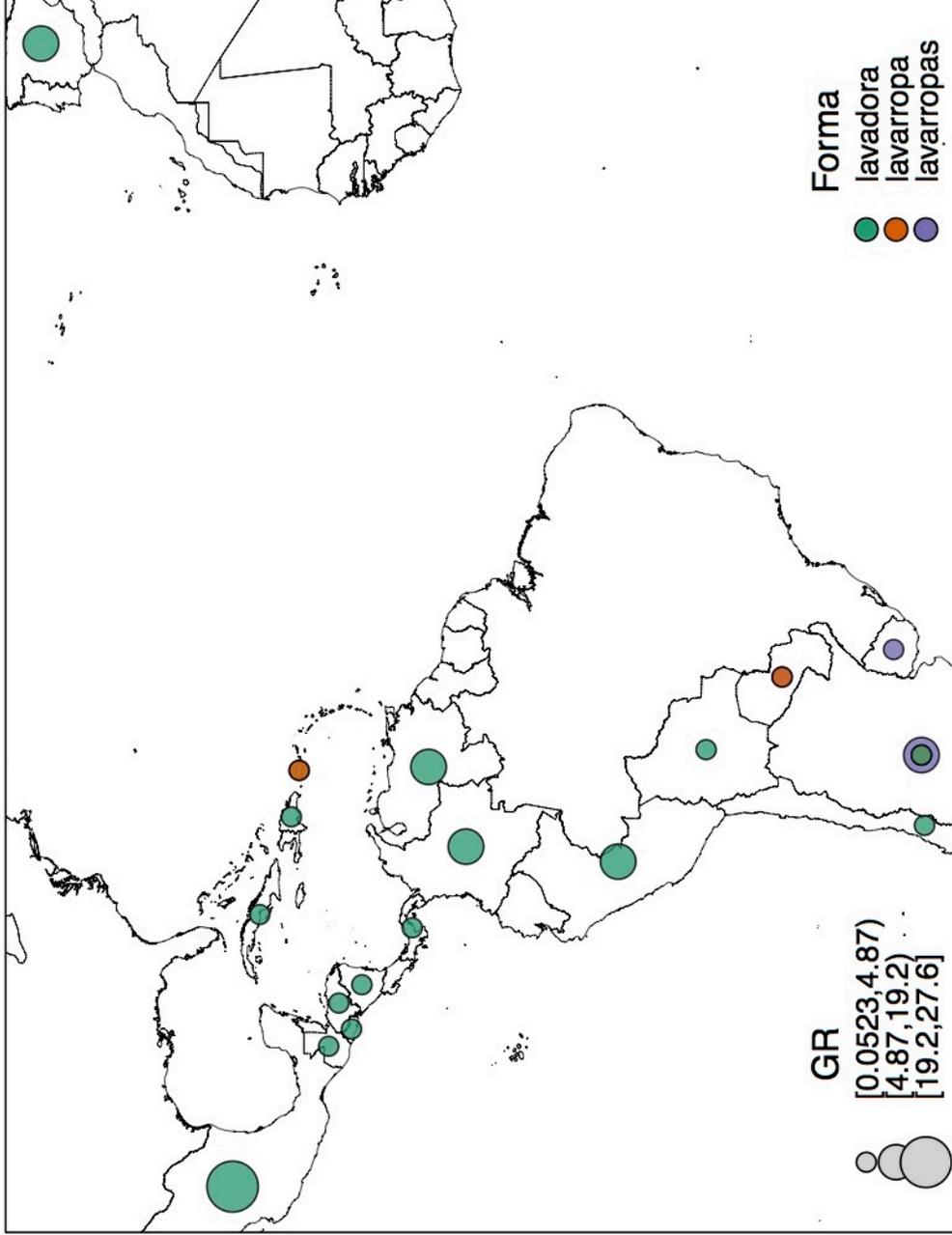
WARDROBE



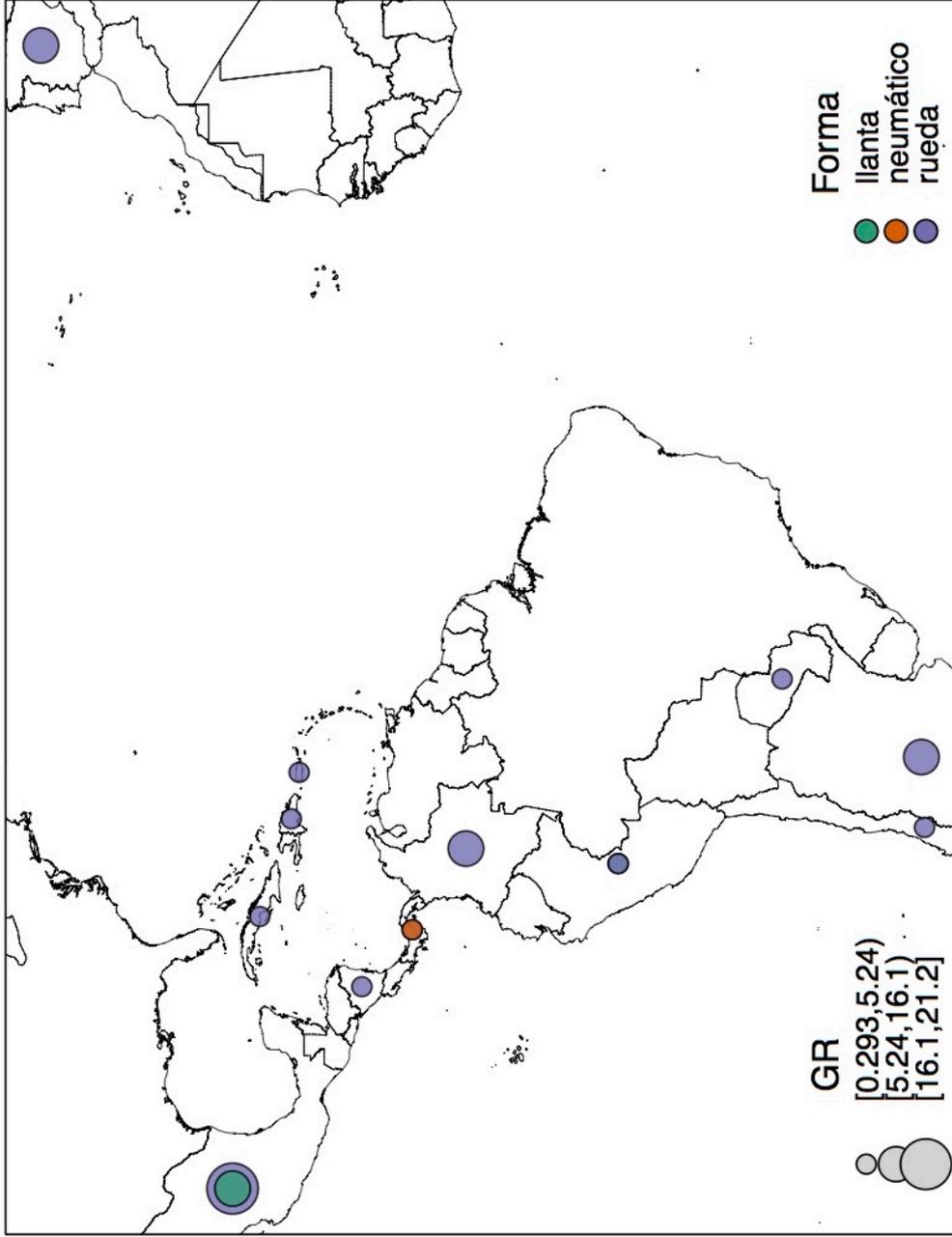
WASHBASIN



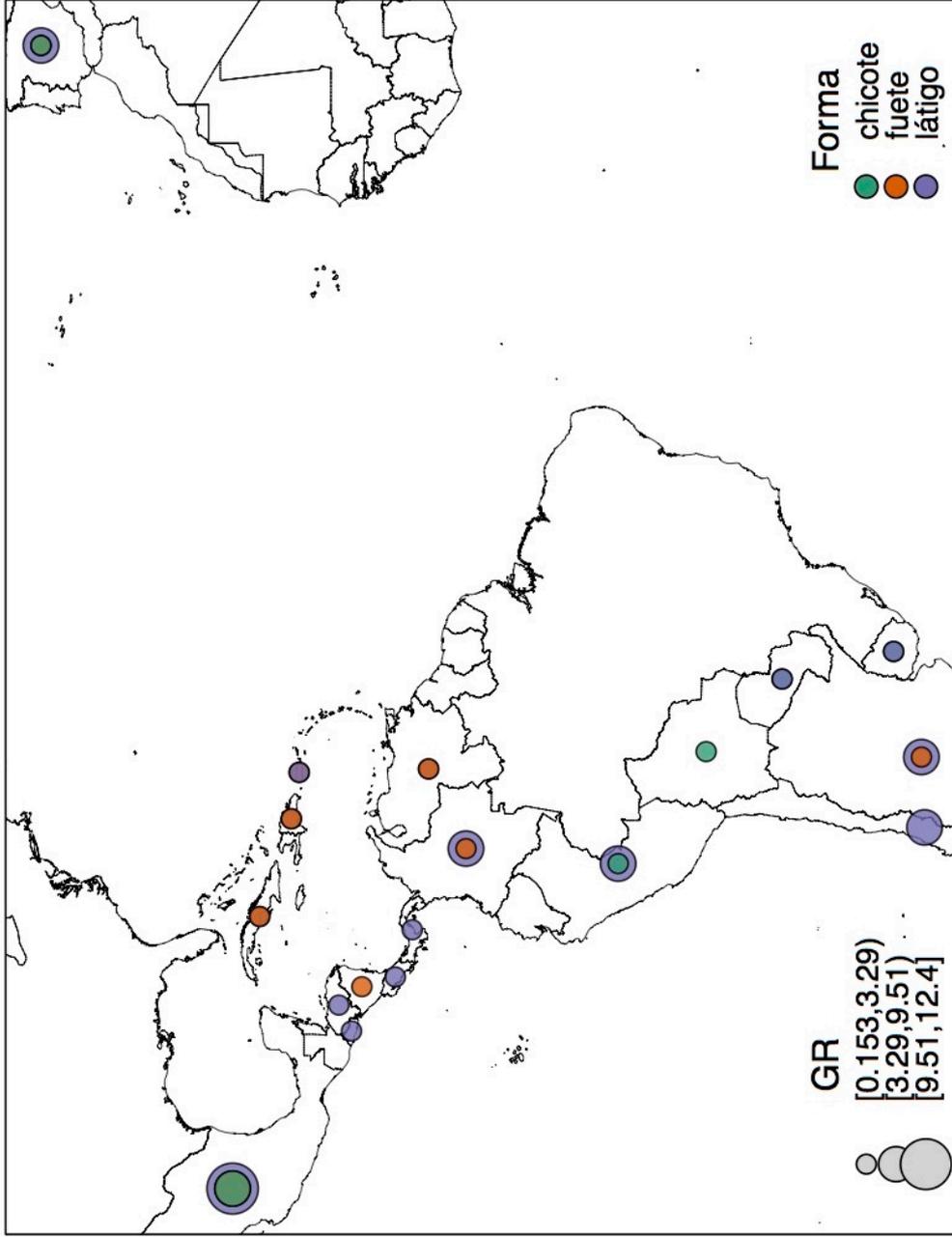
WASHER



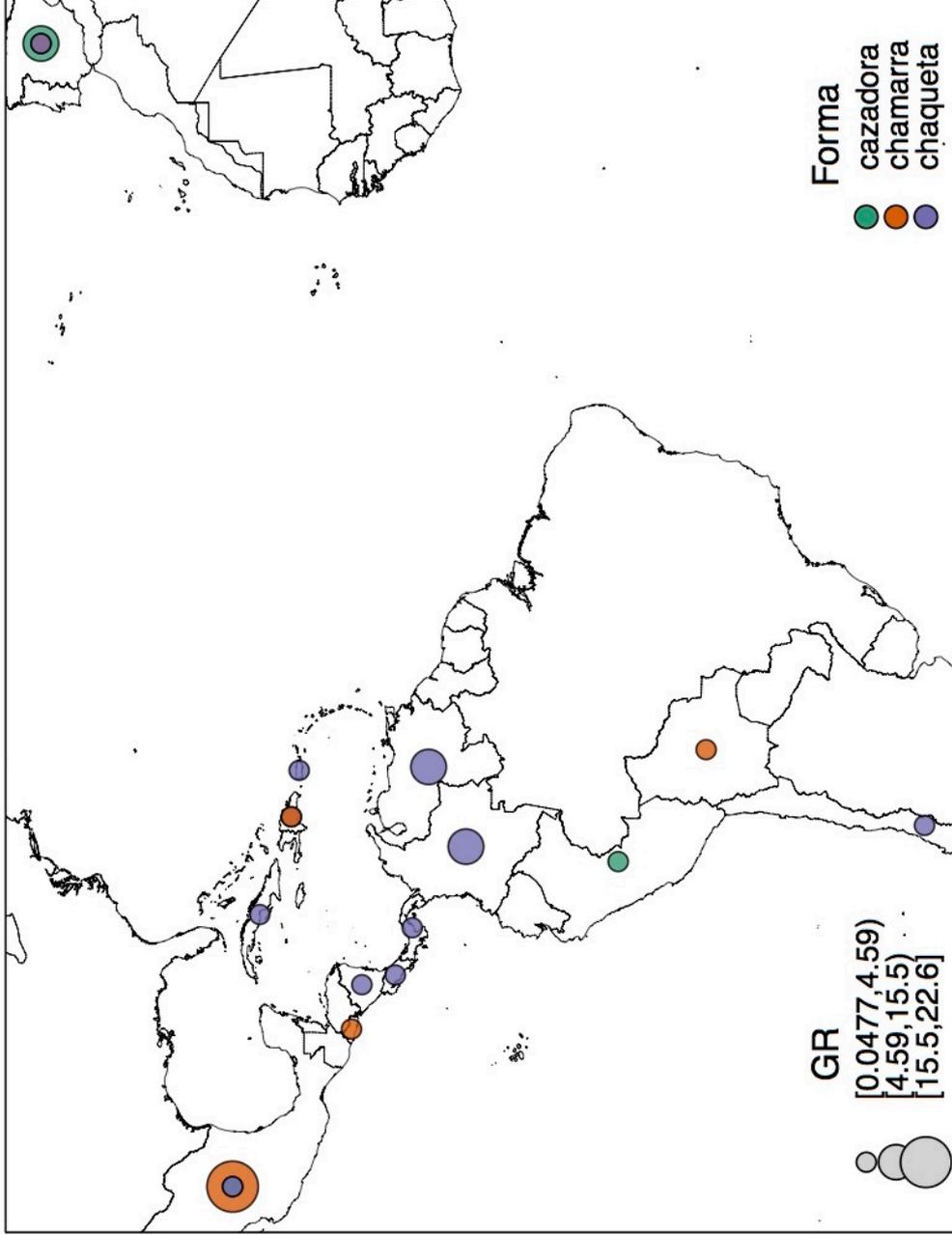
WHEEL



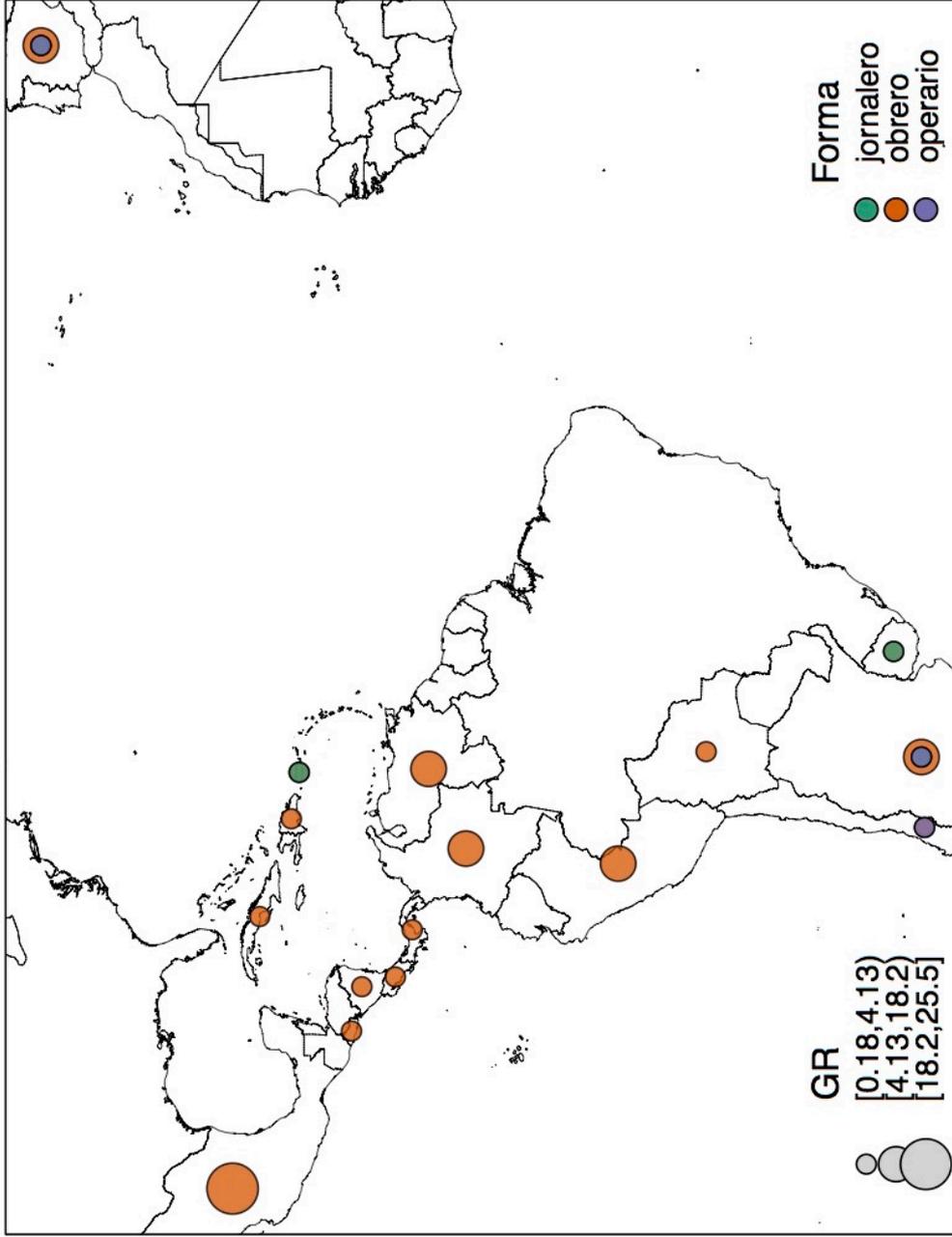
WHIP



WINDBREAKER



WORKER



YOUNG ADULT

