

Marcación morfosintáctica de las categorías de tiempo, aspecto y modo (TAM) en lenguas de Sudamérica: Una aproximación computacional preliminar.

Tatiana Maldonado Muñoz
th.mmunoz@gmail.com
Universidad de Chile

1. INTRODUCCIÓN

Sudamérica presenta una gran diversidad de lenguas, las cuales se organizan en diferentes familias y áreas lingüísticas (Birchall, 2014; Campbell, 2017). Estas pueden tener rasgos compartidos o no, lo que permite establecer la distancia tipológica existente entre ellas respecto a una determinada categoría lingüística (Adelaar & Muysken, 2004)

Nuestro trabajo se centra en las categorías de tiempo, aspecto y modo (TAM), por ser rasgos que pueden expresarse de formas distintas en las lenguas del mundo (Bybee et al., 1994). En el caso de las lenguas de Sudamérica, nos interesa observar la expresión gramatical de estas categorías empleando técnicas de análisis computacionales. Esto es posible al existir bases de datos que registran información de este tipo, como SAILS (empleada en nuestra investigación). Por esta razón, nos proponemos resolver las siguientes preguntas: a) ¿cuál es la tendencia de las lenguas de Sudamérica respecto a la marcación morfosintáctica para rasgos de las categorías TAM? y b) Considerando la información obtenida de la pregunta anterior, ¿cuál es el índice de marcación morfosintáctica para TAM por familias lingüísticas? ¿existe alguna relación con su ubicación geográfica?

Para responderlas nos proponemos como objetivo clasificar las lenguas de Sudamérica según el índice de marcación para las categorías TAM determinando la distancia tipológica entre familias lingüísticas a través de técnicas computacionales. A su vez, consideramos dos hipótesis: 1) La mayoría de

las lenguas de Sudamérica (consideradas en nuestro análisis) tienden a no marcar morfosintácticamente las categorías TAM, siendo una característica general de esta zona (Müller, 2016). 2) Puede establecerse una relación entre la distancia tipológica entre familias lingüísticas de Sudamérica y las zonas geográficas en que se ubican para las categorías TAM, a partir de un índice de marcación.

Los datos de nuestro trabajo, como se mencionó previamente, se obtuvieron de SAILS y se consideraron aquellos proporcionados por Müller (2016) sobre categorías TAM en 64 lenguas de Sudamérica. Se realizó un análisis descriptivo para determinar la tendencia de marcación morfosintáctica de las lenguas en base a un porcentaje calculado por cada rasgo lingüístico. Posteriormente, se agrupó las lenguas por familia lingüística para calcular un índice de marcación, es decir, un promedio de rasgos marcados por cada una para determinar la distancia tipológica existente entre lenguas y ver si existe una relación con la zona geográfica en que se ubican. Nuestra metodología se basó, principalmente, en el empleo de técnicas computacionales para el pre-procesamiento y análisis de los datos.

Nuestros resultados comprueban la primera hipótesis, es decir, existe una tendencia, en la mayoría de los rasgos, a la no marcación de TAM en las lenguas analizadas. Respecto a la segunda hipótesis, se observa que sí existe una distancia tipológica entre familias lingüísticas, pero no se presenta una relación con su distribución geográfica.

El trabajo se estructura en las siguientes secciones: 2) marco conceptual, 3) marco metodológico, 4) metodología, 5) resultados, 6) conclusiones, 7) bibliografía y 8) anexos.

MARCO CONCEPTUAL

1.1. TAM: tiempo, aspecto y modo

La sigla TAM es la abreviación para las categorías de tiempo, aspecto y modo (Dahl,

2015), conceptos que abordaremos desde su dimensión gramatical.

El tiempo es una expresión gramaticalizada de la información deíctica sobre la localización temporal de un determinado evento (Klein, 2009; RAE y ASALE, 2010). Si bien las lenguas pueden expresar esta categoría a través de diversas formas, nos delimitaremos al tiempo gramatical y sus distinciones de: presente, pasado y futuro. Al interactuar el punto de habla (momento en que la oración es producida) y el punto del evento (momento en que la acción referida en la oración sucede) se da una relación temporal. Al tomar el punto de habla como referencia, tenemos que: a) si el evento ocurre antes, corresponde a pasado, si ocurre simultáneamente, corresponde a presente y si ocurre posteriormente, corresponde a futuro (Müller, 2013; Reichenbach, 1984).

Por su parte, la categoría de aspecto -en oposición a la de tiempo- corresponde a la información temporal no deíctica sobre cómo se desarrolla y distribuye un evento en el tiempo (Bosque & Demonte, 1999; García, 1998). Se suele distinguir entre aspecto léxico y aspecto gramatical, pues el primero se relaciona con la información aspectual inherente que aporta el verbo, mientras que el segundo, se refiere a la manifestación morfológica de las diversas formas en que se puede observar un evento (Comrie, 1989). Dentro del aspecto gramatical, se realiza una clasificación general entre: aspecto perfectivo (si el evento puede visualizarse de principio a fin) y aspecto imperfectivo (si solo puede visualizarse una parte interna del evento) (Smith, 1997).

Por último, el modo es una categoría gramatical relacionada con el estado de la proposición que describe el evento. Una distinción es la de *realis* e *irrealis*, aunque controversial dentro del dominio de la modalidad, permite obtener información de lo que un hablante percibe como real o irreal. De esta forma, el modo *realis* retrata situaciones actualizadas, ocurridas o que están

ocurriendo, las cuales son conocidas a través de la percepción directa. El modo *irrealis*, retrata situaciones que son propias del pensamiento, conocidas a través de la imaginación (Palmer, 2001).

2. MARCO METODOLÓGICO

2.1. Áreas lingüísticas

Nuestro trabajo considera la distribución geográfica presente en Birchall (2014). A pesar de no usar el término “áreas lingüísticas” como tal, las regiones mayores propuestas por el autor, están en coherencia con las áreas lingüísticas definidas por Campbell (2017). Existen siete regiones mayores¹: 1) Andes del Norte, 2) Andes Central, 3) Amazonia del Norte, 4) Amazonia Occidental, 5) Amazonia del Sur, 6) Chaco-Planalto y 7) Cono Sur. En base a esta distribución se organiza el análisis de datos de la sección 5.2.

2.2. Marcadores

Un marcador corresponde a una marca con contenido morfológico que permite expresar una determinada categoría gramatical. Para efectos de los datos analizados, cuando se dice que existe marcación morfosintáctica, es porque la categoría se expresa mediante: un afijo, clítico, partícula o un auxiliar que posea información de tiempo, aspecto o modo (Müller, 2013).

2.3. Distancia tipológica

La distancia tipológica es un concepto relevante para la última etapa del análisis de los datos de este trabajo. Corresponde a la diferencia tipológica entre dos o más lenguas (de una misma familia lingüística o diferente) y permite medir cuantitativamente la distancia existente entre estas lenguas respecto a un mismo rasgo (Chiswick & Miller, 2004)

Como la mayoría de las bases de datos se conforma de un conjunto de lenguas y rasgos

¹ Para profundizar en la clasificación de las regiones mayores y las familias que las componen, ver Birchall (2014, p. 7-12).

pueden establecerse valores para ser asignados a dichos rasgos y lenguas. Es posible expresar esto a través de la fórmula $L_x[i]$, donde X es la lengua y [i] es el valor del rasgo a analizar. Lo anterior permite ver qué tan similares o disímiles son las lenguas en un rango de 0.0 (si son idénticas) a 1 (si son completamente distintas) (Hammarström & O'Connor, 2013). Existen diversas fórmulas para este cálculo como la de Hamming o Levenshtein. Sin embargo, para efectos de nuestro trabajo, se asignó a los rasgos TAM un valor en un rango entre 0 y 1 y se representó gráficamente a través de un histograma de frecuencias. Este gráfico muestra, en forma de barras, la distribución de la frecuencia (n° de lenguas en este caso, en el eje y) de los valores continuos de una variable (índice de marcación, eje X) (Pierdant & Rodríguez, 2006).

4. METODOLOGÍA

4.1. Los datos y la muestra

Nuestro trabajo se basa en la sistematización de rasgos TAME proporcionados por Müller (2016), datos disponibles en The South American Indigenous Language Structures (SAILS). Se empleó la muestra definida por la autora conformada por 64 lenguas de Sudamérica (ver la distribución geográfica de las lenguas en el anexo 2). A su vez, clasificamos las familias lingüísticas en regiones mayores siguiendo a Birchall (2014). Las familias lingüísticas de la muestra, el número de lenguas por familia y la clasificación por regiones mayores se presenta en el anexo 1. En nuestros análisis se consideran las lenguas de forma individual y también agrupadas por familias.

4.1.1. Rasgos lingüísticos de la muestra

Müller (2013), realiza una sistematización² de rasgos lingüísticos TAME en relación a la presencia o ausencia de marcadores

morfosintácticos para su expresión. Para nuestro trabajo seleccionamos rasgos asociados a tiempo, aspecto y modo, los que se resumen en Tabla 2.

| RASGOS LINGÜÍSTICOS | | VALORES |
|---------------------|--------------------------|------------------------|
| Tiempo | Presente, Pasado, Futuro | Marcado/ No Marcado |
| Aspecto | Perfectivo, Imperfectivo | |
| Modo | Realis, Irrealis | |

Tabla 1. Rasgos lingüísticos de la muestra.

4.2. Técnicas computacionales para el pre-procesamiento de los datos

Para la conformación y análisis del *dataframe* del presente trabajo, se realizaron dos grandes pasos:

1. Se descargaron los datos de la muestra asociados a los rasgos lingüísticos definidos en la sección anterior desde la base de datos SAILS. Cada rasgo y sus respectivos datos, se registraron en archivos individuales para facilitar su análisis en formato .csv³.
2. Se empleó la aplicación JupyterLab y la librería Pandas (2020) para la implementación de los códigos que permitieron visualizar cada uno de los archivos y clasificar nuestros datos por categorías: tiempo, aspecto y modo.

4.3. Técnicas computacionales para el análisis de los datos

Respecto al análisis de los datos, se emplearon las siguientes técnicas computacionales:

1. Se realizó una clasificación de las lenguas para determinar la tendencia de marcación morfosintáctica TAM por cada rasgo lingüístico, mediante la implementación de diccionarios y creación de tablas. Estos datos se representaron en porcentajes para realizar análisis descriptivos.
2. Con la ayuda de la librería GeoPandas y Matplotlib (Hunter, 2007), se elaboraron mapas para graficar la distribución

² Datos disponibles en SAILS: <https://sails.clld.org/>

³ Los archivos y códigos empleados para el análisis de los datos, se encuentran disponibles en el siguiente link:

<https://github.com/TatianaMaldonadoM/TAM-en-Lenguas-de-Sudam-rica.git>

geográfica de las lenguas conforme a su tendencia de marcación por cada rasgo TAM analizado (ver anexos 2-9). Los pasos 1 y 2 se implementaron para dar respuesta a nuestra primera pregunta de investigación y comprobar la hipótesis 1.

3. Se determinó el índice de marcación de la tendencia morfosintáctica de rasgos TAM por cada lengua. Para esto, se sumó la cantidad de rasgos marcados por lengua y se dividió por el total de rasgos analizados (7). Los valores se registraron en una tabla.
4. El índice de marcación se graficó a través de un histograma de frecuencia con ayuda de la librería Seaborn y Matplotlib.
5. Finalmente, se agruparon las lenguas por familia y se determinó el promedio del índice de marcación por cada una.

5. Resultados

5.1. Descripción de los datos

En esta sección se describen los resultados obtenidos respecto a la tendencia de marcación morfosintáctica de las lenguas por cada categoría⁴.

La tabla 3 resume la tendencia de marcación de los rasgos de la categoría tiempo, respecto a los valores marcado y no marcado. El rasgo presente es el con menor tendencia de marcación de la categoría: un 28% (18 lenguas) tiene el valor marcado y un 72% (46 lenguas), tiene el valor no marcado. Para pasado y futuro se da una tendencia opuesta, pues son los rasgos con mayor tendencia de marcación en todas las categorías con un porcentaje de 70% (45 lenguas) y 86% (55 lenguas), respectivamente. De esta forma, las lenguas con valor no marcado para pasado son un 30% (19 lenguas) y para futuro son solo un 14% (9 lenguas). Estos rasgos se alejan de nuestra

hipótesis inicial y se establecen como los rasgos más codificados en las lenguas de Sudamérica.

| TIEMPO | | | |
|-------------------|----------|----------|----------|
| Valor | Presente | Pasado | Futuro |
| Marcado | 18 (28%) | 45 (70%) | 55 (86%) |
| No marcado | 46 (72%) | 19 (30%) | 9 (14 %) |

Tabla 3. Porcentaje de marcación categoría tiempo.

En la categoría de aspecto, se observa una distribución equitativa de los valores en ambos rasgos, es decir, existe un mismo porcentaje de lenguas que marca perfecto e imperfecto como se observa en la Tabla 4. En este caso, los rasgos tienen una menor tendencia a la marcación con un porcentaje de 61% (39 lenguas) versus un porcentaje de 39% (25 lenguas) con tendencia a la marcación.

| ASPECTO | | |
|-------------------|------------|--------------|
| RASGO | Perfectivo | Imperfectivo |
| Marcado | 25 (39%) | 25 (39%) |
| No marcado | 39 (61%) | 39 (61%) |

Tabla 4. Porcentaje de marcación categoría aspecto

Por último, en la Tabla 5 se registran los resultados obtenidos para la categoría de modo. El rasgo realis tiene la menor tendencia de marcación con un 12,5% (8 lenguas), predominando el valor no marcado con un porcentaje de 87,5% (56 lenguas). Se posiciona como el rasgo TAM menos codificado en las lenguas de Sudamérica analizadas. Para el rasgo irrealis, la tendencia de marcación arroja un porcentaje de 47% (30 lenguas) versus una tendencia a la no marcación con 53% (34 lenguas). sin marcación para este rasgo.

| MODO | | |
|-------------------|------------|----------|
| RASGO | Realis | Irrealis |
| Marcado | 8 (12.5%) | 30 (47%) |
| No marcado | 56 (87.5%) | 34 (53%) |

Tabla 5. Porcentaje de marcación categoría modo

Los resultados presentados en esta sección comprueban nuestra primera hipótesis y responden la pregunta 1. Existe una tendencia a

⁴ Ver sección 8 (Anexos) para ver la distribución geográfica de los valores de cada rasgo por lenguas. El valor marcado se representa

con amarillo, mientras que, el valor no marcado se representa con rojo.

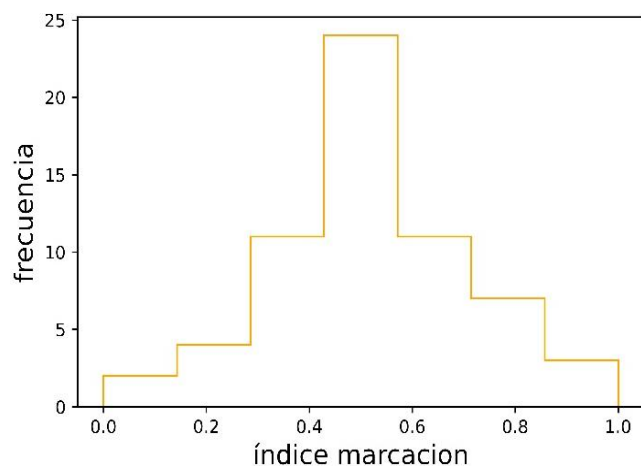
la no marcación de rasgos TAM en la mayoría de las categorías analizadas, pero también existen un par de rasgos que se alejan de esta tendencia: tiempo pasado y futuro. Sobre esto, en una investigación futura sería interesante centrarse en estos datos para analizar si son las mismas lenguas que codifican gramaticalmente estos rasgos o son diferentes. La tendencia de marcación de las categorías TAM (de menor a mayor), se sintetiza en la tabla 6.

| RASGO | % TENDENCIA MARCACIÓN |
|-----------------|-----------------------|
| Modo Realis | 12.5% |
| Tiempo Presente | 28% |
| Aspecto | 39% |
| Modo Irrealis | 47% |
| Tiempo Pasado | 70% |
| Tiempo Futuro | 86% |

Tabla 2. Rasgos lingüísticos ordenados por porcentaje de marcación morfosintáctica.

5.2. Análisis de los datos

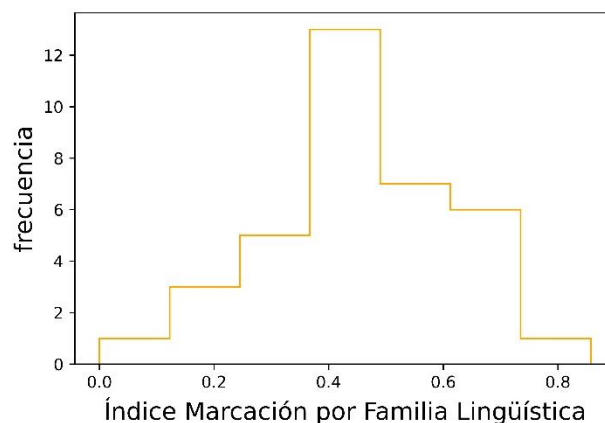
En base a estos resultados, realizamos un análisis cuantitativo para determinar el índice de marcación de cada lengua. Para esto, se sumó la cantidad de rasgos marcados por lengua y se dividió por el total de rasgos TAM considerados en la muestra (7). Los valores obtenidos se representan en el histograma que se presenta a continuación:



Los resultados muestran que las lenguas se concentran en un índice de marcación de 0.5 de un rango de 0 a 1. Por tanto, la mayoría de las

lenguas analizadas tienden a marcar morfosintácticamente, al menos la mitad de los rasgos analizados. También, se encuentran lenguas que registran un bajo índice de marcación entre 0.0 y 0.2, es decir, tienden a no marcar morfosintácticamente rasgos TAM. Por otro lado, encontramos el fenómeno opuesto. Existen lenguas con un alto índice de marcación entre 0.8 y 1.0, es decir, tienen una tendencia a la marcación morfosintáctica de los rasgos TAM.

Determinado el índice de marcación por lengua, consideramos apropiado agruparlas por familia lingüística y calcular el promedio de marcación. Los resultados obtenidos, se representan en el siguiente histograma:



El promedio de marcación por familia lingüística se concentra mayormente entre los rangos 0,4 – 0,6. También, encontramos familias distribuidas en rangos menores 0 – 0,3 y en rangos mayores 0.7 – 0.8.

En la tabla 7, se presenta la descripción de los datos por familia lingüística. Se considera el promedio de marcación por cada familia y la zona geográfica en que se ubican.

| Familia | Promedio marcación | Zona geográfica |
|-------------|--------------------|-------------------------|
| Caribe | 0,8 | Amazonas del Norte |
| Chapacura | 0,7 | Amazonas del Sur |
| Chono | 0,7 | Cono Sur |
| Matacoan | 0,7 | Chaco |
| Chocó | 0,6 | Andes del Norte |
| Quechua | 0,6 | Andes Central |
| Tucana | 0,6 | Amazonas Occidental |
| Aisladas | 0,5 | Amazonas/ Andes Central |
| Barbacoanas | 0,5 | Norte Andes |

| | | |
|-------------|-----|-------------------------|
| Macro-ye | 0,5 | Amazonas del Sur/ Chaco |
| Nambikwára | 0,5 | Amazonas del Sur |
| Pano Tacana | 0,5 | Amazonas Occidental |
| Arahuaca | 0,4 | Amazonas del Norte |
| Arahua | 0,4 | Amazonas Occidental |
| Aymara | 0,4 | Andes Central |
| Borán | 0,4 | Amazonas Occidental |
| Chibchas | 0,4 | Andes del Norte |
| Jívara | 0,4 | Amazonas Occidental |
| Mosetas | 0,4 | Amazonas del Sur |
| Nadahup | 0,4 | Amazonas Occidental |
| Tupí | 0,4 | Amazonas del Sur/ Chaco |
| Araucana | 0,3 | Cono Sur |
| Yanomana | 0,3 | Amazonas del Norte |
| Paez | 0,1 | Andes del Norte |
| Guaicurú | 0 | Chaco |

Tabla 7. Familias lingüísticas: promedio de índice marcación y distribución por regiones mayores.

A partir de los datos anteriormente, se infiere que existe distancia tipológica entre familias lingüísticas. Por ejemplo, si tomamos como referencia la familia quechua que tiene un promedio de marcación de 0.6, se diferencia en mayor grado con familias como la Guaicurú o Paez que tienen un promedio de 0 y 0,1, respectivamente. Por tanto, hay diferencias tipológicas. Esta referencia es solo para ilustrar análisis que pueden realizarse en investigación posterior, pues requiere de profundización y comparación entre cada una de las familias lingüísticas.

Respecto a la relación entre distancia tipológica y ubicación geográfica, a nivel descriptivo, no se evidencia. Los promedios de índices de marcación están ordenados en orden decreciente y la ubicación varía por familia. Se hubiese esperado que, por ejemplo, existiese una diferencia entre familias de Amazonia y Andes Central, pero los promedios se van intercalando entre zonas.

6. Conclusiones y perspectivas futuras

A partir de nuestros análisis podemos concluir lo siguiente:

En relación a nuestra primera pregunta de investigación e hipótesis, los resultados obtenidos comprueban que existe una tendencia a la no marcación de los rasgos TAM considerados en esta investigación. Esta

tendencia se registra en la mayoría de las lenguas. No obstante, esta afirmación involucra un problema de análisis respecto a la muestra, pues se realiza en base a 64 lenguas de Sudamérica cuando el número es mucho mayor (alrededor de 600). Por otra parte, vemos que los rasgos de tiempo pasado y tiempo futuro presentan una tendencia opuesta, es decir, tienden a marcarse morfosintácticamente en la mayoría de las lenguas. Este hallazgo requiere ser profundizado en una investigación posterior para explicar por qué estos rasgos están más gramaticalizados respecto a otros de las categorías TAM. Otro problema de análisis es que no se consideró una comparación para estos rasgos que permitiese identificar si una misma lengua codifica ambos rasgos o no.

En el caso de nuestra pregunta e hipótesis n° 2, nuestros resultados nos permiten responderla de forma parcial. Los histogramas empleados para graficar los resultados del cálculo del índice de marcación de rasgos TAM por lenguas individuales y agrupadas por familia, muestran que, la mayoría de las lenguas codifica al menos la mitad de los rasgos TAM analizados. Posteriormente, el promedio de marcación por familia lingüística permitió obtener un valor que permite calcular distancias tipológicas entre familias. Sin embargo, este último procedimiento no se llevó a cabo en el presente y es materia de investigación para un trabajo posterior. Entre las limitaciones de este análisis puede mencionarse que, no existe una distribución equitativa de lenguas por familia; así, puede haber familias conformadas por una única lengua y en otros casos, hasta por nueve. Por ende, la comparación resulta asimétrica y dificulta el análisis.

No puede establecerse una relación entre el índice de marcación y la ubicación geográfica. Lo que demuestra que nuestra segunda hipótesis no es correcta. Existe distancia tipológica entre familias, pero no se correlaciona con la ubicación. Por ejemplo, no puede realizarse una oposición entre familias lingüísticas de la Amazonia y de Andes Central

porque hay diferentes índices de marcación. Este aspecto requiere de análisis posteriores.

Considerando estos resultados, proponemos dos pasos siguientes para profundizar en el análisis. Por una parte, profundizar en los rasgos de pasado y futuro, comparando las lenguas que codifican estos rasgos y graficar su distribución geográfica por mapas. Este análisis también puede aplicarse por familia lingüística. Por otra parte, es necesario calcular la distancia tipológica propiamente tal para profundizar en las diferencias de marcación morfosintáctica de TAM en las lenguas de nuestra muestra. En este trabajo, solo determinamos un índice de marcación e inferimos algunas posibles relaciones de distancia. Al aplicar estos pasos, esperamos tener una perspectiva más completa sobre la marcación o no marcación morfosintáctica en lenguas de Sudamérica,

El trabajo presentado es solo una aproximación preliminar al análisis de categorías TAM en Sudamérica desde una perspectiva computacional. Como se expuso anteriormente, deben superarse algunos problemas de análisis y profundizar en diversos hallazgos en una investigación posterior.

7. Bibliografía

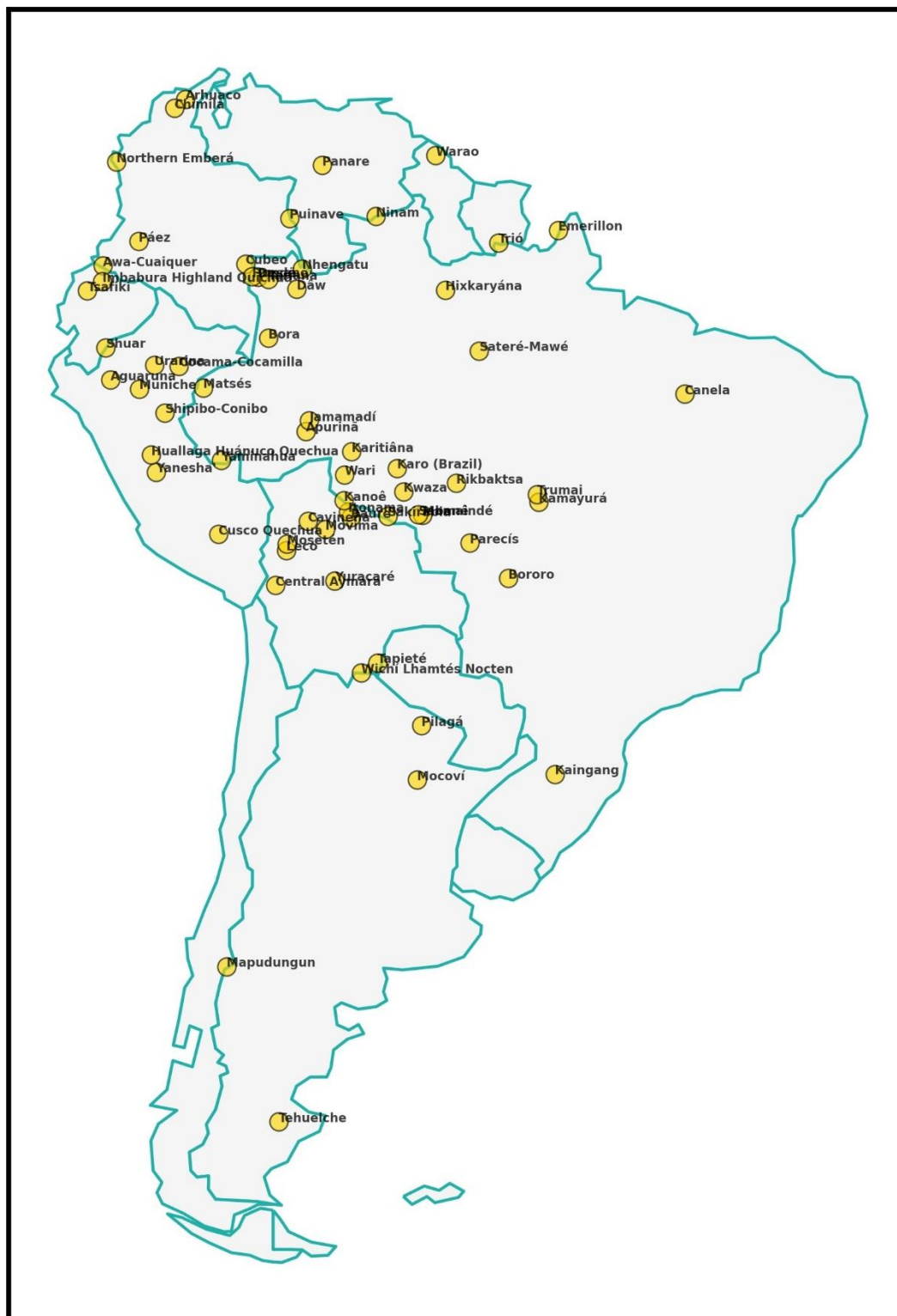
- Adelaar, W., & Muysken, P. (2004). *The languages of the Andes*. Cambridge University Press.
- Birchall, J. (2014). *Argument Marking Patterns in South American Languages*. LOT.
- Bosque, I., & Demonte, V. (1999). *Gramática descriptiva de la lengua española 2. Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales*. (Vol. II). Espasa.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. The University of Chicago Press.
- Campbell, L. (2017). Why is it so Hard to Define a Linguistic Area? In R. Hickey (Ed.), *Handbook of Areal Linguistics* (pp. 19-39). Cambridge University Press.
- Chiswick, B., & Miller, P. (2004). *Linguistic Distance: A Quantitative Measure of the Distance Between English and Other Languages*. IZA.
- Comrie, B. (1989). *Universales de lenguaje y tipología lingüística: sintaxis y morfología* (Vol. 365). Gredos.
- Dahl, Ö. (2015). Tense, Aspect, Mood and Evidentiality. In J. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 210-213). Elsevier.
- García, L. (1998). *El aspecto gramatical en la conjugación*. Arco/Libros.
- Hammarström, H., & O'Connor, L. (2013). Dependency-sensitive typological distance. In S. Anju & B. Lars (Eds.), *Approaches to Measuring Linguistics Differences*. De Gruyter.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 90-95.
- Klein, W. (2009). How time is encoded. In W. Klein & L. Ping (Eds.), *The Expression of Time* (pp. 39-82). De Gruyter Mouton.
- Müller, N. (2013). *Tense, Aspect, Modality and Evidentiality Marking in South American Indigenous Languages*. LOT.
- Müller, N. (2016). Tense-Aspect-Mood-Evidentiality (TAME). In P. e. a. e. Muysken (Ed.), *South American Indian Language Structures (SAILS) Online*. Max Planck Institute for the Science of Human History. <https://sails.clld.org>
- Palmer, F. (2001). *Mood and Modality*. Cambridge University Press.
- Pierdant, A., & Rodríguez, J. (2006). *Elementos básicos de estadística para ciencias sociales*. Universidad Autónoma Metropolitana.
- RAE y ASALE. (2010). *Nueva gramática de la lengua española. Manual*. Espasa.
- Reichenbach, H. (1984). *Elements of Symbolic Logic*. The MacMillan Company.
- Smith, C. (1997). *Parameter of aspect*. Kluwer Academic Publisher.
- The Pandas Development Team. (2020). *Pandas*. In (Version latest) Zenodo.

8. ANEXOS

Anexo 1. Muestra: familias lingüísticas, lenguas y zonas geográficas.

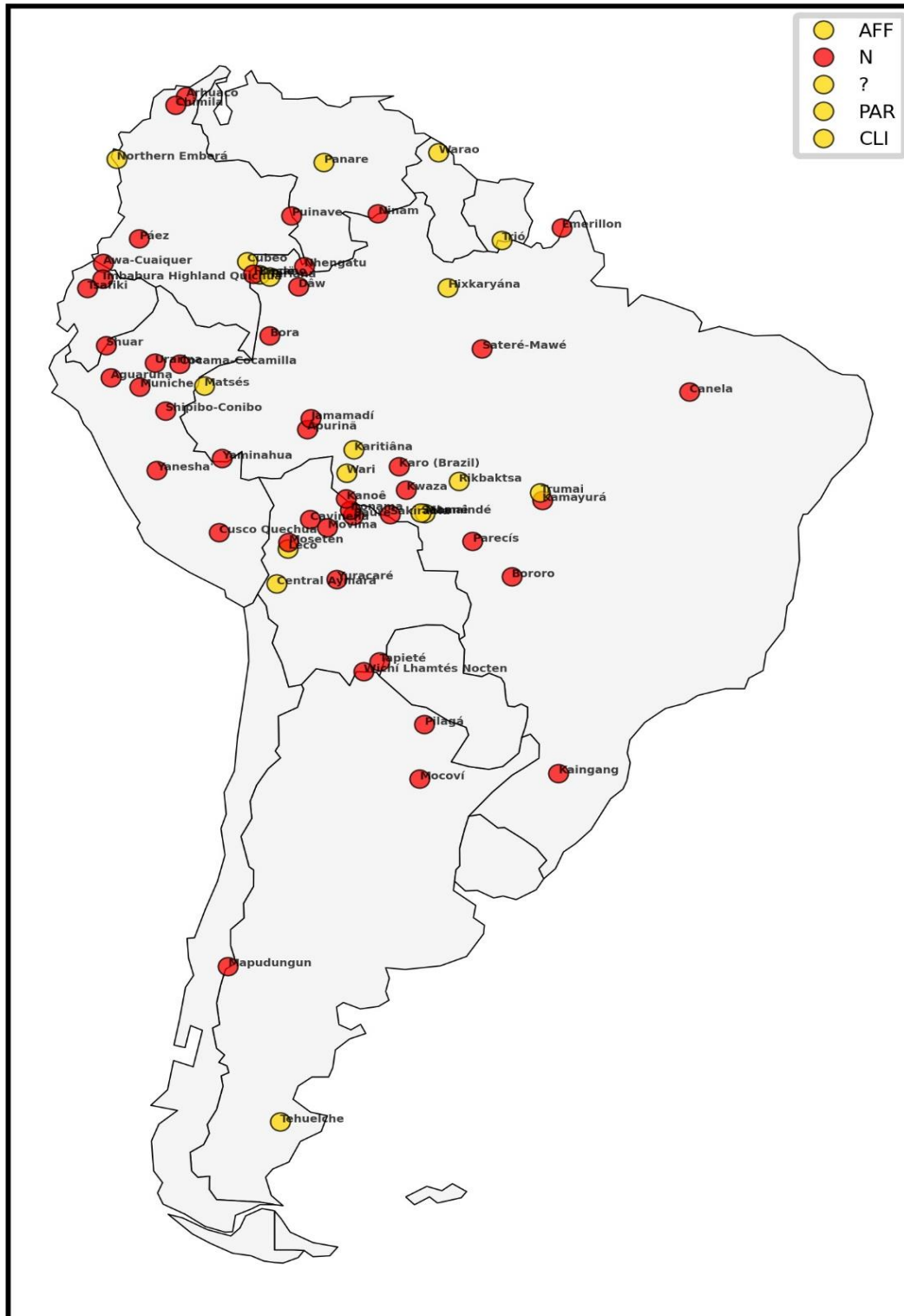
| FAMILIA LINGÜÍSTICA | LENGUAS | REGIÓN GEOGRÁFICA |
|----------------------------|--|--|
| Arahuaca | 1.Apurinã,2.Baure,3.Tariana,4.Yanesha,5.Paresí. | Amazonia |
| Arahuana | 1.Taíno (Jarawara) | Amazonia Occidental |
| Araucana | 1.Mapudungun | Cono Sur |
| Aymara | 1.Aymara | Andes Central |
| Barbacoana | 1.Awá pit, 2)Tsafiki. | Andes Norte |
| Borán | 1.Miraña | Amazonia Occidental |
| Caribe | 1.Hixkaryána, 2.Panare, 3.Tiriyó. | Amazonia Norte |
| Chapacura | 1.Wari | Amazonia Sur |
| Chibchas | 1.Chimila, 2.Iku | Andes Norte |
| Chocó | 1.Emberá | Andes Norte |
| Chono | 1.Tehuelche | Cono Sur |
| Guaicurú | 1.Mocoví,2.Pílaga | Chaco |
| Jívaras | 1.Aguaruna | Amazonia Occidental |
| Macro-ye | 1.Bororo,2.Rikbaktsá,3.Káingang, 4.Timbira | Chaco Amazonia Sur |
| Matacoan | 1.Wichí | Chaco |
| Mosetenas | 1.Mosetén | Amazonia Sur |
| Nadahup | 1.Dâw, 2.Hup | Amazonia Occidental |
| Nambikwára | 1.Mamaindê,2.Sabanê | Amazonia Sur |
| Paez | 1.Nasa Yuwe | Andes Norte |
| Pano | 1.Matsés, 2.Shipibo, 3.Yaminahua | Amazonia Occidental |
| Quechua | 1.Huallaga Quechua, 2.Imbabura Quechua, 3.Cuzco Quechua | Andes Central |
| Tacana | 1.Cavineña | Amazonia Sur |
| Tucana | 1.Cubeo, 2.Desano | Amazonia Occidental |
| Tupí | 1.Karitiana, 2.Karo, 3.Mekéns, 4.Sateré-Mawé, 5.Emérillon, 6.Kamaiurá, 7.Cocama-Cocamilla, 8.Tapiete, 9.Nheegatú | Amazonia |
| Yanomana | 1.Yanam | Amazonia Norte |
| Aisladas | 1.Itomana, 2.Movima,3.Yurakaré, 4.Kanoê,5.Kwaza, 6.Trumai 7.Munichi, 8.Puinave, 9.Urarina, 10.Warao 11.Leko. | Amazonia Sur Amazonia Occidental Andes Central |

Lenguas analizadas



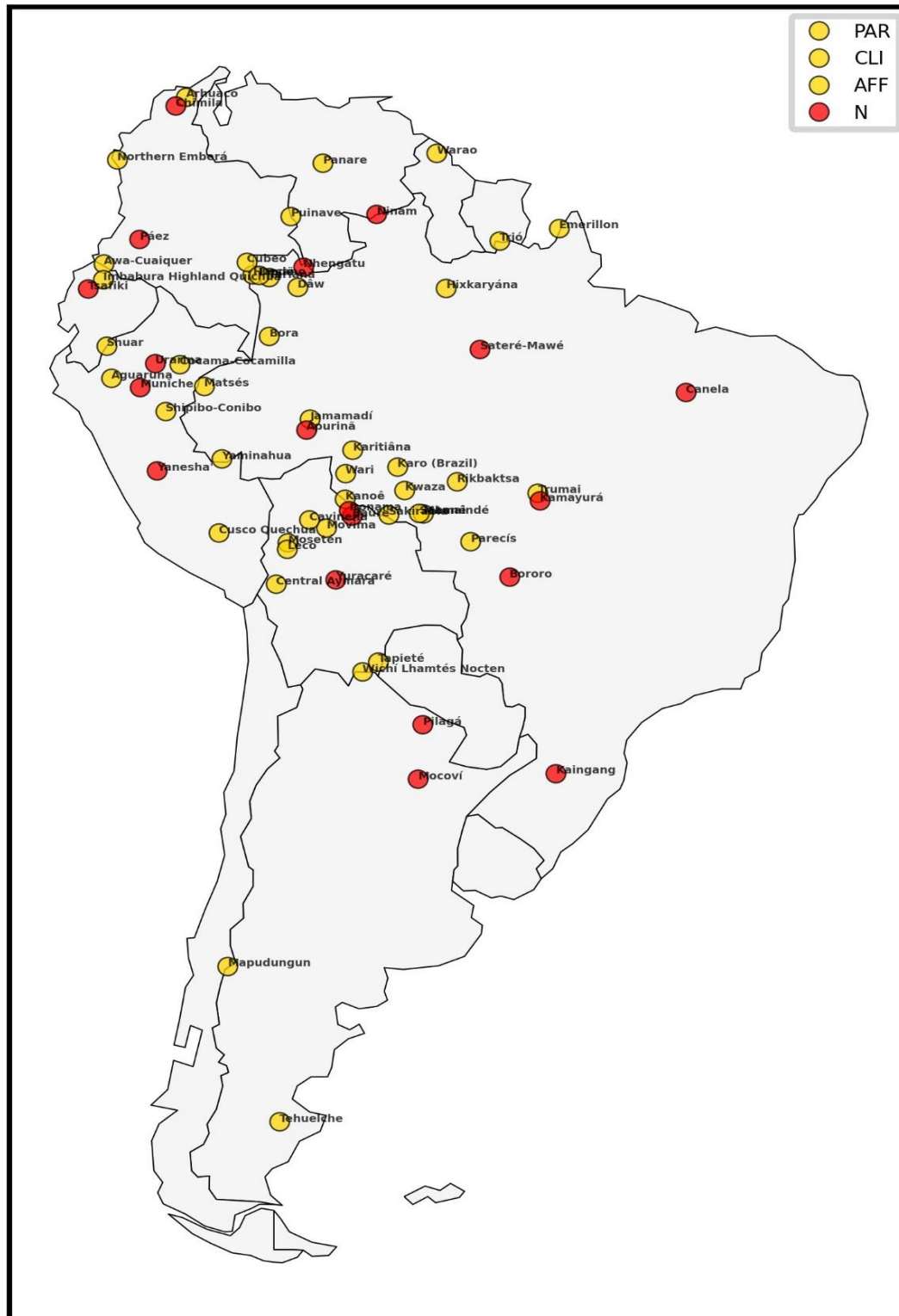
morfosintáctica para tiempo presente

Marcación Presente



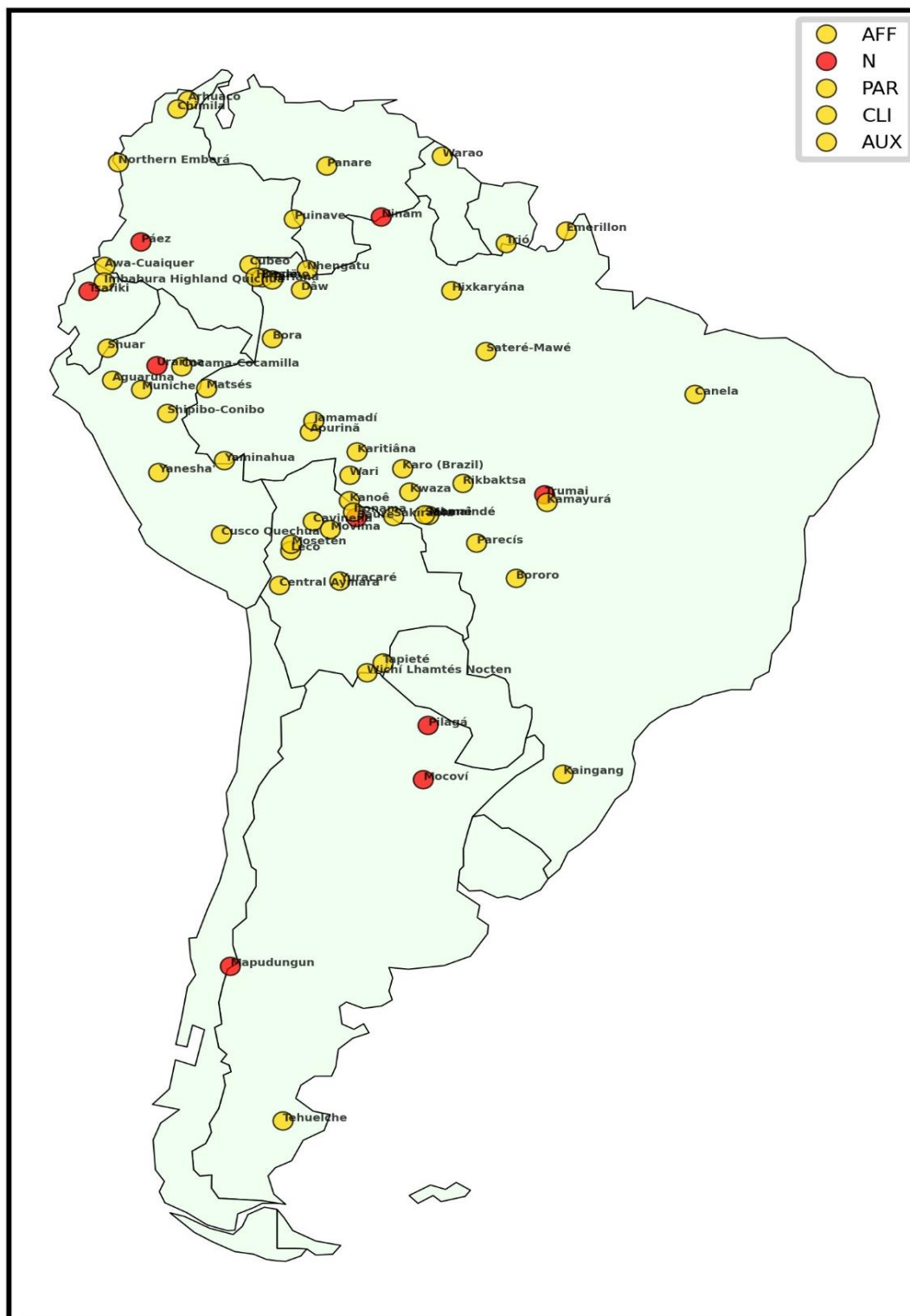
Anexo 2. Mapa Marcación Pasado. Distribución lenguas con marcas morfosintácticas para tiempo pasado.

Marcación Pasado



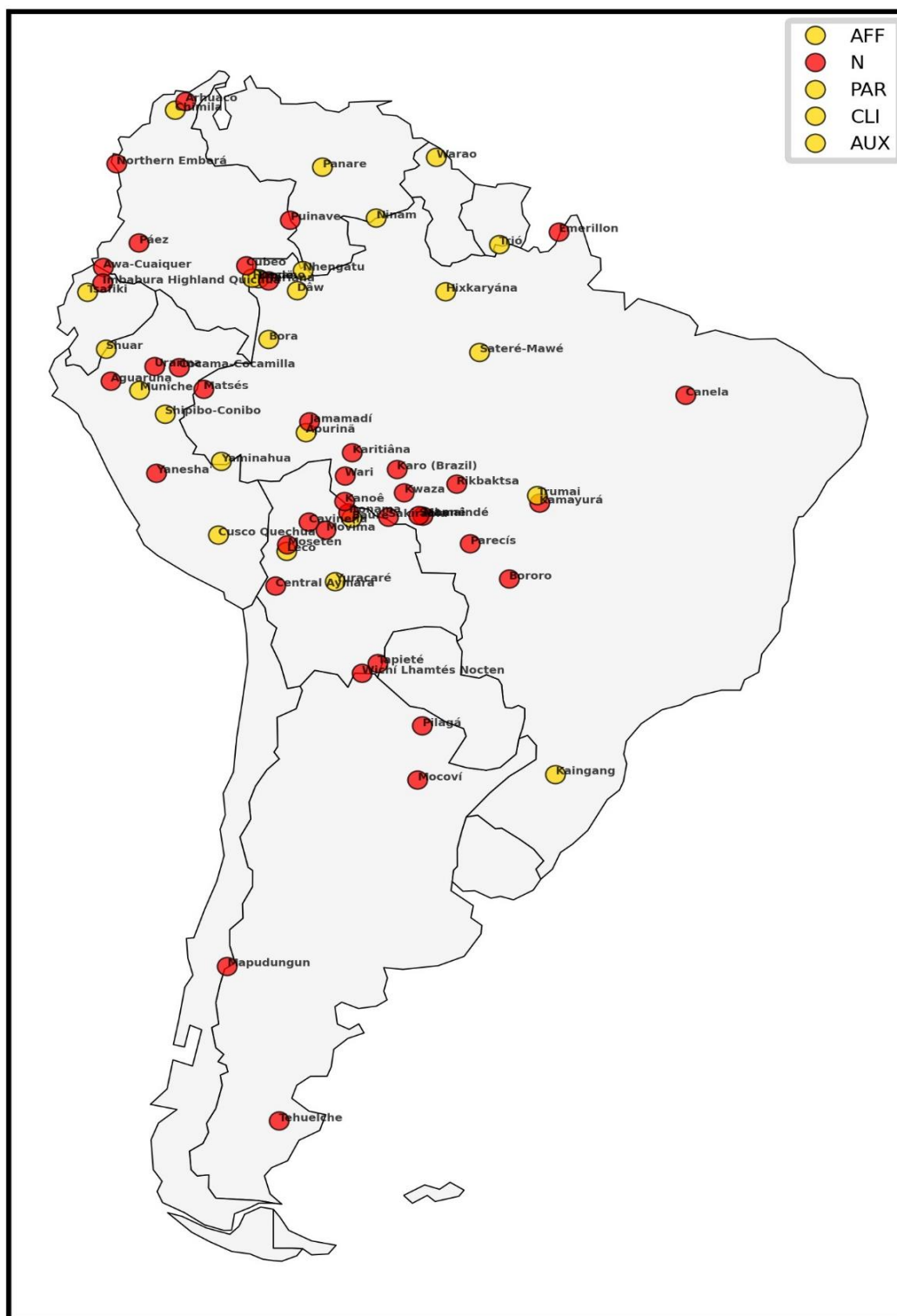
Anexo 3. Mapa Marcación Futuro. Distribución lenguas con marcas morfosintácticas para tiempo futuro.

Marcación Futuro



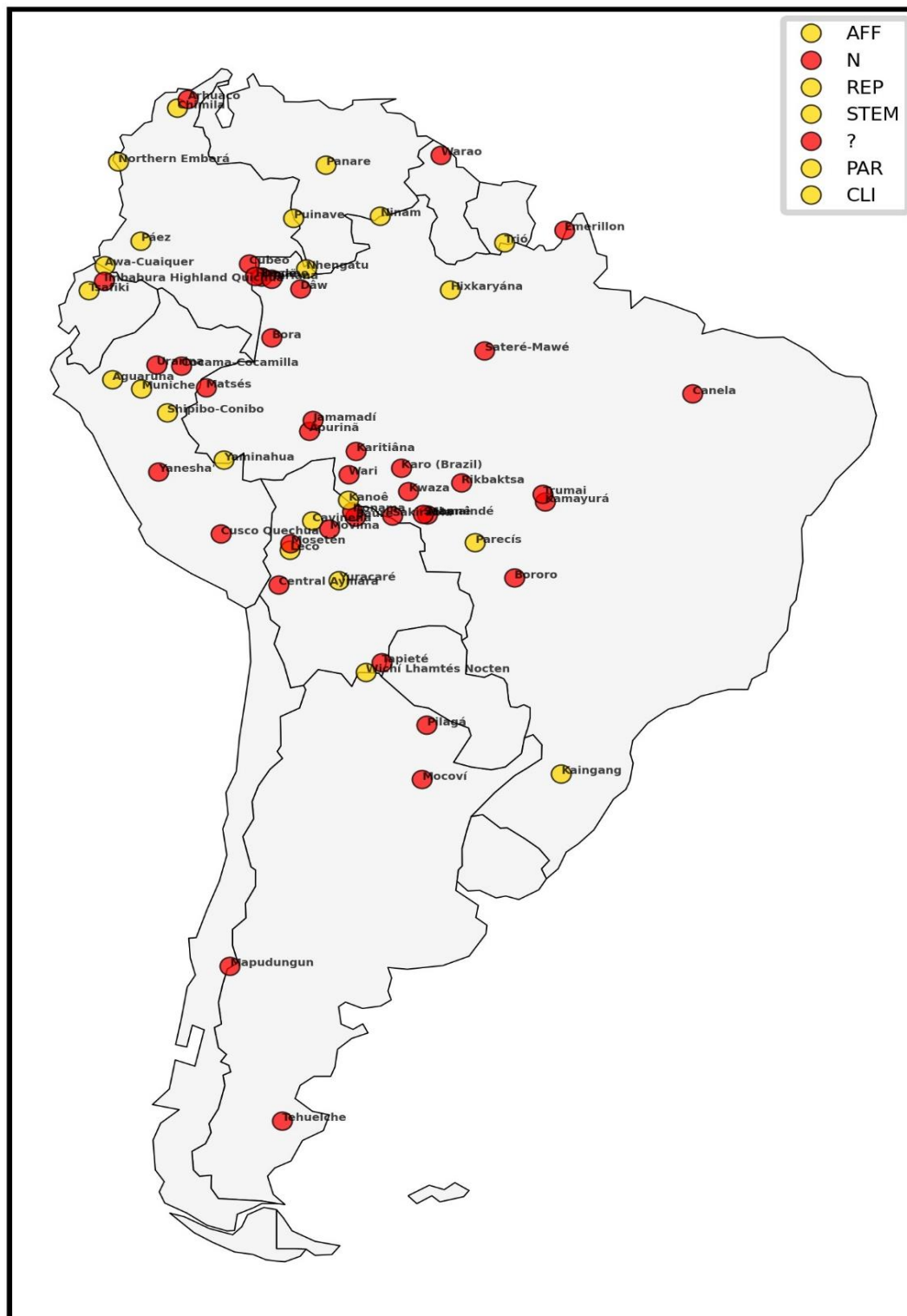
Anexo 4. Mapa Marcación Perfectivo. Distribución lenguas con marcas morfosintácticas para aspecto perfectivo.

Marcación Perfectivo



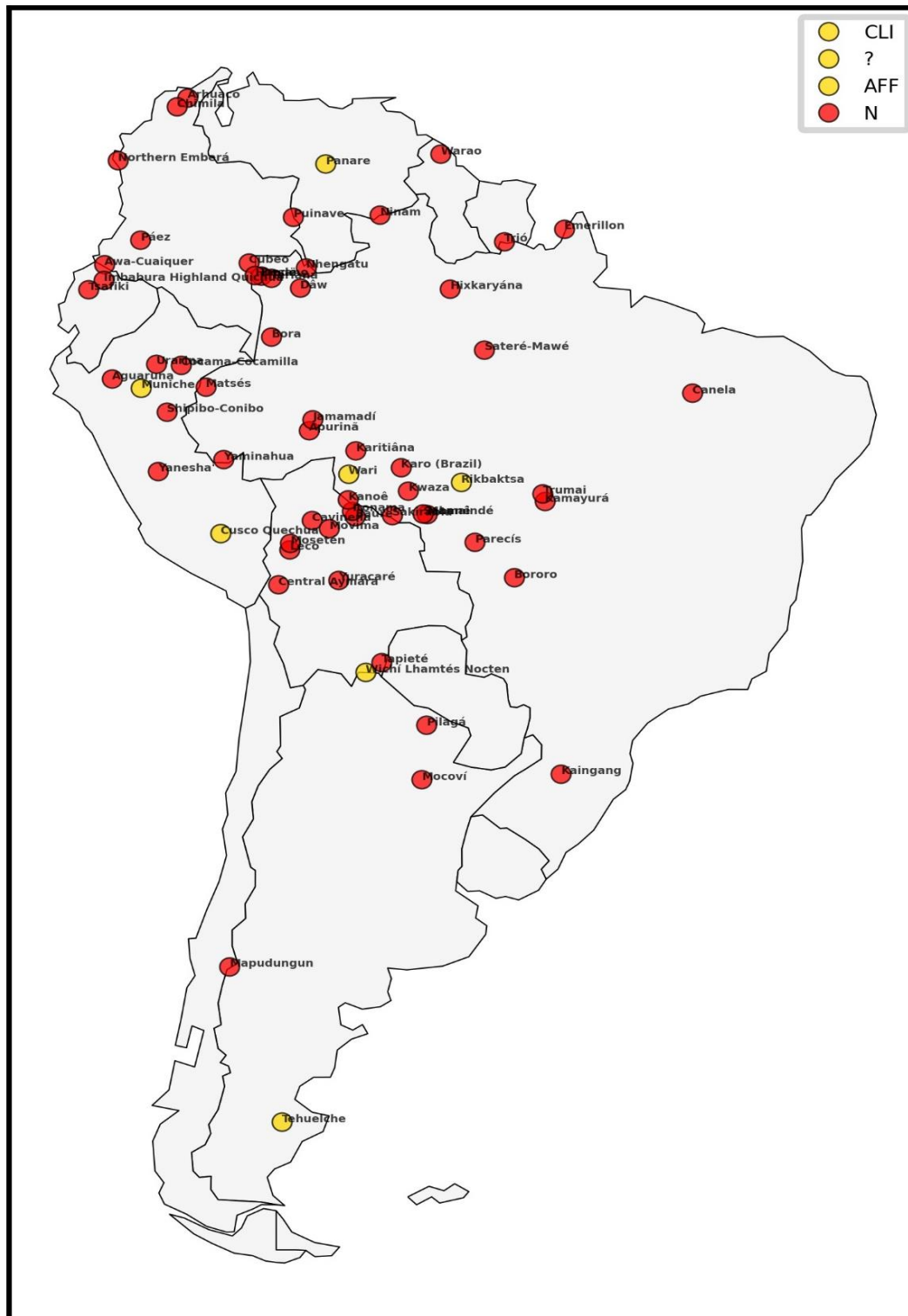
Anexo 5. Mapa Marcación Imperfectivo. Distribución lenguas con marcas morfosintácticas para aspecto imperfectivo.

Marcación Imperfectivo



Anexo 6. Mapa Marcación Realis. Distribución lenguas con marcas morfosintácticas para modo realis.

Marcación Realis



Anexo 7. Mapa Marcación Irrealis. Distribución lenguas con marcas morfosintácticas para modo irrealis.

Marcación Irrealis

