



HAL
open science

UN RETRATO DEL ARTÍCULO VASCO EN EL AÑO 1895, MEDIANTE EL PROGRAMA VDM

Ekaitz Santazilia

► **To cite this version:**

Ekaitz Santazilia. UN RETRATO DEL ARTÍCULO VASCO EN EL AÑO 1895, MEDIANTE EL PROGRAMA VDM. Gotzon Aurekoetxea & José Luis Ormaetxea. Tools for Linguistic Variation, UPV/EHU, pp.175-205, 2010, Supplements of the Anuario de Filología VASca "Julio de Urquijo", LIII. artxibo-00559398v2

HAL Id: artxibo-00559398

<https://artxiker.ccsd.cnrs.fr/artxibo-00559398v2>

Submitted on 12 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UN RETRATO DEL ARTÍCULO VASCO EN EL AÑO 1895, MEDIANTE EL PROGRAMA VDM

Ekaitz Santazilia

UPV/EHU

Abstract

In this paper we offer a concrete application of the computer program VDM, which, given the input of the late XIXth century Bourciez corpus, provides us with clear information about the geographical and functional distribution of the article in Basque at that time. After describing the advantages and defects of the method employed, we discuss a series of linguistic theories about the article, to see how they fit in with the new data provided. Furthermore, by means of this method, which provides data about the geographical distribution of the article, we offer valuable data to explain its progressive introduction into the Basque language. At the same time, we show how the diatopic, as well as the diachronic use of the determiner depends on the type of clause, making it possible to establish a hierarchy and a relative chronology.

Keywords: *Article, Basque, Bourciez collection, VDM program, variation.*

Osaba Jesusendako
Para el tío Jesús
Sortzen denak hiltzea zor

0. Introducción¹

Son pocas las reflexiones realizadas a propósito de los métodos empleados para la investigación lingüística, al menos en lo que al vasco se refiere. La renovación, crítica y mejora de las teorías son el pan de cada día de la actividad científica, pero es necesario, al mismo tiempo, realizar un análisis de los métodos empleados a tal efecto. Son todavía menos los simposios que he conocido como ponente (cualidad intrínseca del neófito, no me puedo preocupar): gracias, por tanto, al grupo de investigación EUDIA,

¹ Este trabajo ha podido llevarse a cabo, en parte, gracias a una beca de colaboración del MEPSyD, en el curso 2008-2009. La ponencia original fue realizada en euskara, bajo el título: “Euskal artikuluen 1895eko argazki bat, VDM programaren bitartez”.

por potenciar dichas reflexiones tan necesarias para la lingüística y por ofrecer a novatos como yo la opción de comenzar a caminar: y todo eso en un sólo simposio.²

Entremos definitivamente en el tema. Hemos realizado una aplicación de un programa informático desarrollado para la dialectometría sincrónica, con el fin de conocer la forma, extensión y estatus del artículo en vasco. No somos nosotros, claro está, los primeros en invertir tiempo pensando en este tema; ni siquiera somos los que más tiempo hemos invertido, pero esta inversión pretende ofrecer un paso más en la descripción de la extensión sincrónica y diacrónica del artículo vasco, empleando para ello los recursos que la informática nos ofrece.

Perdone, pues, el lector, desde el primer momento, los errores que pueda encontrar en el texto y prosigamos.

1. Un poco de luz a la oscuridad del título

Este capítulo pretende ofrecer algunas aclaraciones para la correcta inteligencia del título de este trabajo. En primer lugar, daremos unas generalidades sobre el artículo; explicaremos después el porqué de 1895 y hablaremos también de ciertas cualidades del programa informático *VDM*.

1.1. El artículo

No tenemos tiempo ni espacio aquí para buscar una definición de *artículo*. De todos modos, todo aquél que haya tenido contacto con la lingüística europea ha oído hablar con frecuencia, o al menos esporádicamente, sobre el artículo. Pero de lengua a lengua la presencia y empleo de éste cambia; es decir, en algunas lenguas es necesario sólo el sintagma nominal (SN), donde en otras, para lograr la misma interpretación semántica, es necesario el sintagma determinante (SD).

Longobardi (2001) hace una clasificación tipológica en función del nivel de empleo del artículo. Divide las lenguas de maneras diferentes, dependiendo del nivel de tolerancia de sintagmas nominales desnudos (sin artículo), y de la interpretación semántica que les corresponde:

- (1) a) Lenguas sin SN desnudos.
- b) Lenguas con contados SN desnudos.
- c) Lenguas con SN desnudos más libres.
- d) Lenguas con SN singulares desnudos indefinidos.
- e) Lenguas con SN desnudos exclusivamente (carecen de artículo y tienen interpretaciones ambiguas).

Longobardi inserta al vasco junto con el francés en el grupo de (1a), asumiendo que mientras que en estas lenguas los SD pueden ser argumentos, los SN no (Lon-

² Quiero expresar mi agradecimiento a Ana Gándara, pues con ella comencé en este tema; a Javier Ormazabal, porque dirigió esos primeros pasos; a Xabier Artiagoitia, Urtzi Etxeberria, Iván Igartua y Julen Manterola, porque las conversaciones mantenidas con ellos han sido realmente esclarecedoras y finalmente, al grupo de investigación EUDIA y en especial a Gotzon Aurrekoetxea y Aitor Iglesias, porque han sido compañeros imprescindibles para que este trabajo haya llegado hasta aquí.

gobardi 2001: 581-582). En cualquier caso, entre las lenguas hay diferencias sincrónicas en lo que a presencia del artículo e interpretación semántica se refiere. Nos parece interesante comprobar qué encontramos en vasco. Más adelante (§3) traeremos algunas discusiones entre vascólogos, con el fin de observar de qué manera les afecta nuestra aportación.³

1.2. Un corpus del año 1895

Para reunir datos sobre el artículo, es necesario seleccionar un corpus. Nosotros hemos elegido la colección de textos recogida por Edouard Bourciez. Antes de describirla con precisión, hablemos un poco de los trabajos de este tipo.

1.2.1. *Las recogidas sistemáticas del siglo XIX*

La política lingüística de Napoleón estimó necesario recabar información sobre las lenguas y *patois* que se hablaban en el Imperio Francés; saber cuál era su extensión geográfica y número de hablantes (Oyharçabal 1992a: 351). Los primeros en desarrollar dicha labor fueron Coquebert de Montbret padre e hijo, durante el primer cuarto del siglo XIX aproximadamente. Por encargo del gobierno, se pidió que la Parábola del Hijo Pródigo (y otros textos, cf. Oyharçabal 1992b) fuera traducida a las lenguas y hablas de cada lugar del Imperio.

Las desavenencias políticas evitaron que el trabajo fuera concluido debidamente (Oyharçabal 1994, 1995; Simoni-Aurembou 1989), pero el método perduró en la dialectología francesa, pues posteriormente también será utilizado.

1.2.2. *El trabajo de Bourciez*

Debemos enmarcar la recopilación realizada por Edouard Bourciez dentro de esa corriente dialectológico-sociolingüística. La realizó a finales del siglo XIX; allá por los años 1984-85. En el trabajo *Récueil des idiomes de la Région Gasconne* recopiló, de nuevo mediante las traducciones de la Parábola del Hijo Pródigo, información sobre las hablas de Gascuña. Logró reunir más de 4.000 traslaciones y fueron los maestros de cada pueblo los que asumieron el encargo de realizar las mencionadas traducciones al habla local (Videgain 2005). En la geografía del vasco se recopilieron 150 textos que fueron publicados recientemente por Aurrekoetxea & Videgain (2004).⁴

Por lo tanto, ¿con qué contamos para extraer información sobre el artículo? con textos de unos 150 pueblos, traducciones de la Parábola del Hijo Pródigo realizadas por gente alfabetizada, a finales del siglo XIX. Esto puede ser una ventaja, teniendo en cuenta que el estatus de los traductores es similar, se tradujeron todos en la misma

³ En vasco el artículo es pospositivo y se inserta al final del sintagma. Haciendo una simplificación, tal vez demasiado osada (pero consciente y necesaria) para el vascoparlante y sobre todo para el vascólogo, podríamos decir que *txakur-Ø* es 'perro', *txakurr-a* 'el perro' y *txakurr-ak* '(los) perros'.

⁴ Por lo tanto, nuestro corpus se circunscribe exclusivamente a la zona de habla vasca al norte de los Pirineos, denominada en euskara *Iparralde*, término que emplearemos aquí. *Iparralde* se subdivide a su vez, en tres territorios, que son de oeste a este, Lapurdi, Baja Navarra y Zuberoa.

época y porque tuvieron todos el mismo texto en francés como modelo. Eso nos permite hacer búsquedas sistemáticas con un grado de coherencia alto. De todos modos, en contra de este corpus está la brevedad de los textos, que nos impide poder encontrar todos los ejemplos que quisiéramos. El nivel cultural de los traductores nos hace pensar en que la variedad lingüística recogida pueda distar del habla estrictamente local (al haber introducido cultismos o estructuras no autóctonas). En algunas localidades han sido recogidas dos traducciones y las diferencias entre ambas son notables. En el texto de Izura por ejemplo, hay diferencias en el léxico, grafía y morfología (Aurrekoetxea & Videgain 2004: 145-146):

- (2) a) *Ordou icit ičan naiñ nihaoun buriaz yabé eta oukhan déçaan sosa.*
 b) *Ordu da izan nadin ene naiñi eta ukhan dezadan dihurüØ.*

Por mentar algunas cosas, aunque tal vez salten a la vista, (2a) emplea <ou> para el sonido /u/, mientras (2b) opta por <u>. Para 'dinero', (2a) tiene *sos* y (2b) *dihuru*. Además, (2a) emplea el artículo, la forma determinada (*sosa*), pero es la forma desnuda la que aparece en (2b): *dihuruØ* (vs. *dihurua*).

También puede resultar un problema el haber traducido demasiado fielmente el texto original, por ejemplo, al haber mantenido el orden de palabras del francés, en lugar del del vasco. Reproducimos a continuación un ejemplo de Anhauze (Aurrekoetxea & Videgain 2004: 175):

- (3) a) *Zenbait egunen burian tzarra seme zen yuan herritic eguinez fierrain eta gabe erran adioric nehorì.*

Tenemos *tzarra seme* con el adjetivo antepuesto, en lugar del esperable *seme tzarra*; *eguinez fierrain* con el verbo antepuesto al objeto, en lugar de *fierrain eguinez* y *gabe erran* en lugar de *erran gabe*, donde la posposición *gabe* ha sido misteriosamente transformada en preposición.

De cualquier modo y aunque el corpus se limite exclusivamente a Iparralde, el de Bourciez es un tesoro que hay que explotar, puesto que son escasas, en el caso del vasco al menos, estas recopilaciones sistemáticas: nos viene a la mente, aparte del de Bourciez, el trabajo de Sacaze. Si bien la edición moderna de esa recopilación realizada en 1887 verá la luz en breve gracias al grupo de investigación EUDIA, J. Allières ya dio a conocer los datos sobre el vasco de dicha recolecta (Sacaze 1887, Allières 1960-1961 y esta página web: <http://www.garae.fr/spip.php?article191>). Regresando al corpus de Bourciez, tenemos que decir que como la edición de Aurrekoetxea y Videgain nos ofrece los textos en soporte digital y transcritos, la realización de búsquedas y bases de datos ha resultado una labor más amena.

1.3. ¿Qué es VDM?

Visual Dialectometry (VDM) es un programa informático desarrollado por el informático E. Haimlerl, bajo la supervisión del lingüista austríaco H. Goebel, para cuya utilización el grupo de investigación EUDIA posee una licencia. La labor del programa es transformar corpus en mapas:

- (4) CORPUS → Datos estadísticos → MAPA

Tras introducir el corpus en tablas y lematizarlo, el programa ofrece la posibilidad de realizar numerosas operaciones estadísticas, cruzando datos de tantas tablas como queramos, para medir la variación dialectal conforme a diversos parámetros. En el último paso tenemos la opción de mostrar los datos en mapas cartográficos. Hemos transcrito el segundo paso de (4) en minúsculas, porque es optativo: en nuestro caso al menos, hemos omitido los datos estadísticos para observar directamente el mapa.

Dejando a un lado las enormes posibilidades que el programa ofrece, nosotros hemos hecho una simple pero productiva aplicación, transformando directamente el corpus en mapas; es decir, hemos usado *VDM* para transformar los datos extraídos pueblo a pueblo del corpus de Bourciez en mapas coropletas de Volnoi. De esa manera, podemos saber con un simple vistazo si de pueblo a pueblo hay coherencia geolingüística en el empleo del artículo. Hemos realizado un mapa por cada frase que hemos elegido, siguiendo el procedimiento que a continuación detallamos.

2. Procedimiento y metodología de trabajo

2.1. La explotación del corpus

Es un trabajo manual. En la edición de Bourciez de Aurrekoetxea & Videgain (2004) los textos están transcritos y en formato digital, pero se han transcrito en su grafía original. Esto nos impide poder hacer búsquedas automáticas con un buscador y hay que mirar los textos uno a uno, ya que como se muestra en (2), la grafía puede variar enormemente, incluso dentro del mismo pueblo.

Para comenzar, elegimos textos de localidades distantes entre sí (unos 5 textos al azar) y buscamos oraciones con sintagmas desnudos. Una vez hallados, observamos cómo se había traducido ese sintagma o esa oración en todos los pueblos. Así pues, por cada frase que incluía un sintagma nominal desnudo interesante, creamos una tabla en *Access*.

Tabla 1

Ejemplo de una tabla de la base de datos

Datubase5-gasna			
_KarteNr	_OrtNr	Testu hitza	Lema
217	1193	etchia bethiaüc muthilez çoïneç baïtute ogia eta arnua, arrolziac eta gasna	2
217	1198	bethia da sehiz, soïneç baïtute oguia eta arnoa, arroltciak eta gasna	2
217	1201	séhis béthia da, çoïgneç baïtouté ogui éta arno, arroltcé eta gasna // 1202: sehiz bethea da, oghi, arno, arrautze eta gasna baïtouté ogui éta arno, arroltcé eta gasna	1
217	1203	bethea da sehiez doutenac oguia eta arnoa arroltceac eta gasna	2

He aquí el fragmento de la tabla correspondiente a una oración del corpus. En la primera columna aparece el código que le hemos dado a la oración que comprobaremos pueblo a pueblo, así como al mapa que extraeremos después (217). En la siguiente vemos el código que corresponde a cada pueblo (el 1193 por ejemplo, pertenece a Lehuntze, etc.). En tercer lugar, tenemos el texto extraído del corpus: el tener el texto en formato electrónico, en CD, facilita enormemente este trabajo, pues no hay más que copiar y pegar. Finalmente, hemos realizado una lematización.

Tal y como se puede inferir de la tabla 1, la lematización la hemos hecho con números. Corresponde el número 1 a las frases con SN desnudos (*ogi*Ø, *arno*Ø, *arraultze*Ø) y el número 2 a los SD con artículo (*ogia*, *arno*a, *arraultze*a**k**). El número de lemas lo podemos incrementar a placer, añadiendo más números. Nosotros, por ejemplo, si en un pueblo no hemos encontrado datos sobre el empleo del artículo en la oración que estamos trabajando (ya sea por que la frase tiene otra estructura, por que es una traducción más libre, etc.), le asignamos el lema 0. Cada lema se representará con un color diferente después en el mapa.

Es así como al final tendremos una tabla por cada oración que hayamos decidido investigar, que reúna los datos de todos los pueblos y los clasifique por lemas.

2.2. Los datos de Access a VDM

Aunque resulte sencillo realizar búsquedas concretas en las tablas de la base de datos, es decir, aunque es fácil consultar los datos de una oración concreta en un pueblo concreto (supongamos que queremos conocer el dato del pueblo de Mugerre en la frase 217: con mirar en la tabla 1, obtendríamos la respuesta), extraer conclusiones generales es más difícil, debido al volumen de datos que contiene cada tabla.

VDM se alimenta de estas tablas de Access y las muestra en mapas. Tomemos como ejemplo la tabla 1: obtendríamos el mapa referente a la oración 217 mediante VDM; esto es, el mapa llamado 217. El programa une el código correspondiente a cada localidad con el número de lema que le hayamos dado, asignando a cada lema un color. Por lo tanto, el pueblo de Lehuntze, cuyo número en el mapa es 1193, aparecerá en rojo mientras que el pueblo cuyo código es 1203 (Azkaine) aparecerá en azul, ya que éste tiene el lema 2, en tanto que a Lehuntze le corresponde el 1. Esto nos ofrece un mapa con tantos colores como queramos. De manera muy gráfica y con un simple vistazo, podremos saber, como mostraremos más tarde, si el empleo de SN desnudos o de SD presenta alguna coherencia geográfica.

2.3. Limitaciones del método y de VDM

Hemos descrito ya en §1.2.2 los problemas intrínsecos del corpus. Aparte de eso, hemos creído conveniente traer aquí, ya que estamos hablando sobre tecnologías para la variación lingüística, las limitaciones que nos plantea tanto el método como VDM, ante la esperanza de que algún día estén en situación de ser superadas.

Este método nos obliga a trabajar en un sistema binario; es decir, a presuponer que en una localidad obtendremos un sola respuesta y a asumir que esta será *a* o *b*, blanco o negro. Teniendo en cuenta las características del corpus, así ha sido en la mayoría de los casos, puesto que tenemos un sólo texto por pueblo y por tanto, un sólo ejemplo de la oración que buscamos. Pero hemos dicho ya al hablar del corpus, que en algunas localidades tenemos dos textos (§1.2.2). Cuando ambos textos dan la misma respuesta, no hay problema: si por ejemplo, en los dos textos, en la oración que estamos estudiando, nos aparece el SN desnudo, sin artículo, no cabe duda de que el lema que asignaremos a esa localidad será 1; por el contrario, si en ambos textos tenemos sintagmas determinados (SD) en esa frase, le asignaremos el lema 2 a esa localidad. Pero la casuística es más amplia. Tomemos un ejemplo:

- (5) a) 1368 (Donapauale): *OguenØ dut.* // 1269 (Donapauale): *OguenØ dut.*
 b) 1315 (Baigorri): *Hoben handia izan nîn (nièn).* // 1316 (Baigorri): *Hoben handia izan nien.*
 c) 1311 (Iholdi): *Oguén handia oukhan nicin.* // 1312 (Iholdi): *Oben haüin-diØ ukhan nuen.*

En los tres pueblos contamos con dos textos; dos respuestas para un solo pueblo. Evidentemente, (5a) y (5b) no son problemáticos, puesto que los resultados de los dos textos concuerdan; en (5a) ámbos son indeterminados, esto es, les corresponde el lema 1, y en (5b) ámbos son determinados, lematizados con 2. El conflicto viene con (5c), ya que los textos difieren: mientras en uno tenemos un SN desnudo, el otro tiene un SD. *VDM* acepta un sólo resultado por localidad, es decir, sólo le puede asignar un color a una localidad. La solución es la proliferación de más lemas: podemos crear tantos lemas como queramos, pero tan sólo podemos asignar uno por localidad. Si 1 es para los indeterminados (SN), 2 para los determinados (SD) y 0 para la falta de datos, 3 será para los datos divergentes: lo emplearemos cuando en localidades con dos textos, los datos de uno y de otro difieran. Esto solventa el problema parcialmente. Visualmente el problema está resuelto, puesto que el que vea el mapa observa claramente esa distribución mediante colores diferentes, pero *VDM*, a la hora de obtener estadísticas, tratará los lemas 1, 2 y 3 como respuestas diferentes, y eso no es estrictamente así: el lema 3, asignado a los datos divergentes, no es una respuesta diferente, sino aquella que toma en cuenta las anteriores dos (con el lema 1 y 2); es decir, el lema 3 no es *c*, sino *a+b*. Pero, como decíamos, para *VDM* es estadísticamente igual la ausencia de datos (lema 0), el tener una sola respuesta (lemas 1 y 2), tener ambas respuestas (lema 3) o cualquier otro tipo de respuesta (lemas 4, 5, 6, etc.). Es una característica del método cuantitativo: lo que cualitativamente es diferente, cualitativamente se mide igual.

Otro problema del sistema cualitativo binario es la imposibilidad de juntar resultados cruzados. Expliquémonos. Para investigar el nivel de utilización del artículo (núcleo D del sintagma), hemos cogido como referencia oraciones de un corpus. Hemos hecho un mapa de cada oración. ¿Cómo comparar entre sí todos esos mapas? No es difícil crear en la base de datos una tabla que una las respuestas de todos los mapas:

Tabla 2

Los resultados de todos los mapas por frases

_OrtNr	KARTIEL	DIRU	GASNA	OGEN	AHATE	NAGUSI	ERO
1193	2	2	2	2	7	1	2
1194	1	2	2	2	2	1	2
1196	0	0	1	2	8	1	0
1198	1	1	2	6	1	1	1
1201	2	1	1	2	3	1	2
1203	0	2	2	6	2	1	0
1204	2	2	2	0	2	1	0
1205	2	1	1	2	7	1	2
1206	1	1	1	2	7	4	0
1207	2	2	2	1	0	1	0
1208	0	0	2	0	0	0	0
1209	2	2	2	2	2	1	2
1210	2	2	2	2	8	1	2

En primer lugar tenemos el código de cada pueblo y después los datos de siete tablas diferentes (*Kartiel*, *Diru*, *Gasna*, *Ahate*, *Nagusi*, *Ero* y *Ogen*). Cada una de las siete corresponde a una oración y su mapa. En esta tabla hemos juntado todas y podemos ver qué ha respondido cada localidad sobre cada frase. El pueblo 1194, Mugerre, nos ha dado los siguientes resultados:

- (6) a) KARTIEL: *Behar da (...) ikus deçadan herriØ*.
 b) DIRU: *Dembora da (...) içan deçadan dirua*.
 c) GASNA: *bethea da sehiz çoinec baituté ogui**a** eta arno**a**, arroltze**ac** eta gasna*.
 d) OGEN: *Hutz handi**a** eguin nuen*.
 e) AHATE: *Hartcen ahalco ditutçué éré oilarr**ac**, ahate**ac** eta ekartcen aratché on bat hiltçeko*.
 f) NAGUSI: *Dembora da içan nadin éné nausiØ*.
 g) ERO: *Soro**a** çare*.

Tal y como muestran los datos de la tabla 2 y de (6), mientras unas oraciones aparecen con el determinante (*Diru*, *Gasna*, *Ahate*, *Ero*, *Ogen*), otras presentan un SN desnudo (*Kartiel*, *Nagusi*). En otras localidades, por supuesto, los resultados serán otros.

¿Cómo hacer un mapa que tenga en cuenta los siete mapas? En la tabla 2 tenemos 7 lemas por localidad y recordemos que *VDM* sólo acepta un lema por pueblo. Las combinaciones posibles entre esos 7 lemas son demasiadas (no hay más que observar la tabla 2, para percatarse de que en ningún pueblo se da la misma combinación de lemas): asignar un lema (un color en el mapa) a cada combinación posible, haría el mapa ilegible; incoherente y lleno de colores.

La solución a este problema parte de fijarse, más que en la cantidad, en la cualidad. Todos los SN desnudos no son iguales; tampoco los SD. La presencia del artículo cambia la semántica y los SN desnudos son sólo posibles en contextos sintácticos concretos. La labor del lingüista es clasificar los SN y los SD recopilados en el corpus, en función de su interpretación semántica y sintáctica. Tal y como mostraremos luego, realizar comparaciones cruzadas entre mapas será más sencillo tras hacer la clasificación.

Para darse cuenta de la importancia de la función sintáctica y semántica, veamos de manera breve lo dicho por algunos lingüistas sobre el empleo del artículo en vasco. Aprovecharemos la síntesis de sus teorías para presentar los mapas correspondientes a cada oración, mostrando si avalan o contradicen esas teorías.

3. Lo dicho por los lingüistas y nuestros mapas

Ya hemos añadido (§1.1) que se han escrito unas cuantas líneas sobre el artículo en vasco. Los sincronistas han hecho un gran esfuerzo para fijar el contexto de empleo y aparición del artículo y los diacronistas han perseguido el objetivo de saber desde cuándo se encuentra en vasco y de qué manera. Esta cita de Michelena servirá para introducirnos en el meollo del asunto:

Es un lugar común de la lingüística histórica vasca la afirmación de que el artículo determinado es entre nosotros de introducción relativamente reciente, como lo es en las lenguas románicas o germánicas. Pero aquí, lo mismo que mucho antes en griego, su aparición es un hecho documentado, mientras que en vasco se trata de una presunción (...) referente a la prehistoria de la lengua. Lo que sí es un hecho plenamente histórico es su proliferación: durante los últimos siglos ha ido y va ganando posiciones en las cuales su empleo tiende a hacerse obligatorio. (Michelena 1978: 208)

Por lo tanto, es imaginable que en la evolución, partiendo de la ausencia absoluta del artículo al paradigma actual, el vasco ha conocido diferentes fases cronológicas como geográficas; esto es, no hallaremos el mismo empleo del artículo en un lugar y época, o en otra. Asimismo, aparecerá en unas estructuras sintácticas antes que en otras. Mientras los diacronistas han tratado de concretar esa cronología, los sincronistas persiguen fijar la distribución sintáctica y geográfica.

Los que han hecho frente al problema, sea desde el punto de vista sincrónico como desde el diacrónico, han dado por buena la separación de los sintagmas en dos grupos, dependiendo de su función. Por un lado estarían los predicados, que cumplen una función atributiva, con un verbo copulativo. Por el otro, tendríamos sintagmas que satisfacen un argumento, puesto que un verbo les asigna un rol temático.

3.1. Sobre los sintagmas con función atributiva

En el trabajo Azkarate & Altuna (2001) se nos dice que los atributos que tengan un verbo auxiliar de raíz *izan* ‘ser’, serán indeterminados si son adjetivos simples; si tienen una estructura de *nombre + adjetivo*, aparecerán definidos con el artículo (Azkarate & Altuna 2001: 74-75). De todos modos, en las hablas del oeste (Bizkaia, Gipuzkoa, etc.), también los adjetivos simples aparecerán principalmente determinados. En las hablas del este, por tanto, la presencia de un nombre obligaría a la presencia del artículo. Sirvan como ejemplo estas frases de nuestra cosecha:

- (7) a) *Mikel aberatsØ da*. ‘Mikel es rico’ (Adjetivo simple, que en el oeste aparecería también con artículo)
 b) *Mikel gizon aberatsa da*. ‘Mikel es un hombre rico’ (estructura *nombre + adjetivo*)

La unión entre la presencia del nombre y el artículo fue expresada ya por Michelena (1978: 213): “...el adjetivo, como tal, queda indeterminado, mientras que la determinación es normal cuando el predicado nominal está formado por un sintagma sustantivo + adjetivo” y un poco más adelante: “...permitiría (...) atribuir la determinación a los sustantivos”. Zabala, por el contrario, da a los nombres o adjetivos simples la opción de ir indeterminados (desnudos), y añade que sólo las estructuras *nombre + adjetivo* tienen obligación de portar artículo (Zabala 2001: 330).

La semántica del predicado también tiene su importancia, como ya han destacado otros trabajos. Zabala mismo nos muestra que lo que Txillardegi o Lafitte definieron como *transitivo / permanente* y que ella define como *individual / stage level*, tiene mucho que decir sobre la presencia o ausencia del artículo (Zabala 2001: 329). La cualidad de los predicados *individual level* es perpetua o intransitoria y en el caso de los *stage level* esa cualidad tiene una limitación temporal; no es para siempre:

- (8) a) *Mikel irakaslea da*. ‘Mikel es profesor’ *Individual level*
 b) *Mikel irakasleØ da(go)*. ‘Mikel está de profesor’ *Stage level*

Como se puede observar, los predicados de tipo *individual level* (8a) se forman con el auxiliar *izan* ‘ser’; los de tipo *stage level*, por el contrario, pueden formarse en vasco con los auxiliares *izan* ‘ser’ o *egon* ‘estar’: la elección de uno u otro depende del dialecto. Los dialectos que emplean *izan* para ambos casos (8 a y b), hacen la diferenciación semántica empleando la ausencia o presencia del artículo. Eguren (s. d.) también utilizará esta diferenciación para demostrar que el morfema *-a* no es el artículo, sino el núcleo de un sintagma predicativo. En todo caso, no es éste el lugar para profundizar en la hipótesis de Eguren.

Resumiendo; la necesidad de un SD o de un SN desnudo se ha asociado a la interpretación semántica, pero en algunas hablas, en palabras de Zabala en las del este (Zabala 2001: 330), en esas mismas que permiten que los sintagmas de interpretación *individual level* puedan ser SN desnudos, la presencia del determinante artículo iría asociada a la presencia del nombre.

Vista la distribución sincrónica, fijémonos en la descripción realizada desde el punto de vista diacrónico. Manterola nos dice claramente que en su opinión, en la fase antigua del vasco no había artículo (cf. la cita de Michelena §3) y que por tanto,

a día de hoy los dialectos del este son los que mayor vínculo conservan con dicha fase antigua (Manterola 2008: 1).

En los trabajos más recientes, como el de Manterola, se han hecho matizaciones a la distribución anterior. Zabala nos decía que en las hablas del este había sintagmas desnudos de carácter *individual level*; Manterola nos confirma que, además de con nombres, con adjetivos simples a veces tampoco son posibles los SN desnudos. Los ejemplos del dialecto bajonavarro son suyos (Manterola 2008: 7).

- (9) a) *Jestu hori / jestu horren⁵ egitea ez da pollitØ*. ‘Hacer ese gesto no es bonito’
 b) **Ume hori ez da politØ vs. Ume hori ez da pollita*. ‘Ese niño no es bonito’

Ambos sintagmas (9a, b) son adjetivos de tipo *individual level*, pero el segundo necesita del artículo en bajonavarro (la presencia del nombre no condiciona nada en este caso, ya que no lo hay). Para justificar esto, Manterola nos recuerda el trabajo de Milsark, que, tras aplicarlo al vasco, le lleva a pensar en una posible relación entre el sujeto y el predicado de la oración: en los ejemplos de (9), alguna propiedad del sujeto sería la que condicione que mientras en (9a) el SN desnudo es posible, no lo sea en (9b). ¿Cuál es, pues, el motivo de que en bajonavarro haya entrado el artículo antes en oraciones del tipo (9b)? que el sujeto déictico de (9b) *ume hori* ‘ese niño’ tendría una “carga referencial” mayor: sería más tangible que el de (9a). El niño tendría más propiedades de sujeto, de agente, que el gesto (Manterola 2008: 8-9).⁶

En nuestro corpus tenemos algún ejemplo de predicado atributivo de raíz *izan* ‘ser’.

En el mapa de la página siguiente tenemos los resultados de la traducción de la oración francesa *Etes-vous fou?*⁷ He aquí sendos ejemplos de las posibles traducciones que hemos recogido en el mapa:

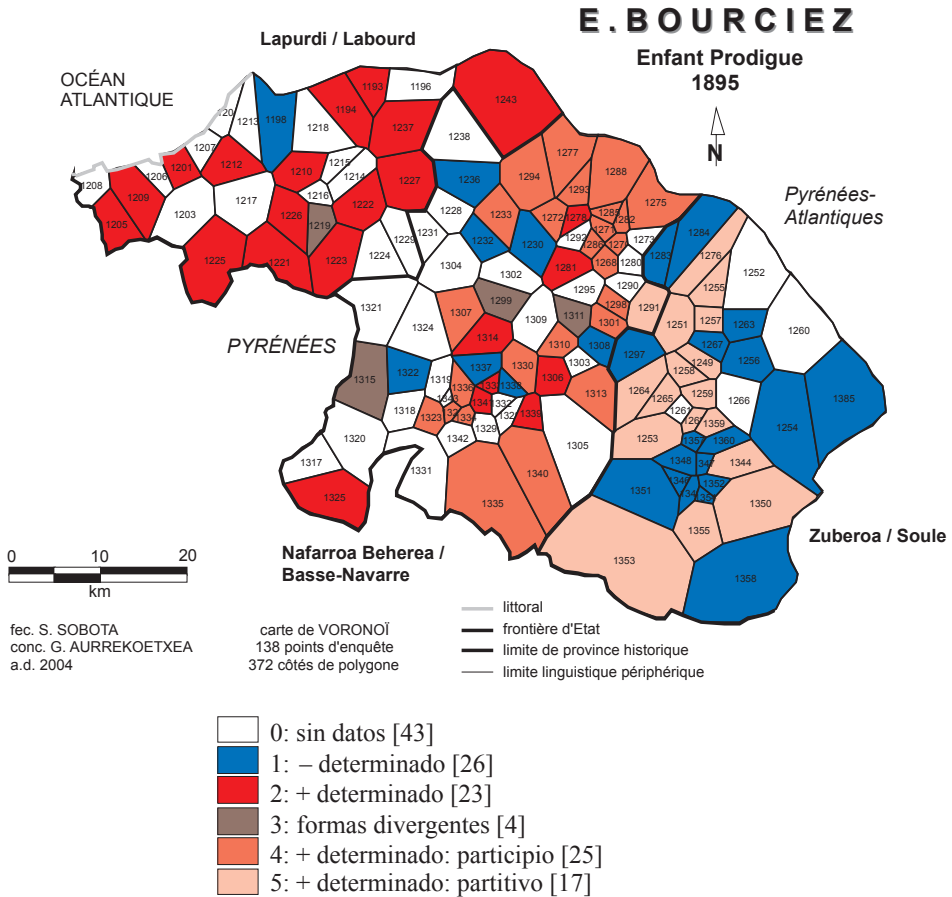
- (10) a) *EroØ zara?*⁸ – determinado
 b) *Eroa zara?* + determinado
 c) *Erotua zara?* + determinado: participio
 d) *Eroturik zara?* + determinado: partitivo

⁵ En las hablas del este el objeto suele aparecer con caso genitivo (*gestu horREN*) en lugar de absoluto (*gestu horI*), cuando el verbo está nominalizado (*egitea* ‘hacer’). Es por eso que Manterola da las dos formas, pero en principio no tiene que ver con el tema que tratamos.

⁶ El propio Manterola reconoce que la argumentación para clasificar un sintagma como fuerte o débil conforme a su carga referencial, no está todavía demasiado clara. Alguien podría pensar que, en lugar de tener en cuenta la carga referencial, es la estructura sintáctica la que condiciona la presencia del artículo, ya que mientras en (9a) el sujeto es una oración (tiene verbo), en (9b) nos encontramos con un sintagma determinante cumpliendo dicha función. En todo caso, y reiterando que ésta es una hipótesis no contrastada, Manterola sospecha que una oración como **Jestu hori ez da pollit* tampoco sería posible a pesar de contar con un SD como sujeto, puesto que éste sería un sintagma débil; de todos modos, este ejemplo que damos en la nota al pie, a diferencia de los de (9), no ha podido ser escuchado por Manterola a hablantes en conversación natural, por lo tanto, es un asunto que queda por dilucidar.

⁷ En Aurrekoetxea & Videgain (2004: 421) viene *Etes-vous devenu fou?* en la transcripción, pero es una errata. Cf. La versión corregida del texto en Santazilia (2007).

⁸ Los ejemplos correspondientes a los mapas los hemos adaptado y estandarizado dejando de lado variantes dialectales, para mostrar de manera más clara el asunto que nos concierne: el de la utilización del artículo. Si alguien quisiera ver la oración original, no tiene más que acudir a Aurrekoetxea & Videgain (2004).



Mapa 1

Predicado atributivo de raíz *izan* 'ser'

El retrato de finales del siglo XIX que nos ofrece el mapa es esclarecedor. En Lapurdi son mayoría las formas determinadas por el artículo (10b) y en Zuberoa las indeterminadas (10a); es más, en esta región no encontramos ninguna forma indeterminada. Además de eso, en este mapa tenemos dos estructuras más; las de las oraciones de (10c) y (10d). La estructura con participio (-tu) lleva artículo; por otro lado, si tomamos el partitivo como determinante, ambas estructuras son determinadas. En el siglo XIX las formas sin artículo (indeterminadas) aparecen con más profusión en Zuberoa. Las estructuras con partitivo y participio también muestran una distribución clara: las primeras las hallamos sólo en Zuberoa y las segundas corresponden sólo a la Baja Navarra. Estas últimas, llevarían el artículo, según Michelena, puesto que si no lo llevaran se podrían confundir con una estructura de *verbo principal + auxiliar*: *Erotu zara* 'has enloquecido' (Michelena 1978: 215).

Siguiendo la pauta de Michelena, podríamos pensar que el sintagma de (10a) es un adjetivo y el de (10b) un nombre; es decir, que mientras el primero considera *erho* ‘loco’ como adjetivo, el segundo lo interpreta como nombre (sustantivo). Si no, mirando al original en francés, podríamos decir que mientras en algunas traducciones le han dado una interpretación como *individual level*, en otras localidades lo han interpretado como *stage level*.⁹ Pero, si miramos el mapa, otro tipo de ideas nos vienen a la cabeza, por la coherencia geográfica que muestra. Si fuera un mero caso de interpretación semántica, no podríamos diferenciar isoglosas claras. Es más económico pensar que para entonces en Lapurdi se habría extendido la tendencia que era ya habitual en las hablas al sur del Pirineo: que el artículo aparezca con todo tipo de atributos. De facto, Michelena mismo reconoce que “se diría que [el artículo] es empleado más profusamente en el centro y oeste del país que en el este, más en general al sur que al norte” (Michelena 1978: 208). Nos parece lo más apropiado pensar que la oración ha sido en todos los lugares interpretada como *stage level*, puesto que las formas con participio (*erotua*) de Baja Navarra y las formas con partitivo (*eroturik*) no admiten más que una interpretación transitoria (*stage*). Siguiendo la opinión de Manterola, podríamos justificar las pocas formas determinadas por el artículo que encontramos en Baja Navarra, aduciendo a una posible interpretación *individual level* por parte de los traductores, en el que se han visto obligados a emplear el artículo, al tener un sujeto con carga referencial grande (*strong*).

De todos modos, Michelena admite que el empleo de ciertos predicados indeterminados es común a la totalidad del vasco (Michelena 1978: 210). Tenemos un ejemplo (en la página siguiente, Mapa 2) de esto en nuestro corpus.

Son las siguientes, las oraciones representadas aquí:

(11) *Il est temps que je sois mon maître.*

- a) *Denbora da izan nadin nere buruaren nagusi*Ø. – determinado
- b) *Denbora da izan nadin ene nagusia*. + determinado
- c) *Ordu da ene buruaz nagusi*Ø *izan nadin*. instrumental

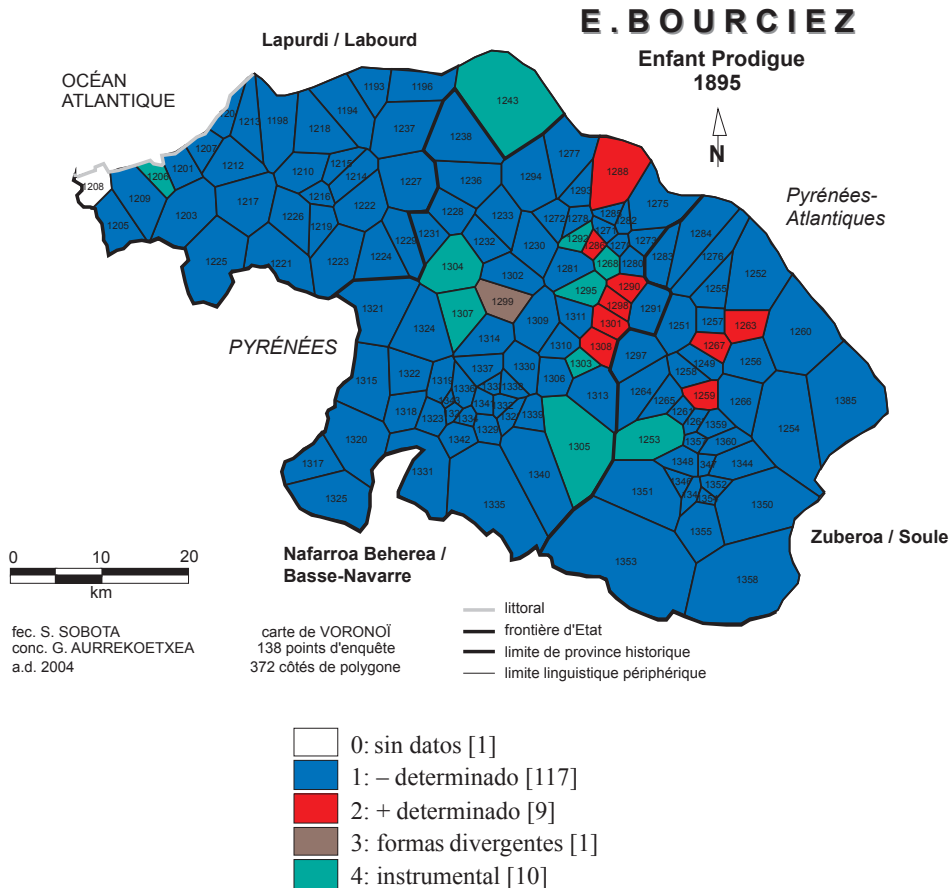
¿Por qué es en este caso general el empleo de la forma desnuda, también en las hablas del sur del Pirineo? A nuestro parecer en este tipo de oración sólo es posible una interpretación *individual level*, por tanto, sería lo habitual que apareciera el artículo. Es más, si asumimos que la presencia de un nombre condiciona la aparición del artículo, podría pensarse que *nagusi* ‘dueño’ es un adjetivo, pero al aparecer con un genitivo (*ene buruaREN* ‘de mí mismo’), queda asegurado su carácter de nombre, puesto que los genitivos sólo pueden aparecer con éstos.

Podríamos tomarlo como una estructura lexicalizada, si tenemos en cuenta que en castellano también es posible esa estructura sin artículo:

(12) a) *Soy* Ø *dueño de mis actos.* a') *Soy el dueño de mis actos.*

Podría proponerse que tanto en castellano como en vasco, la ausencia de artículo sugiriera un cambio semántico progresivo hacia un *stage level*, que transformaría la cualidad de propiedad en algo transitivo, a pesar de formarse con el verbo *izan* ‘ser’

⁹ Recordemos que en francés ambas interpretaciones se realizan con el verbo *être* ‘ser/estar’.



Mapa 2

Ejemplo de profusión de indeterminados

(y no con *egon* 'estar'). No seremos nosotros quienes aportemos la solución definitiva al asunto, pero nos parece que puede abrirse una vía de investigación en torno a la estructura semántica de *nagusi* 'dueño'. Diríamos que es un nombre con estructura argumental y por tanto, tanto los sintagmas con genitivo como los de instrumental son necesarios para que la frase satisfaga dicha estructura argumental. Haremos una anotación a modo de anexo, para marcar la importancia de esos complementos en genitivo e instrumental: los pueblos que han utilizado formas determinadas (9 pueblos) no han utilizado como complemento una forma reflexiva (*nerre burua* 'mí mismo'), sino un pronombre simple (*nerre* 'mi'); pero sí hay quienes aunque no hayan usado la forma reflexiva, han empleado la forma indeterminada, sin artículo:

- (13) a) *Denbora da izan nadin enel/neure nagusi*Ø.
a') *Denbora da izan nadin enel/neure nagusia*.

- b) *Denbora da izan nadin ene buruaren/neure buruaren nagusi*∅.
 b') ***Denbora da izan nadin ene buruaren/neure buruaren nagusia*.

Sin duda, este asunto merecería un análisis más profundo.

3.2. Sobre los sintagmas de función complemento

Las demás oraciones extraídas del corpus y sus mapas corresponden a sintagmas de función complemento, es decir, son argumentos de un verbo y cumplen la función de complemento directo. Hemos elegido ejemplos de esa función sintáctica, puesto que no hemos encontrado en el corpus sintagmas argumentales con otra función, contruidos mediante SN desnudos. En el vasco estándar de hoy es imposible encontrar sintagmas argumentales sin artículo (Artiagoitia 2004: 27-28, Zabala 2001: 328), pero en el vasco histórico mismo y en otras lenguas tenemos ejemplos: fijémonos en la oración inglesa *Dogs are dangerous*, donde el sujeto *Dogs* no lleva artículo; o veamos estos ejemplos extraídos por Michelena del texto de Axular, del siglo xvii (Michelena 1978: 211-212):

- (14) a) *Aita saindu*∅-ri. Complemento indirecto, con marca de dativo (-ri) y sin artículo (∅).
 b) *Ainguru*∅-c. Sujeto, con marca de ergativo (-c) y sin artículo (∅).

Como decíamos, en el texto de Bourciez no hay ni rastro de SN desnudos en función de sujeto ni de objeto indirecto, pero sí que hay, aunque no sean posibles en el estándar de hoy, SN desnudos con función de objeto directo: es en éstos en los que nos fijaremos.

Asumiendo, pues, que los dialectos orientales son los más arcaizantes en este aspecto y que el artículo, además de ser tardío, se ha extendido de oeste a este (Manterola 2006: 11), tendríamos la hipotética esperanza de encontrar el mayor número de objetos desnudos en el este. Desde el punto de vista diacrónico, siguiendo a Manterola, los nombres incontables son tipológicamente los últimos en recibir el artículo; primero lo adquieren los sintagmas de interpretación plural indeterminada (Manterola 2006: 7):

- (15) a) *Nik ardoa erosi dut.* a') *Nik ardo*∅ *erosi dut.* 'Yo he comprado vino'
 b) *Nik arraultzeak erosi ditut.* b') *Nik arraultze*∅ *erosi dut.*¹⁰ 'Yo he comprado huevos'

Los objetos de las primeras oraciones (15a y 15a') son conceptos de masa; incontables. Los otros dos son contables, pero tienen interpretación indeterminada, genérica: no sabemos a qué huevos hacemos referencia, ni necesitamos precisarlo. En castellano también son indeterminados estos sintagmas: *He comprado* ∅ *huevos*. En los dialectos conservadores del este, a día de hoy todavía esperaríamos ejemplos como los de (15a') y (15b'). En el trabajo de Urtzi Etxeberria y Rikardo Etxepare, limitan la existencia de SN desnudos al dialecto suletino (de Zuberoa)

¹⁰ Aunque no entraremos en este tema, es interesante en el caso de (15b) y (15b'), cómo la presencia o ausencia del artículo condiciona la concordancia verbal de objeto, siendo plural en (15b) (*ditut*) y singular en (15b') (*dout*). Hay algo sobre este asunto en el trabajo Etxeberria & Etxepare (2008).

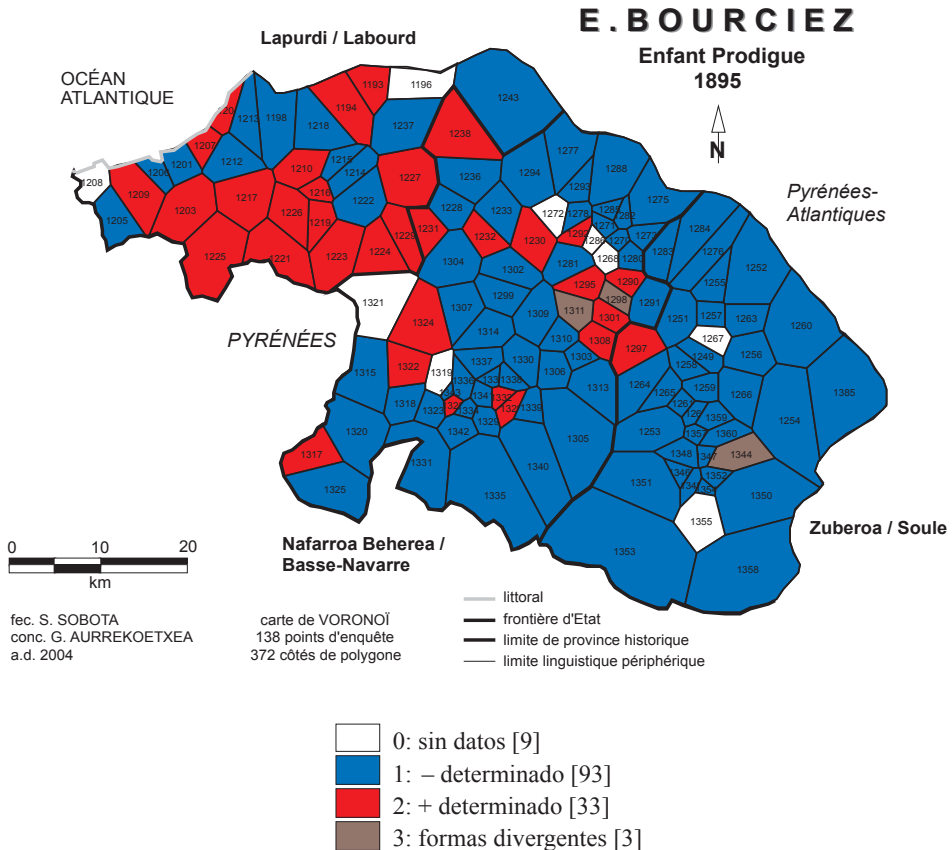
(Etxeberria & Etxepare 2008).¹¹ Cuando mostremos los mapas, trataremos de fijar las fronteras de ese “este” en el siglo XIX, atendiendo a la oposición recién descrita de *contable / incontable*. A su vez, dentro de los incontables separaremos los sintagmas con interpretación existencial (genérica) y específica.

3.2.1. Los sintagmas incontables

En el mapa de debajo veremos la distribución de las siguientes oraciones:

(16) *Il est temps (...) que j'aie de l'argent.*

- a) *Denbora da (...) izan dezadan diruØ.* – determinado
- b) *Denbora da (...) izan dezadan dirua.* + determinado



Mapa 3

Concepto de masa

¹¹ De todos modos, el objetivo de Etxeberria & Etxepare en ese trabajo no es establecer isoglosas claras, por lo tanto, no se tome *dialecto suletino*, demasiado *sensu stricto*.

El dinero es un concepto de masa; no se puede contar (**tres dineros*). Como vemos en el mapa, el empleo de SN desnudos con conceptos de masa está bastante extendido en el corpus de Bourciez. Las formas determinadas aparecen principalmente en el territorio del labortano, aunque también tenemos ejemplos en bajonavarro. Teniendo en cuenta que Manterola y Etxepare & Etxeberria circunscriben ese fenómeno a día de hoy a los dialectos orientales exclusivamente, parece ser que el artículo habría entrado en estos contextos en época muy reciente, debido a la rápida evolución que ha tenido desde finales del siglo XIX hasta hoy, si comparamos el mapa con la distribución actual.

3.2.2. *Los sintagmas contables*

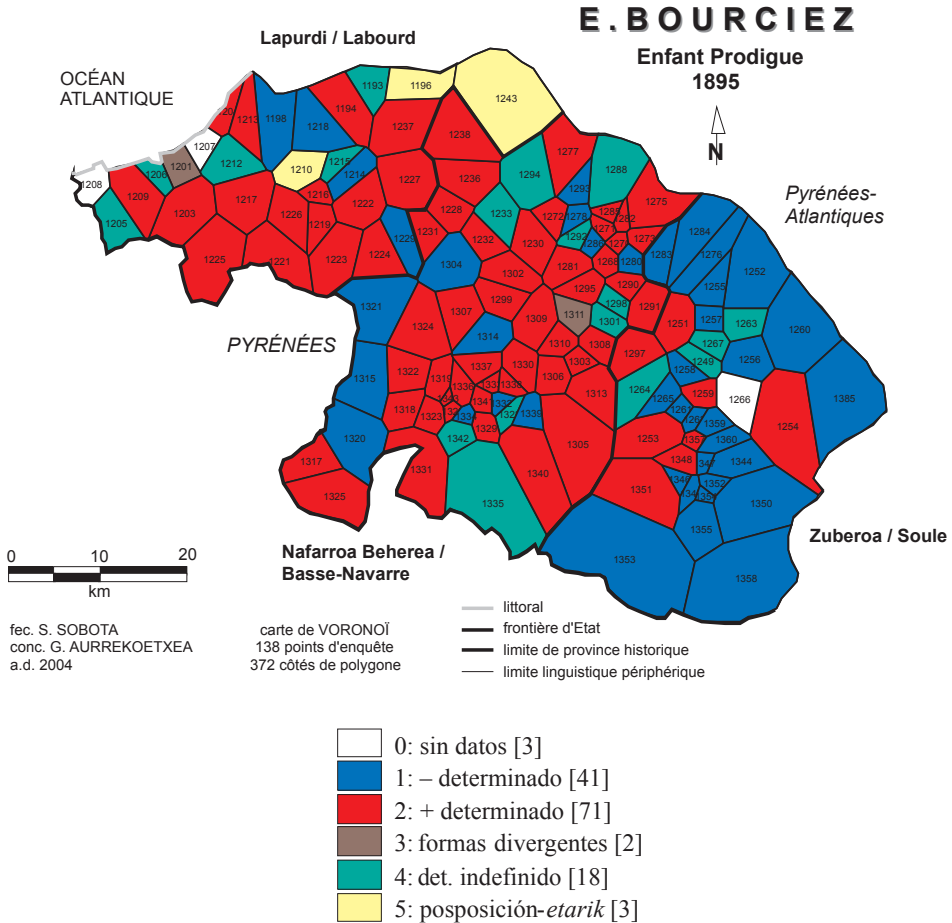
Tenemos oraciones como éstas:

(17) *Vous pourrez aussi prendre des coqs, des canards, et amener un veau (...).*

- a) *Hartzen ahal duzue oilarØ, ahateØ, eta ekartzen txahal bat (...).*
– determinado
- b) *Hartzen ahal dituzue oilarrak, ahateak, eta ekartzen txahal bat (...).*
+ determinado
- c) *Hartzen ahal dituzue oilar eta ahate zenbait, eta ekartzen txahal bat (...).*
det. indefinido
- d) *Hartzen ahal dituzue oilar eta ahateetarik, eta ekartzen txahal bat (...).*
posposición *-etarik*

Las oraciones de (17), por lo tanto, tienen sintagmas con conceptos contables, pero mediante la falta de artículo expresan el carácter genérico del sintagma; nos da igual qué pato o gallo coger. Las oraciones (17c) y (17d) son claros ejemplos de esta interpretación genérica, ya que el determinante indefinido (*zenbait* ‘algunos’) y la posposición *-etarik* no permiten otra interpretación.

Según muestra el mapa 4, mientras que las formas indefinidas son más abundantes en Zuberoa, encontramos las determinadas en Lapurdi y la Baja Navarra principalmente. Lo más interesante es comparar este mapa con el mapa 3: a finales del siglo XIX es bastante más común la presencia de SN desnudos con conceptos incontables que con contables, ya que con estos últimos no encontramos sintagmas indeterminados de manera general, más que en Zuberoa. Que la jerarquía de Manterola mencionada arriba (§3.2) se cumple en estos mapas, es evidente: han tomado antes el artículo los nombres contables (por eso está más extendido el uso del artículo con éstos), a pesar de tener una interpretación genérica, que los conceptos de masa.



Mapa 4

Sintagmas contables genéricos

3.2.3. Los sintagmas contables e incontables

Nos ha parecido interesante hacer el mapa de una oración que presenta a la vez nombres contables e incontables.

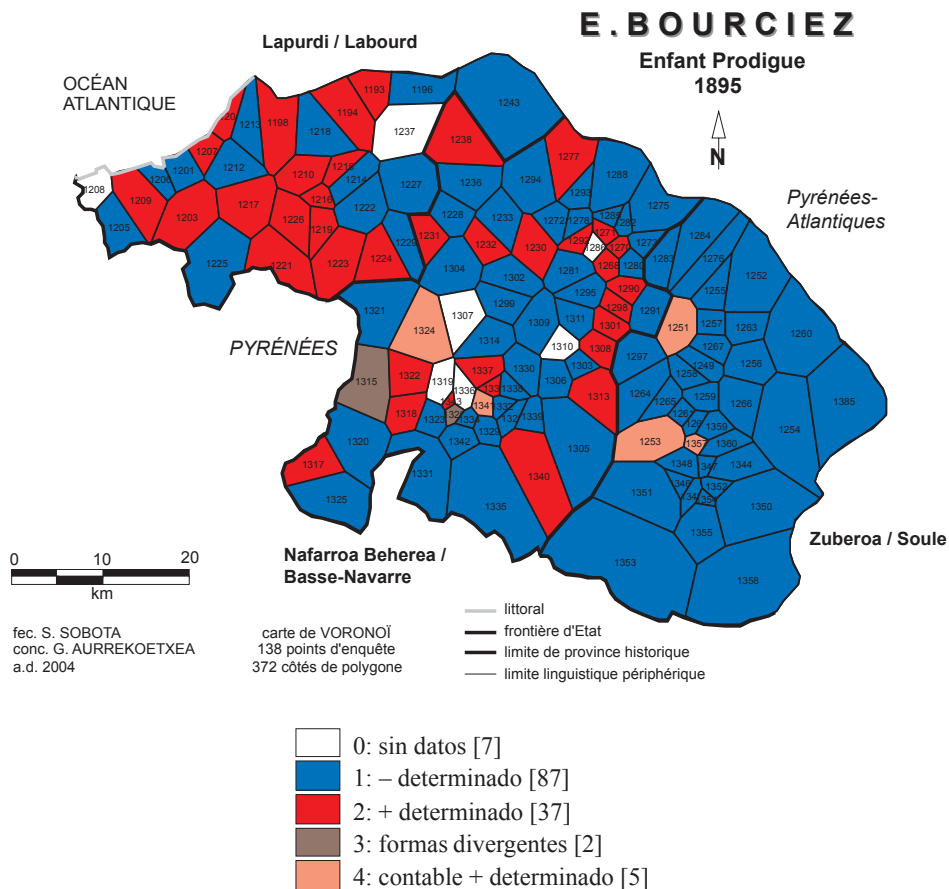
Como veremos a continuación en los ejemplos, tres son las combinaciones posibles:

(18) *[La maison de mon père] est pleine de domestiques qui ont du pain et du vin, des oeufs et du fromage.*

a) *Betea da mutilez, zeinek baitute ogia eta arnoa, arraultzeak eta gasna.*
+ determinado

- b) *Betea da mutilez, zeinek baitute ogiØ eta arnoØ, arraultzeØ eta gasna.* – determinado
- c) *Betea da mutilez, zeinek baitute ogiØ eta arnoØ, arraultzeak eta gasna.* contable + determinado

Y he aquí el mapa que hemos obtenido:



Mapa 5

Conceptos contables e incontables

En estos ejemplos de (18), los dos primeros sintagmas (*ogia* ‘pan’, *arnea* ‘vino’) son incontables y el tercero (*arraultzeak* ‘huevos’) contable.¹² Las opciones son tres: todos

¹² Hemos dejado de lado el sintagma *gasna* ‘queso’, porque la -a final forma parte de la raíz y no cambia, sea determinado o no: *gasna* + -a = *gasna*; *gasna* + -Ø = *gasna*. Por lo tanto, no hay manera de saber si el ejemplo de la oración lleva artículo o no.

los sintagmas con artículo (18a); todos indeterminados (18b) o todos indeterminados salvo los contables (18c). En teoría podríamos esperar una cuarta opción: que aparecieran los incontables con artículo y los contables sin él, pero no hemos hallado en la práctica ningún ejemplo de eso, por lo que la oración siguiente debemos darla con asterisco:

(19) a) **Betea da mutilez, zeinek baitute ogia eta arnoa, arraultzeØ eta gasna.*

Pero no es casualidad que estos casos teóricamente posibles no aparezcan, y un hallazgo tal, no hace sino fortalecer la hipótesis defendida junto a Manterola en este trabajo: si los nombres incontables han adquirido el artículo diacrónicamente después de los contables, podemos diferenciar tres épocas o situaciones de la lengua:

- (20) a) Los contables y los incontables son indeterminados.
 b) Mientras los contables son determinados, los incontables permanecen indeterminados.
 c) Tanto los contables como los incontables aparecen con el artículo.

Si quisiéramos representar esto en una tabla tetracórica (Croft 1993: 48), tendríamos algo así:

Tabla 3

Diagrama tetracórico de los sintagmas contables e incontables

	Contables e incontables	Sólo incontables
+ determinado	✓	✗
- determinado	✓	✓

Por lo tanto, esta tabla nos descubriría una relación de implicación: si una lengua tiene nombres incontables determinados, entonces tendrá también nombres contables determinados.

El mapa nos atestigua las tres situaciones descritas en (20): no puede existir, en nuestra opinión, ni un lugar ni una época en la que los sintagmas contables fueran indeterminados y los incontables determinados, tal y como confirma el mapa.

De facto, se ve perfectamente en el mapa 5 que el este es más conservador. Tanto los nombres contables como los incontables con artículo los encontraremos de manera más frecuente en el territorio del labortano y también en el del bajonavarro tenemos unos cuantos ejemplos: sin embargo, no hay ni uno en Zuberoa, donde todos son indeterminados. Conforme vayamos al oeste, nos irá apareciendo el artículo de manera progresiva.

Son pocos los ejemplos de (20b), una situación de transición entre sistemas: no son representativos geográficamente, pero sí tipológicamente, puesto que podemos estar seguros de que si en esos pueblos que están entre el sistema indeterminado completo y el determinado completo, ha sucedido algún cambio, ha sido a hacer todas las formas con el artículo, y no a potenciar las formas indeterminadas; esto es, el camino será de (20b) a (20c), nunca a (20a).

Es interesante observar el asunto desde el prisma de lo marcado/no marcado. Ese concepto de *markedness* cuya entrada en vigor puede ser atribuida a los neogramáticos, ha conocido a posteriori una infinidad de empleos imprecisos y ambiguos que han llevado a devaluar el concepto como término científico. No obstante, Andersen recupera ese valor de *markedness* como un concepto universal, para crear el *Principle of Markedness Agreement*, según el cual los elementos marcados intervendrán antes en entornos marcados, mientras que los no marcados lo harán en los no marcados (Andersen 2001a). Se considerará marcado aquello que el hablante no tenga asimilado en su gramática base, y cuya utilización conozca por reglas de empleo superficiales.

Si partimos de la hipótesis de que el artículo es de incorporación reciente en vasco, su ausencia sería, al menos históricamente, el suceso no marcado. Debemos pensar que en determinado momento se introduce el artículo en vasco, un morfema marcado y posiblemente apre(he)ndido,¹³ inicialmente como cuantificador en sintagmas contables, produciéndose una alternancia sincrónica. El contar con este morfema para cuantificar sintagmas contables haría finalmente obligatoria su presencia en sintagmas contables (más fácilmente cuantificables que los incontables), considerando marcados los contextos en los que tenemos sintagmas contables sin cuantificar (es decir, sin artículo), ya que necesitarán de ese morfema marcado, cumpliendo así el *Principle of Markedness Agreement*: el morfema marcado aparecerá primero en un contexto que también lo es, para luego extenderse a los contextos no marcados (Andersen 2001: 31). El mapa mostraría esas tres fases: en Zuberoa el morfema marcado apenas habría comenzado a aparecer, salvo en unos pocos ejemplos donde lo encontramos en los contextos marcados. Conforme nos aproximamos a Lapurdi, el morfema marcado (el artículo) se habría extendido ya incluso a los contextos no marcados, que serían los de los sintagmas incontables.

En resumidas cuentas y asumiendo la cronología sobre la direccionalidad del cambio evolutivo propuesta por Andersen (Andersen 2001b), los hablantes del vasco labortano habrían comenzado a emplear el artículo para cuantificar sintagmas contables, al principio esporádicamente y en alternancia con la forma tradicional indeterminada. A medida que el uso de ese morfema marcado se va extendiendo, se considera el SN contable desnudo como contexto marcado, y se generaliza el empleo del artículo en él. A posteriori, serán los contextos no marcados los que tomarán el artículo. Y este proceso irá sucediendo de igual manera de oeste a este. Así las cosas, se producirá al final un *markedness shift* (Andersen 2001b: 238) y lo que en principio era marcado y sucedía en la gramática superficial del hablante, provocará un reanálisis y una actualización de su gramática base, pasando a ser considerado no marcado, en detrimento de la forma antigua (sin artículo) que será considerada marcada (por arcaica). Este *markedness shift* sólo se producirá cuando la nueva forma haya sido asimilada completamente por la gramática base del hablante.

El labortano presentaría así una situación más evolucionada donde el artículo sería ya un morfema no marcado con el *markedness shift* ya concluido y extendido a todos los contextos, mientras que de camino a Zuberoa encontraríamos fases más arcaicas donde el empleo del artículo no estaría aún completamente gramatical-

¹³ No entraremos aquí a discutir si se trata de un préstamo del romance o no.

zado. Esto demuestra, obviamente, que el artículo en vasco se ha ido extendiendo de oeste a este.

3.2.4. *Un caso particular: el sintagma de nombre + adjetivo*

El corpus siempre es escaso para el investigador; nunca le dará tantos ejemplos como sean necesarios para obtener conclusiones completas. De todos modos, en el texto de Bourciez hemos encontrado unas traducciones que pueden ser un buen tema sobre el cual hablar:

(21) *J'eus grand tort.*

- a) *Ogen handia ukan nuen.* + determinado
- b) *Ogen handiØ ukan nuen.* – determinado
- c) *Anitz ogen ukan nuen.* con adverbio

La particularidad de este sintagma reside en la presencia del adjetivo por un lado y en las características del argumento por otro. Manterola muestra, contra la creencia de Michelena, que hay estructuras *nombre + adjetivo* predicativas indeterminadas en los textos antiguos, como el de Etxepare en el siglo xvi (Manterola 2008: 18), pero, ¿y a modo de complemento en el corpus de Bourciez?

Veamos el mapa 6 de la página siguiente.

Si tomamos el sintagma *ogen handi(a)* ‘gran culpa’ como complemento del verbo *ukan* ‘tener’, lo deberíamos tratar como concepto incontable. Aunque el adjetivo *handi* ‘grande’ puede aparecer con nombres contables pero no con incontables (cf. **el gas grande*), no parece apropiado algo como **tres culpas grandes*. Si es incontable, deberíamos esperar encontrarnos una utilización indeterminada amplia (§3.2.1): pero no es así. La forma determinada es hegemónica sin duda.

Hay que subrayar el carácter especial de un nombre como *ogen* ‘culpa’. Por un lado, como hemos dicho ya, puede tomar el adjetivo *handi* ‘grande’, propio de los nombres contables, pero luego no se puede contar (**Hiru ogen handi* ‘tres grandes culpas’). Por otro lado, aunque sea un argumento del verbo *ukan* ‘tener’, tiene una semántica predicativa grande; es fácilmente parafraseable de la siguiente manera: *ogendun handia izan nintzen* ‘fui un gran culpable’. No sería pues, demasiado inverosímil tomarlo como atributo. Haciendo referencia a la limitación que hemos mencionado ya varias veces, las estructuras de *nombre + adjetivo* indeterminadas serían imposibles o muy esporádicas en los predicados ya para el siglo xix, es decir, para la época de nuestro corpus.

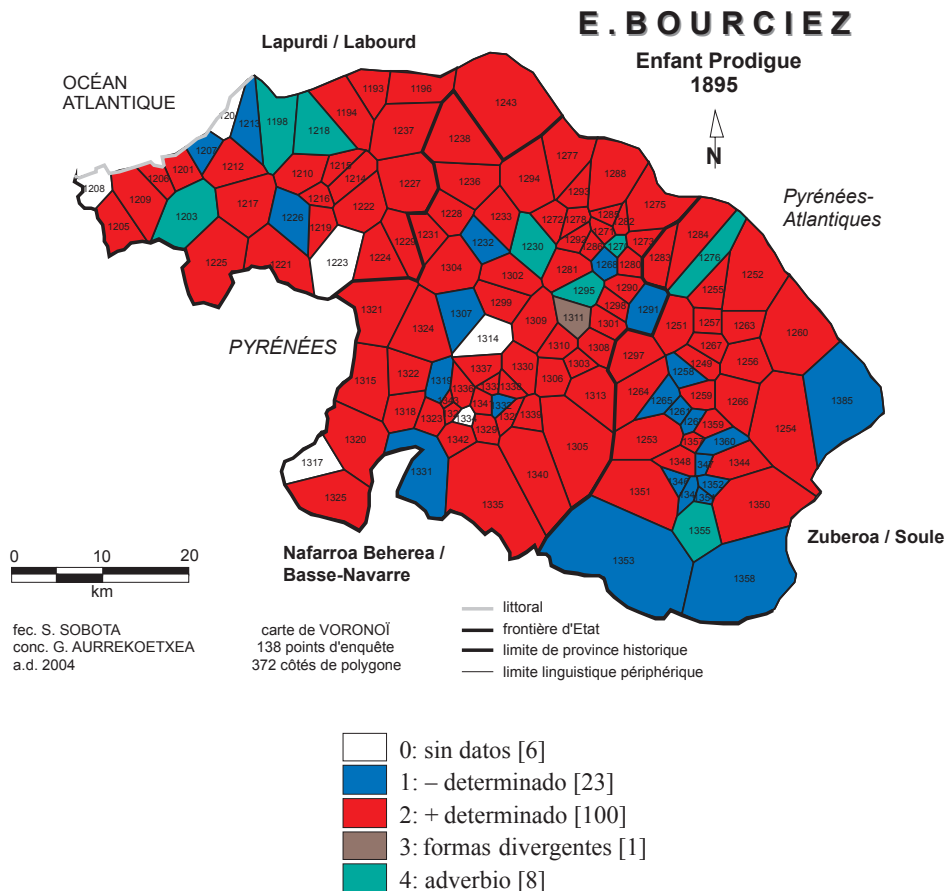
Las pocas apariciones indeterminadas que hay se podrían explicar también mirando el texto francés original. Dice *J'eus grand tort*, forma indeterminada, sin el artículo *du* ni otros. Ésta sería, también en francés, una forma fosilizada.

3.2.5. *Un sintagma contable digno de mencionar*

En las oraciones de debajo tenemos un sintagma que tomaríamos por contable:

(22) *Il faut que (...) je vois du pays.*

- a) *Behar dut (...) ikusi herriØ.* – determinado
- b) *Behar ditut ikusi herriak.* + determinado

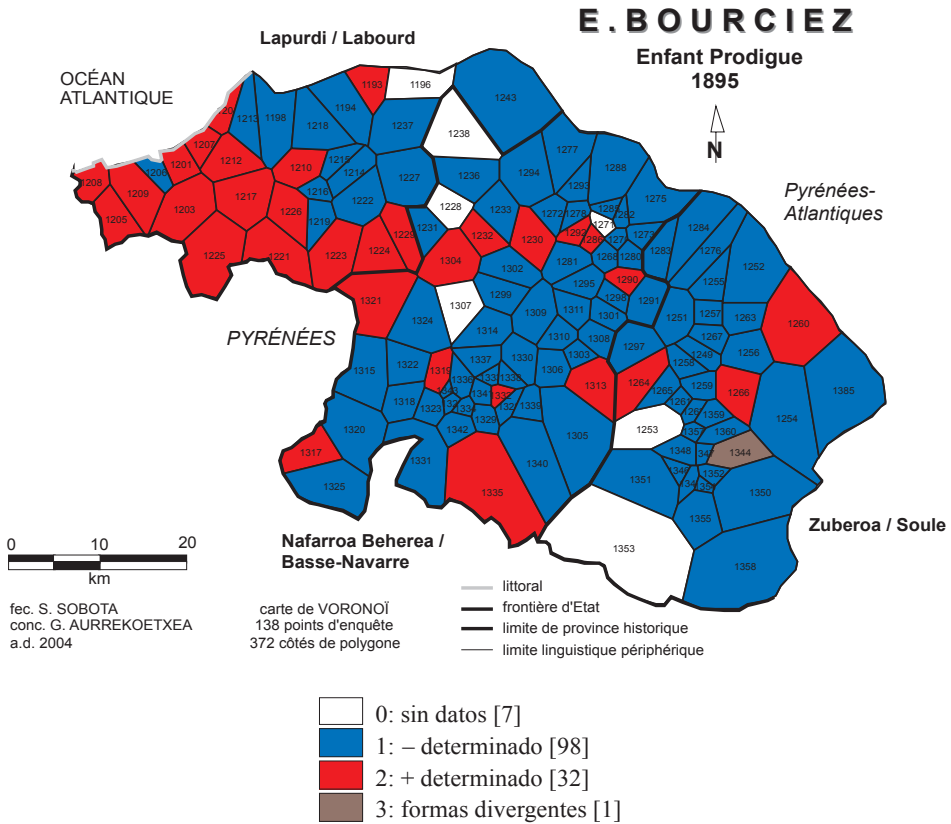


Mapa 6

Estructuras de *nombre + adjetivo*

Los pueblos o países son contables: podríamos decir sin inconveniente *cinco pueblos*. En estos casos de (22a), el empleo de la forma indeterminada se le debe achacar a la voluntad de no precisar los pueblos o países en concreto. De momento no hay ninguna razón para decir que las oraciones de (22) sean algo diferente a las de (17). Así las cosas, esperaríamos que el mapa que a continuación traemos fuera bastante similar al Mapa 4, puesto que ámbos hacen referencia a términos contables. Pero veamos el Mapa 7 de la página siguiente.

Para empezar, en este mapa no tenemos ni un solo sintagma que contenga algún determinante indefinido, como lo tenemos en (17c): No nos topamos con nada como *Behar dut ikusi kartiel zenbait*: todos son SN desnudos o SD con artículo. A su vez, las formas indeterminadas son muchas, sobre todo en Zuberoa, pero en Baja Navarra y en el territorio de Lapurdi de habla bajonavarra también son numerosas, cosa que no sucede en el mapa 4. En el territorio del labortano, por supuesto, hay más formas determinadas.



Mapa 7

Concepto contable peculiar

¿Por qué son tan diferentes los dos mapas? o, es más, ¿por qué son tan similares este mapa de nombres contables y el que hemos mostrado para los nombres incontables? (cf. Mapa 3). En nuestra opinión la respuesta se halla en el texto original en francés: aquél que Bourciez distribuyó para que fuera traducido. En ése, la oración francesa dice así: *Il faut (...) que je voie **du** pays*.

Hemos marcado el determinante en **negrita**, para que el lector constate que es el mismo que se emplea con nombres incontables. Es decir, en el texto en francés *pays* no es contable (si no llevaría el determinante *des*), sino incontable. A nuestro parecer, tener una forma incontable en el modelo a traducir habría ayudado a que en las traducciones en vasco la forma indeterminada tenga tal expansión. Las hablas que a la sazón tenían bastante debilitada la opción de emplear la forma indeterminada como recurso para expresar el valor no específico, habrían visto en el modelo francés una excusa para recuperar esa utilización indeterminada, aunque en §3.2.2 no lo hayan hecho.

Para terminar, es digno de remarcar hasta qué punto están extendidas las formas con artículo en el sur de Lapurdi, que tienen necesidad de emplear el artículo incluso superando al modelo francés, que no lo hace.

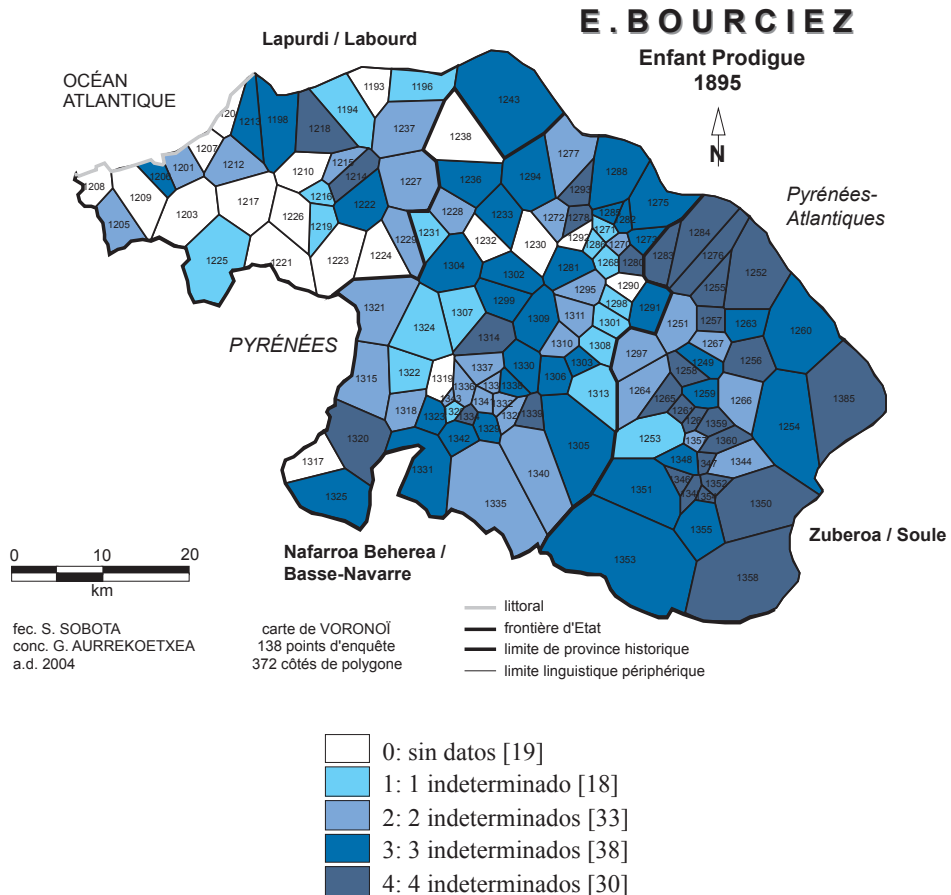
3.3. Mapas con respecto al carácter semántico y sintáctico, o “mapas de los mapas”

Como hemos explicado, el programa *VDM* no permite más que un sólo resultado por localidad (§2.3). ¿Cómo hacer, por ejemplo, la síntesis de todos los mapas con función de complemento directo (§3.2) en un sólo mapa, para poder extraer conclusiones generales?

Lo explicaremos al presentar cada mapa. Se podrían hacer numerosos mapas así y nosotros hemos traído aquí algunos.

3.3.1. Mapas sobre los sintagmas argumentales

En este apartado queremos ofrecer la síntesis de los mapas empleados en §3.2, es decir, los mapas que a continuación se muestran están basados en los mapas de sintagmas argumentales de función objeto directo. Veamos el primero:

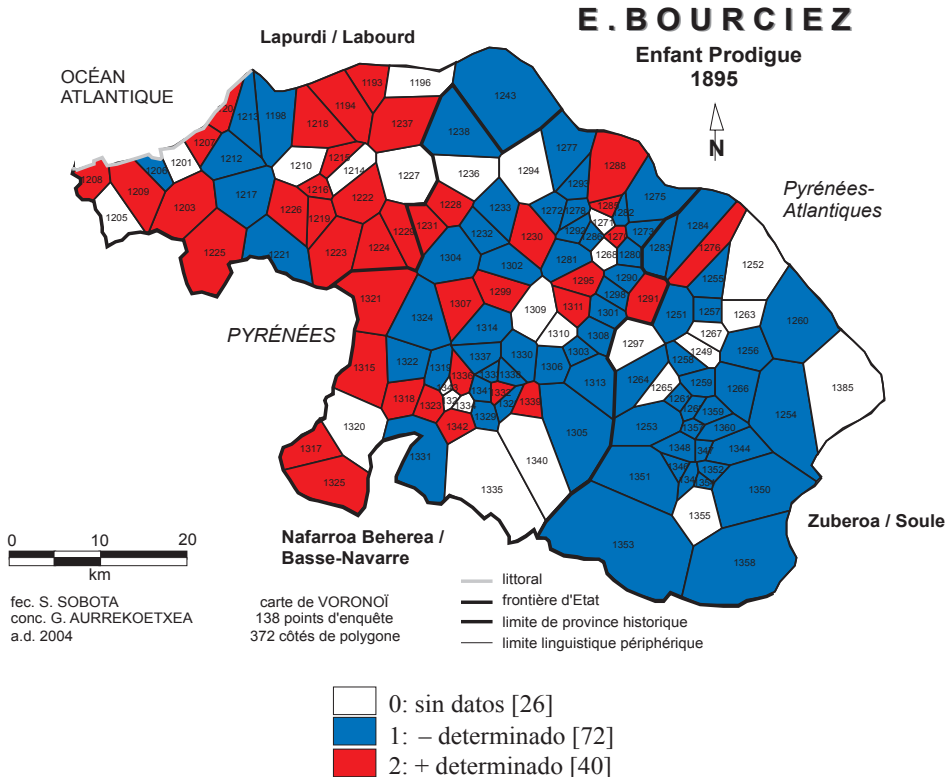


Mapa 8

Número de sintagmas indeterminados por localidad

En este mapa hemos tomado cuatro oraciones cuyo sintagma sometido a estudio tiene función de objeto directo; los del apartado §3.2, y hemos contado cuántos de ellos aparecen sin artículo.¹⁴ Así las cosas, a las localidades que han hecho las cuatro oraciones con el sintagma indeterminado les hemos asignado el mismo lema (y color); lo mismo a las que han hecho con el indefinido tres oraciones de cuatro, etc. En este mapa, por lo tanto, hemos tratado por igual todas las oraciones; no hemos hecho divisiones por semántica o forma (contable vs. incontable, etc.). De todas formas, en este mapa se puede ver de manera clara cómo va penetrando el artículo de oeste a este, siendo una vez más el territorio del labortano el que menos formas indeterminadas presenta, y el del suletino el que más SN desnudos tiene.

En el siguiente mapa hemos trabajado de nuevo con sintagmas argumentales, pero hemos contrastado las formas determinadas y las indeterminadas. A las localidades que han hecho la mayoría de las cuatro oraciones con el artículo les corresponde el color rojo: por el contrario, a las localidades que tienen más oraciones indeterminadas, el azul. Éste es el resultado:



Mapa 9

Sintagmas argumentales determinados vs. indeterminados

¹⁴ No hemos tenido en cuenta el mapa de §3.2.4, porque para cuando lo preparamos ya teníamos hecho éste.

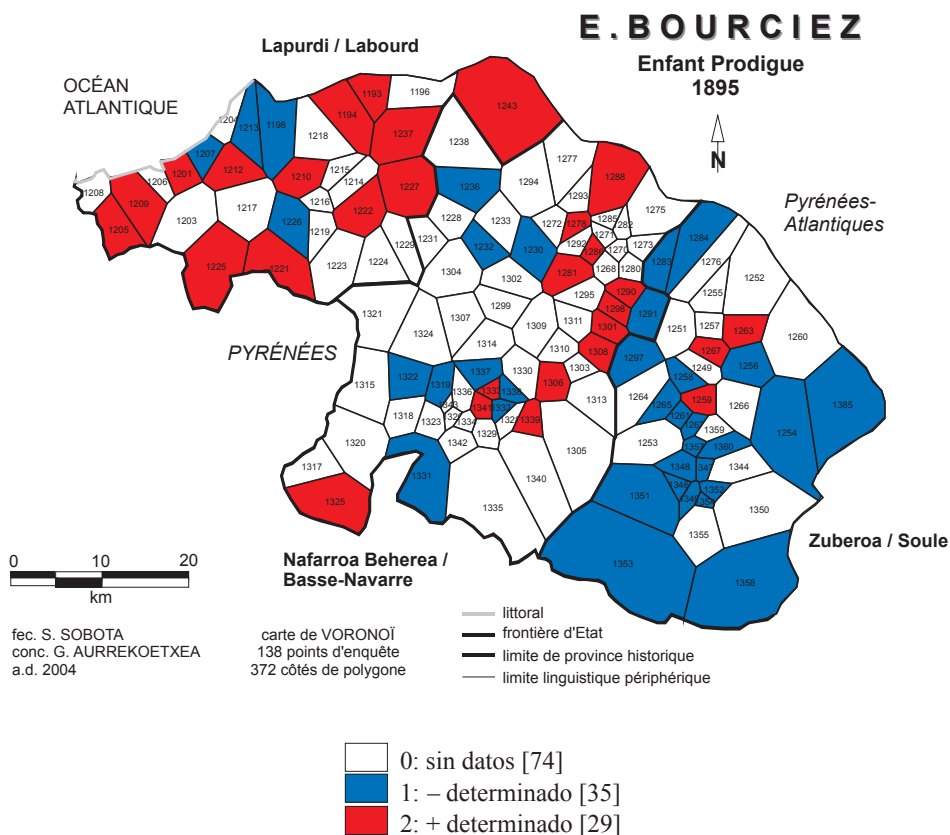
En éste también, los resultados son clarificantes. Mientras que en Zuberoa la mayoría de oraciones se han realizado sin artículo, en Lapurdi es mayoría la presencia de SD con artículo. Baja Navarra es una zona de transición evidente, como se puede ver en el mapa.

3.3.2. Mapas de los sintagmas de función atributiva

Son menos los ejemplos de sintagmas que no son argumentales (§3.1) y es necesario tener eso en cuenta a la hora de valorar los datos.

Así las cosas, esta vez no hemos realizado un mapa como el 8, pues al tener sólo dos ejemplos y ser uno de ellos casi por completo con formas indeterminadas (cf. Mapa 2), el resultado no sería representativo.

Sí hemos hecho, al igual que con los sintagmas de complemento directo, un mapa de mayorías. Corresponde el mismo lema y color a las localidades que han realizado los dos atributos sin artículo y lo mismo para los que emplean atributos determinados. El resto de combinaciones posibles han sido eliminadas, y se representan en blanco en el mapa.



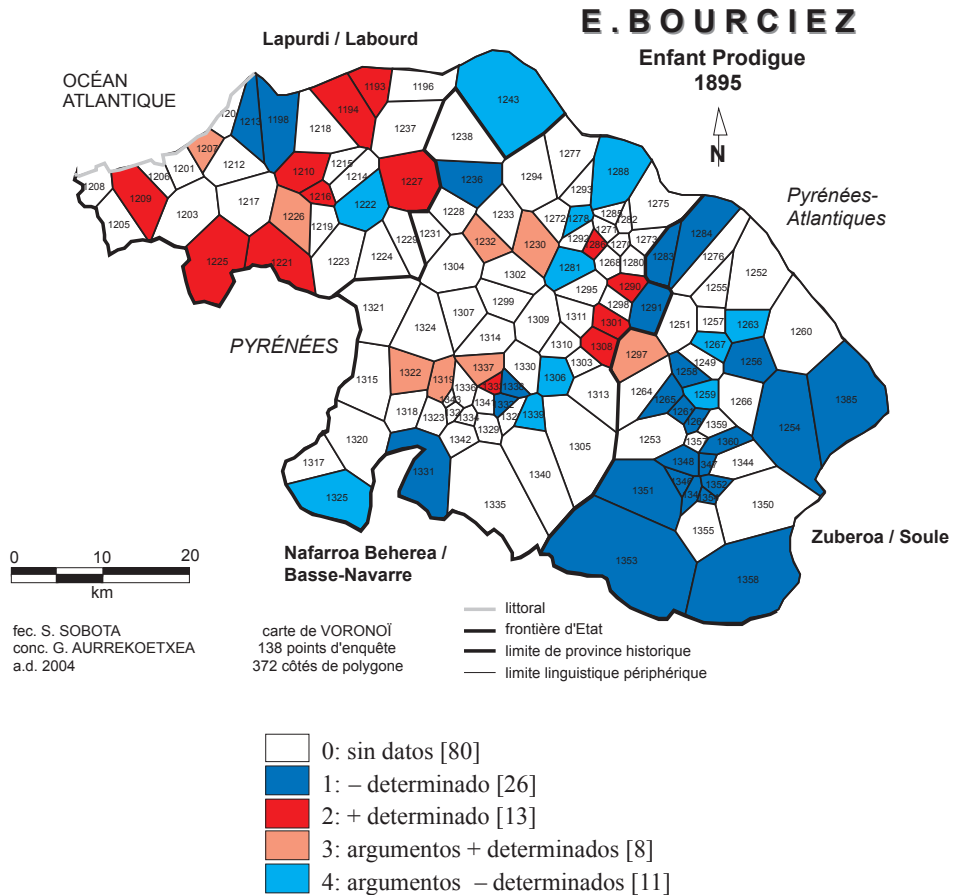
Mapa 10

Sintagmas atributivos determinados vs. indeterminados

Son muchos los espacios en blanco, porque en muchas ocasiones, a pesar de dar sin artículo las oraciones representadas en el mapa 2, las del mapa 1 son determinadas, lo que nos deja en una situación de empate entre los determinados y los indeterminados. Pero lo llamativo es lo siguiente: solamente (y casi exclusivamente) en los lados del mapa tenemos formas extremas, es decir, sobre todo en Lapurdi se han empleado siempre formas determinadas, y en Zuberoa las indeterminadas. Los grandes espacios en blanco de la Baja Navarra dejan entrever la prominencia de combinaciones intermedias, en las cuales, mientras una forma es determinada, la otra no.

3.3.3. Comparando los argumentos con los predicados

Ahora que estamos presentando mapas generales, no podemos dejar de comparar los sintagmas de función atributiva con los de complemento directo. A ello se dispone este mapa, el último, por cierto.



Mapa 11

Sintagmas proposicionales vs. argumentales

Como salta a la vista, son cuatro las opciones posibles en este mapa, aparte de la falta de datos, claro está. Han tomado tono azul las localidades que han utilizado principalmente formas indeterminadas, tanto con sintagmas de función complemento directo, como en sintagmas de función atributo. Les hemos dedicado el rojo a las que han preferido los sintagmas determinados, sean atributos o complementos. El rosa es para las localidades en las que, aunque en las oraciones de atributo prevalezcan las formas indeterminadas, presenten más formas determinadas entre las oraciones de complemento directo. Finalmente, el azul claro ha sido asignado a las localidades donde, siendo mayoría los sintagmas indeterminados para los atributos, realizan los complementos directos sobre todo con el artículo.

Como viene siendo habitual, en ambos costados, al este y al oeste, tenemos los colores más prominentes y en medio el resto de combinaciones posibles. De nuevo toparemos con la coherencia lingüística de Zuberoa, que, al igual que en éste, es visible también en el resto de mapas.

4. Conclusiones

Es menester ya resumir todo lo dicho hasta ahora. Partiendo de la base de que podemos clasificar las lenguas tipológicamente dependiendo de la utilización del artículo que hagan, y sabiendo que, ya sea de lengua a lengua o de dialecto a dialecto, dicha utilización cambia, prolifera o se reduce, hemos querido traer a estas páginas un ejemplo del vasco.

Escogido el corpus de Bourciez como muestra de la variación diatópica, hemos mencionado los trabajos de recogida de la época y los hemos aprovechado para presentar nuestro corpus, al mismo tiempo que lo hemos situado en su contexto geográfico y temporal correspondiente, mencionando a su vez las ventajas e inconvenientes que nos brinda.

Hemos dado una visión general del programa *VDM*, describiendo el uso que hemos realizado nosotros. Siguiendo esa línea, hemos descrito paso a paso la metodología y procedimiento de trabajo, a la vez que hemos expuesto los problemas a los que debemos enfrentarnos.

Para que la interpretación de los resultados del corpus sea correcta, hemos descrito las oraciones que hemos elegido para nuestra investigación, basándonos en los trabajos de diversos lingüistas. De esa manera, hemos tenido la oportunidad de comprobar hasta qué punto nuestros datos convergen con lo dicho hasta el momento, añadiendo a su vez, nuevo material para futuras investigaciones.

Hemos comentado los mapas uno a uno, describiendo lo que muestran y explicando de qué manera concuerdan o difieren con las teorías de los lingüistas que acabamos de mencionar. Aparte de eso, hemos expuesto las dificultades que presentan algunas oraciones del corpus y hemos intentado dar una explicación plausible a incoherencias geográficas que nos han podido surgir en los mapas.

También hemos confeccionado unos mapas generales basados en los mapas extraídos directamente del corpus. Gracias a ellos hemos podido sacar unas conclusiones más globales, resumiendo en unas pocas fotografías el empleo del artículo a finales del siglo XIX.

Desgraciadamente, la brevedad del corpus no nos ha permitido reflejar toda la caustística sobre el empleo del artículo, pero hemos podido extraer algunas conclusiones claras:

- Las formas con artículo son más numerosas en el oeste que en el este.
- Aunque antaño los SN desnudos eran posibles en cualquier posición argumental, han pervivido durante más tiempo en función de complemento directo y en vasco estándar, como en muchos dialectos, ya no son posibles ni siquiera en dicha posición.
- En lo que respecta a la aparición del artículo, los sintagmas de función complemento y los atributos presentan una geografía diferente.
- Dentro de los sintagmas de complemento directo, presentan una geografía diferente los nombres contables y los incontables.
- Desde el punto de vista diacrónico, las diferencias de isoglosas mencionadas en los puntos anteriores abogarían por una expansión tardía del artículo del oeste al este, teniendo incluso la posibilidad de distinguir diferentes fases de gramaticalización.
- Mirando a la geografía, Zuberoa muestra una coherencia lingüística muy grande en favor de los indeterminados y el territorio del labortano (el sur de Lapurdi) puede diferenciarse muy bien mediante el empleo del artículo, aunque no es tan claro como el de Zuberoa. La Baja Navarra actúa a menudo como zona de transición.
- Los mapas parecen mostrar un estadio de la lengua en el cual el artículo presenta diferentes grados de gramaticalización. El *markedness shift* habría sucedido ya en algunos puntos de Lapurdi, mientras que en Zuberoa prácticamente no habría comenzado ni siquiera la alternancia entre formas desnudas y determinadas.
- Desde el lado metodológico es conveniente mentar la fiabilidad de los mapas. No se puede realizar una interpretación pueblo a pueblo y hay que tomar el mapa en su totalidad: debido a las características del sistema de recogida de datos (un texto por localidad, traducido por gente alfabetizada, etc.), en algunas localidades aparecen incoherencias dialectológicas. Es labor de los lingüistas hacer interpretaciones más generales, por encima de esas carencias.

No quisiera terminar, sin hacer una proclama en contra de la tradicional (al menos desde tiempos de Saussure) división entre diacronía y sincronía. Entre los objetivos de este trabajo, aunque sea humildemente, estaba mostrar la debilidad de esa frontera. Que la metodología sincrónica utilizada haya proporcionado también a los diacronistas algunas conclusiones a tener en cuenta sería un sueldo más que digno a cambio del trabajo realizado.

5. Referencias bibliográficas

- Allières, J., 1960-1961, «Petit atlas linguistique basque-français <Sacaze> I-II», *Via Domitia VII-VIII*, Faculté des Lettres et Sciences humaines de Toulouse, Toulouse, 82-126 & 205-224.

- Andersen, H., 2001a, «Markedness and the Theory of Linguistic Change», in H. Andersen (ed.), *Actualization. Linguistic Change in Progress*, John Benjamins Publishing Company, Amsterdam & Philadelphia, 21-57.
- , 2001b, «Actualization and the (Uni)directionality of Change», in H. Andersen (ed.), *Actualization. Linguistic Change in Progress*, John Benjamins Publishing Company, Amsterdam & Philadelphia, 226-248.
- Artiagoitia, X., 2004, «Izen Sintagmaren birziklatzea: IS-tik inguruko funtzio buruetara», in P. Albizu & B. Fernández (eds.), *Euskal gramatika XXI. mendearen atarian: arazo zaharrak, azterbide berriak*, Diputación Foral de Álava, Vitoria-Gasteiz, 11-38.
- Aurrekoetxea, G. & X. Videgain, 2004, *Haur prodigoaren parabola Ipar Euskal Herriko 150 bertsiotan*, ASJUren gehigarriak, XLIX, UPV/EHU, Bilbao.
- Azkarate, M., & P. Altuna, 2001, «3.3.3. Predikatu sintagmak: atributu sintagmak», in *Euskal Morfologiaren Historia*, Elkarlanean, Donostia, 74-76.
- Croft, W., 1993² [1990], *Typology and universals*, Cambridge U. P., Cambridge.
- Eguren, L., s.d., «Non-canonical uses of the article in Basque». Manuscrito de la UAM.
- Etxeberria, U. & R. Etxepare, 2008, «Zenbatzaileak komunztatzen ez direnean: Hiru sistema», in *Aldaketak, Aldaerak, Bariazioak Euskararen eta Euskal Testugintzan*, Abenduaren 12-13a, IKER-CNRS, Baiona. Handout del congreso.
- Gándara, A. & E. Santazilia, 2007, «Zehaztapen batzuk artikuluen erabileraz Bourciez korpusean». Manuscrito de la UPV/EHU.
- Longobardi, G., 2001, «The structure of DPs: some Principles, Parameters and Problems», in C. Collins & M. Baltin (eds.). *The Handbook of Contemporary Syntactic Theory*, Blackwell, USA/UK.
- Manterola, J., 2006, «-a euskal artikuluen definituaren gainean zenbait ohar», on line en la base de datos ARTXIKER: <http://artxiker.ccsd.cnrs.fr/artxibo-00142080/eu/> [consulta: 2009-09-03]. Publicado también en *ASJU* 40: 1-2, 651-676.
- , 2008, «-a morfemaren erabilera (eza) ekialdeko euskaretan», on line en la base de datos ARTXIKER: <http://artxiker.ccsd.cnrs.fr/artxibo-00352804/eu/> [consulta: 2009-09-03].
- Mitxelena, L., 1978, «II. Vasc. ON DA/GAUZA ONA DA», in «Miscelánea filológica vasca», *FLV* X: 29, 208-218.
- Oyharçabal, B., 1992a, «Euskararen mugez egin lehen mapak (1806-1807)», in *Luis Villan-teri omenaldia*, Iker-6, Euskaltzaindia, Bilbao, 349-366.
- , 1992b, «Lehenbiziko inkesta geo-linguistikoak Euskal Herrian frantses lehen Inperioaren denboran: ipar aldean bildu dokumentuak», in *Nazioarteko Dialektologia biltzarra. Agiriak*, Iker-7, Euskaltzaindia, Bilbao, 285-298.
- , 1994, «Les documents recueillis lors des enquêtes linguistiques en Pays Basque durant la période révolutionnaire et le Premier Empire», in J. B. Orpustan (ed.), *La révolution française dans l'histoire et la littérature basques du XIXème siècle*, Izpegi, Baigorri, 62-119.
- , 1995, «Euskararen mugak hego aldean 1807.ean: Coquebert de Montbrer-ek bildu dokumentuak», in R. Gómez & J. A. Lakarra (eds.), *Euskal Dialektologiako kongresua (Donostia, 1991ko Irailak 2-6)*, *Suplementos de ASJU* XXVIII, Diputación Foral de Gipuzkoa, Donostia.
- Sacaze, J., 1887, *Recueil de linguistique et de toponymie des Pyrénées*.
- Santazilia, E., 2007, «Edouard Bourciezen *Haur prodigoaren parabola*: edizio paleografiko konparatua», *ASJU* 41-1, 397-402.
- Simoni-Aurembou, M.-R., 1989, «La couverture géolinguistique de l'Empire français: l'enquête de l'enfant prodigue», in *Espaces Romains: Études de dialectologie et de géolinguistique offertes à Gaston Tuailon*, Université Stendal-Grenoble, Grenoble, II, 114-135.
- Videgain, X., 2005, «Présentation du recueil Bourciez», *Lapurdum* X, 315-324.
- Zabala, I., 2001, «4.2 Nominal Predication: copulative sentences and secondary predication», in J. I. Hualde & J. Ortiz de Urbina (eds.), *A Grammar of Basque*, Mouton de Gruyter, Berlin, 426-447.