

**EL ANÁLISIS DEL LÉXICO DE BILINGÜES PRECOSES A TRAVÉS DE
DIFERENTES MÉTODOS CUANTITATIVOS**

**LEXICAL ANALYSIS OF EARLY BILINGUAL PEOPLE WITH DIFFERENT
QUANTITATIVE METHODS**

RESUMEN

Son pocos los trabajos y metodologías de análisis con las que se comparan las diferencias léxicas en función de la lengua materna de los hablantes y de la lengua utilizada en contextos bilingües. La finalidad de este trabajo es constatar el empleo del léxico de las informantes, según su lengua materna midiendo el tiempo de reacción desde la visualización del estímulo hasta la emisión léxica correspondiente. Para ello, se han obtenido los datos de la emisión léxica de 40 alumnas de la Facultad de Educación de Bilbao empleando un cuestionario de 34 estímulos por medio de una secuencia de diapositivas, al que las informantes debían responder lo más rápido posible. Se ha realizado la clasificación de las informantes según los datos obtenidos mediante los siguientes métodos cuantitativos objetivables método de conglomeración jerárquico, K-medias y el algoritmo FCM basado en la lógica difusa. En los resultados obtenidos se observa la clasificación de las informantes en grupos heterogéneos.

PALABRAS CLAVE: bilingüismo precoz; léxico; métodos cuantitativos.

ABSTRACT

There are few works and methodologies of analysis where the lexical differences are compared according to the mother tongue of the speakers and the language used in bilingual contexts. Therefore, the purpose of this work is to verify the use of the informants' lexicon, according to their mother tongue, by measuring the reaction time from the visualization of the stimulus

and the corresponding lexical emission. For this purpose, the data of the lexical emission were obtained of 40 female students of the Faculty of Education of Bilbao using a questionnaire of 34 entries by a sequence of slides, which the informants had to respond as quickly as possible. The classification of the informants was carried out according to the data obtained by means of the following objective quantitative methods hierarchical method, K-means and the FCM based on Fuzzy C-means algorithm. In the results obtained, is observed the classification of the informants in heterogeneous groups.

KEY WORDS: Early bilingualism; lexicon; quantitative methods

1. INTRODUCCIÓN

En este trabajo presentamos un análisis introductorio sobre algunos aspectos de la riqueza léxica en la lengua vasca. Los trabajos sobre el léxico en euskera se han venido realizando en dos vertientes principales: confección de diccionarios generales en la lengua estándar o diccionarios específicos sobre ciertas materias (biología, física, música, deportes, etc.) y monografías sobre el léxico dialectal. En general, en este tipo de trabajos no se tiene en cuenta la cuestión relativa a la riqueza léxica, aspecto este que entronca directamente con la capacidad léxica de los hablantes y que debería tener su impacto en la enseñanza de la lengua. Una de las aportaciones relevantes a este respecto, desde nuestro punto de vista, es la metodología novedosa que se emplea en el trabajo para la obtención de los datos.

En la comunidad lingüística vasca se han dado cambios significativos en los últimos años. Esos factores de cambio han sido, por un lado, la estandarización del euskera y por otro, lado la adopción de los diversos modelos educativos bilingües dentro del marco formal y su

posterior crecimiento.

El bilingüismo precoz surge cuando los hablantes adquieren su lengua materna en el hogar y la segunda lengua dentro del contexto formal educativo y/o en la sociedad. Actualmente, dentro del bilingüismo precoz se diferencian principalmente dos grupos en función del lugar de adquisición de cada una de las dos lenguas oficiales. Así, el primero de ellos es el compuesto por alumnado que ha adquirido una de las variedades tradicionales del euskera en el entorno familiar, y que, a su vez, también ha adquirido el español (o francés) a edades muy tempranas, bien en el ámbito educativo, bien en la sociedad. El segundo grupo de bilingües precoces es el formado por el alumnado que ha adquirido el español por transmisión familiar y el euskera en el sistema educativo a edad muy temprana.

Esto ha traído consigo el interés por investigar y querer indagar más en el bilingüismo precoz en el contexto que describimos. Los trabajos que han analizado las producciones y emisiones de los jóvenes bilingües precoces las han abordado desde distintos enfoques y perspectivas (Etxebarria, Gaminde, Olalde & Gaminde, 2016; Eguskiza, De Pablo & Gaminde, 2017; Gaminde, 2010, 2011; Gaminde & Eguskiza, 2017; Gaminde, 2011). Sin embargo, estas investigaciones apenas han abarcado el campo del léxico (Gaminde y otros, 2016a; 2016b), campo que ha recibido atención únicamente de manera tangencial.

Son pocos los trabajos y metodologías de análisis con las que se comparan las diferencias léxicas en función de la lengua materna de los hablantes y de la lengua utilizada en contextos bilingües (Gaminde y otros, 2016a; 2016b).

Por tanto, el objetivo de este trabajo es investigar la riqueza léxica de las alumnas de la Facultad de Educación teniendo en cuenta tanto la lengua materna de las informantes como

la lengua de reproducción.

El trabajo se divide en tres apartados a partir de éste. En el primero de ellos se describen el corpus y la metodología utilizada en este estudio; a continuación, se presentarán los resultados y, por último, se explicarán las conclusiones.

2. CORPUS Y METODOLOGÍA

Con la finalidad de alcanzar los objetivos establecidos, se ha llevado a cabo la recogida de datos que forman el corpus. Para ello, se ha contado con la participación y colaboración de 40 alumnas de la Facultad de Educación de Bilbao. Se ha elegido este grupo de informantes por el efecto multiplicador que este colectivo debe tener en el futuro, ya que serán las responsables de la transmisión lingüística a las siguientes generaciones.

Todas ellas han nacido entre los años 1992 y 1998, han cursado todos sus estudios en el modelo D¹ y son bilingües precoces, es decir, han adquirido una de las dos lenguas (euskera o castellano) por medio de la transmisión familiar y la segunda lengua en el contexto formal educativo a edades muy tempranas. Veinte de ellas tienen como lengua materna el euskera y las otras veinte el castellano.

Para la obtención de los datos, se ha empleado la misma metodología en ambas lenguas. Primeramente, se les dio las instrucciones del procedimiento. Se les explicó que una vez visualizado un estímulo debían responder con la mayor brevedad posible con la palabra correspondiente a dicho estímulo. Los estímulos aparecían en una pantalla de ordenador por medio de un PowerPoint, cada vez que se visualizaba un estímulo sonaba un pitido al unísono

¹ Modelo D: Enseñanza en euskera como lengua vehicular, excepto en la asignatura de Lengua y Literatura Castellana e Inglés.

que sirve posteriormente como referencia para calcular el tiempo desde la audición del pitido hasta la emisión léxica correspondiente a cada uno de ellos. Los estímulos del cuestionario se corresponden al campo semántico compuesto por las partes del cuerpo humano. Estos estímulos fueron elegidos por ser los de mayor frecuencia en ambas lenguas y, a priori, por su accesibilidad léxica. De esta manera, el corpus queda constituido por los 34 estímulos los cuales se muestran en la Cuadro 1.

CUADRO 1. Orden de los estímulos visuales realizados

(1) oreja, (2) mano, (3) dedo, (4) boca, (5) nariz, (6) cuello, (7) brazo, (8) barbilla, (9) corazón, (10) ojo, (11) espalda, (12) riñón, (13) muela, (14) uña, (15) muslo, (16) colmillo, (17) frente, (18) lengua, (19) moratón, (20) campanilla, (21) rodilla, (22) pestañas, (23) cabeza, (24) tobillo, (25) dedo del pie, (26) talón, (27) muñeca, (28) pierna, (29) pantorrilla, (30) codo, (31) axila, (32) venas, (33) orzuelo y (34) ombligo

Por lo tanto, se ha conseguido un corpus total de 1360 datos, conformados por 34 estímulos y 40 informantes. 10 de las informantes de lengua materna euskera han realizado sus producciones en su L1 y las otras 10 en su L2. Del mismo modo, la mitad de las informantes de lengua materna español han realizado sus emisiones en su lengua materna y la mitad restante en su L2. En el Esquema 1 mostramos la distribución de las informantes según su lengua materna y la lengua de producción léxica en la que se ha obtenido el corpus.

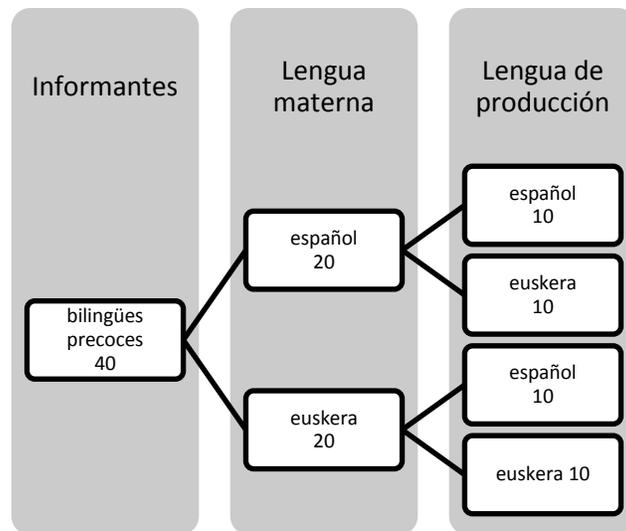


FIGURA 1. Distribución de las informantes según LM y lengua de producción

Todas las grabaciones han sido realizadas en contextos silenciosos empleando un ordenador y un micrófono externo omnidireccional de condensador (BOYA B1-M1). Las grabaciones se han realizado en el formato .wav para el posterior anotado y etiquetado mediante el programa Praat (Boersma & Weenink, 2016). Cada emisión acústica ha sido anotada y etiquetada desde el pitido hasta el comienzo de la producción léxica de las informantes. Los datos anotados y etiquetados han sido volcados a una matriz previamente creada mediante un script. El script ha facilitado la extracción automática de los datos en milisegundos (ms). Se han recogido todas las emisiones con excepción de aquellas que las informantes no conocían la respuesta, etiquetando estas últimas como nulas.

El análisis de los datos se ha efectuado en dos planos. Por una parte, teniendo en cuenta la lengua materna y la lengua de emisión de las informantes, ambas se consideran como variables en función de las cuales se efectúan los diversos análisis sobre el tiempo de reacción de las informantes. Por otra parte, según los tiempos de reacción se clasifican las informantes

siguiendo tres métodos objetivables de conglomeración: el método jerárquico, el método de K medias y el método de C medias basado en el algoritmo de la lógica difusa (FCM). En todos los casos la distancia utilizada ha sido la distancia euclídea y en el método de conglomeración jerárquico se ha utilizado el método de Ward, por ser éste el que mejor explota las similitudes y disimilitudes existentes entre los datos (Aurrekoetxea, 1995, 2005; Goebel, 1992, 2010; Pardo & Ruiz, 2002, Pérez López, 2005).

El primer método clasifica a las informantes en un único grupo, es decir, si el dendograma creara dos grupos, una determinada informante sería parte de uno de los dos grupos, no dando información de la posible parte perteneciente al otro. La clasificación de las informantes es realizada en función de la distancia que hay entre el resto de elementos y sin anteriormente establecer la cantidad de grupos que se desea obtener. Así, a través de la realización del mismo se obtiene un dendograma.

En el método K-medias, a diferencia del anterior, hay que elegir previamente la cantidad de conglomerados que deseamos obtener. Este método destaca las igualdades y las diferencias entre los grupos. De manera que en éste es el investigador quien toma la decisión de la elección de la cantidad de grupos.

En el tercer método (FCM), además de mantener la elección en lo que respecta al número de conglomerados se muestra la función de pertenencia que se denota de la siguiente manera:

$$\mu_A: X \rightarrow [0,1]$$

La pertenencia al grupo abarca desde 0 a 1, siendo 0 la medida que nos demuestra que no pertenece a ese grupo en absoluto y 1 sería cuando pertenece completamente al mismo. De

manera que, el informante que tenga más de 0'5 en uno de los grupos, por ejemplo 0,7, significa que pertenece mayoritariamente a ese grupo, en un 70%, pero también es, en menor medida, parte del otro, en un 30%. Asimismo, la aplicación de este método permite al investigador/a la elección de (1) la cantidad de grupos que desea crear y (2) el índice de pertenencia; esto es, nos permite decidir a partir de cuándo es altamente perteneciente a ese grupo (en nuestro caso, hemos decidido poner la frontera en 0'85). Aquellas informantes que estén entre 0'5 y 0'85, se les denomina informantes de transición, pues sólo pertenecen en cierta medida al grupo.

3. RESULTADOS

A continuación, se procederá a explicar los resultados en cuatro apartados. Primeramente, se desarrollará el apartado del análisis del tiempo de reacción de las informantes para después clasificarlas según los diferentes métodos expuestos en el apartado metodológico; en cuyo caso se llevará a cabo el desarrollo de este análisis atendiendo a los resultados obtenidos con las respuestas en su lengua materna y con las de la segunda lengua.

Asimismo, quisiéramos señalar que los resultados se expondrán haciendo referencia en todo momento a los siguientes cuatro grupos:

- 1) Informantes con lengua materna español y lengua de reproducción español.
- 2) Informantes con lengua materna español y lengua de reproducción euskera.
- 3) Informantes con lengua materna euskera y lengua de reproducción español.
- 4) Informantes con lengua materna euskera y lengua de reproducción euskera.

3.1. Análisis del tiempo de reacción

En este apartado, primeramente, se muestran los datos generales (respuestas válidas y nulas) en función de los cuatro grupos identificados previamente. Después, se muestra el análisis de las respuestas nulas de cada una de las informantes.

En la Cuadro 2 mostramos la cantidad y porcentaje de las respuestas válidas y nulas en los cuatros grupos.

CUADRO 2. Número y porcentajes de respuestas nulas y válidas por lengua materna y lengua de reproducción

Grupos	Válidas		nulas		total	
	N	%	N	%	N	%
1	331	97,35	9	2,65	340	100
2	284	83,53	56	16,47	340	100
3	314	92,35	26	7,65	340	100
4	318	93,53	22	6,47	340	100

En cada uno de los grupos ha habido un total de 340 respuestas posibles, pero en aquellos casos en los que las informantes no han respondido al estímulo expuesto, la respuesta se ha considerado nula, como hemos dicho anteriormente. De manera que, de las respuestas posibles, en el grupo uno se han obtenido 331 respuestas válidas y 9 nulas, en el segundo grupo, 284 válidas y 56 nulas, en el tercer grupo 314 válidas y 26 nulas, y en el cuarto de los grupos, 318 válidas y 22 nulas.

Si analizamos los datos generales en función de la lengua materna y lengua de reproducción,

los resultados que obtenemos son que las respuestas válidas en el caso de aquellas informantes que respondían en su lengua materna era de media 324,5, es decir, un 95,44%; mientras que cuando respondían en un idioma diferente al de su lengua materna se obtuvieron una media de 299 respuestas válidas, esto es, un 87,94%. Por lo tanto, se puede concluir que hay más respuestas válidas cuando las informantes respondieron en su lengua materna que cuando no lo hicieron.

Si observamos en función de si la lengua materna es español o euskera, podemos ver cómo las informantes de lengua materna español obtienen un 90,44% de respuestas válidas frente al 92,44% de las personas con lengua materna euskera, logrando estas últimas un mayor porcentaje de respuestas válidas.

En cuanto a las respuestas nulas, para conocer si hay diferencias estadísticamente significativas se ha realizado la prueba de Mann-Whitney con la cual hemos obtenido los datos presentados en la Cuadro 3.

CUADRO 3. Significación (Mann-Whitney) según las respuestas nulas

	1	2	3	4
1		*0,02	*0,012	0,684
2	*0,02		0,144	*0,020
3	*0,012	0,144		0,338
4	0,684	*0,020	0,338	

Como se ve en el cuadro 3 los resultados estadísticamente significativos son el grupo 1 con el 2, el 1 con el 3 y el 2 con el 4.

En lo que respecta al tiempo de reacción, en el Gráfico 1 y Cuadro 4 exponemos la desviación estándar, mínimo, máximo y rango de las respuestas en función del grupo.

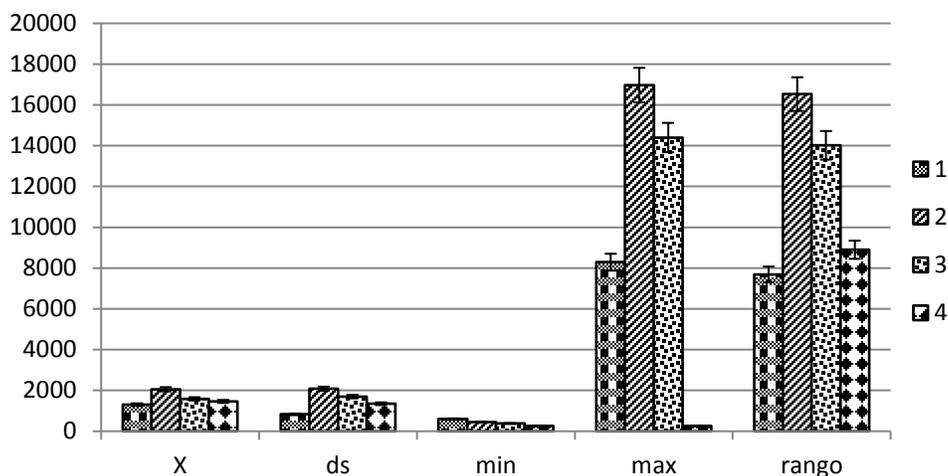


GRAFICO 1. Media, desviación estándar, mínimo y máximo de las respuestas

CUADRO 4. Media, desviación estándar, mínimo y máximo de las respuestas.

Grupos	\bar{x}	ds	min.	max.	rango
1	1304,41	830,26	600,00	8290,00	7690,00
2	2059,71	2082,66	440,00	16970,00	16530,00
3	1589,58	1697,38	380,00	14400,00	14020,00
4	1463,83	1349,09	260,00	260,00	8900,00

Dentro del análisis del tiempo de reacción se ha aplicado la prueba HSD de Tukey para conocer si hay diferencias estadísticamente significativas entre los grupos. Los datos obtenidos se muestran en la Cuadro 5.

CUADRO 5. Significación (HSD de Tukey) según el tiempo de reacción

	1	2	3	4
1		*0,000	0,086	0, 549
2	*0,000		*0,001	*0,000
3	0,086	*0,001		0,733
4	0, 549	*0,000	0,733	

Los grupos que muestran significación estadística entre sí son el grupo 1 con el 2, el 2 con el 3 y el 2 con el 4.

3.2. Clasificación de las informantes según la lengua materna y la lengua de reproducción

Tal y como exponíamos en el apartado dedicado a la metodología, para realizar la clasificación de las informantes según la lengua materna y la lengua de reproducción, se han utilizado tres métodos cuantitativos de conglomeración: el clúster jerárquico, K-medias y Fuzzy C-means.

3.2.1. Clúster jerárquico

Comenzaremos exponiendo los datos obtenidos mediante el análisis de conglomerados jerárquicos, por medio del cual se ha obtenido el dendograma de la Figura 2.

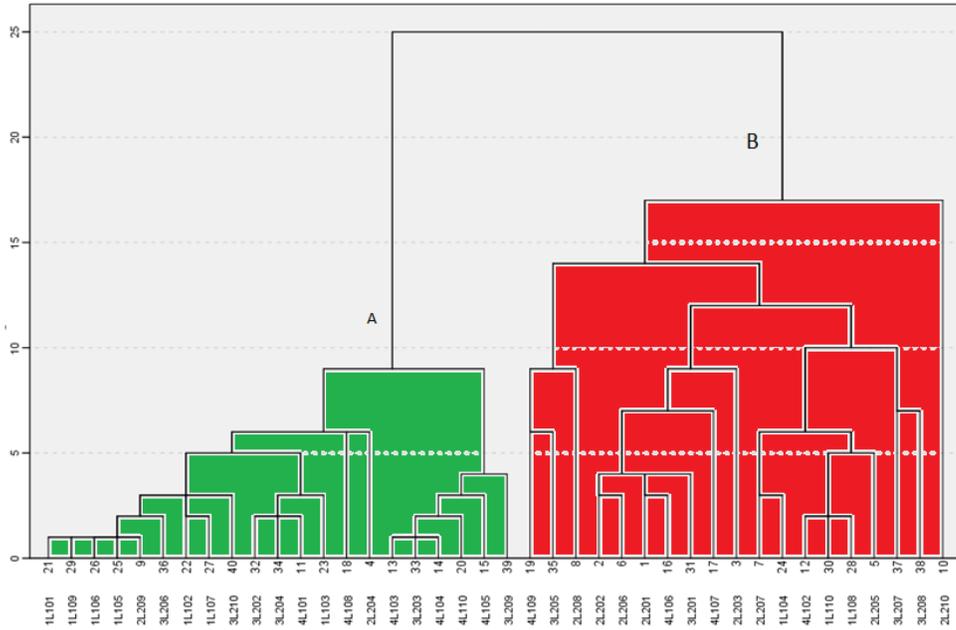


FIGURA 2. Dendrograma del clúster jerárquico

Se pueden ver claramente los dos principales grupos creados por medio del análisis jerárquico, a los que hemos denominado A y B. Con los resultados obtenidos en el dendrograma se ha construido la Cuadro 6 donde se indica la pertenencia del número de informantes por cada conglomerado. Los datos obtenidos no son significativos estadísticamente (χ^2 (a.m. 3) 5,915; $p = 0,116$).

CUADRO 6. Cuantificación de los datos obtenidos en el dendrograma del clúster jerárquico

	LM			L2		
	1	4	TOTAL	2	3	TOTAL
A	7	6	13	2	6	8
B	3	4	7	8	4	12

Tal y como se indica en la tabla las informantes con LM español y lengua de reproducción

español, se han clasificado 7 en el grupo A y 3 en el grupo B. Asimismo, las informantes con LM español y lengua de reproducción euskera, se han clasificado 6 en el grupo A y 4 en el grupo B. En el caso de las informantes con LM euskera y lengua de reproducción euskera, se han encontrado 2 en el grupo A y 8 en el grupo B. Por último, aquellas informantes con LM euskera y lengua de reproducción euskera, se han situado 6 en el grupo A y 4 en el grupo B.

3.2.2. K-medias

Para la aplicación de este método de conglomeración, se ha optado por crear dos grupos en base al análisis jerárquico. En la Cuadro 7 se muestra el número de informantes por conglomerado.

CUADRO 7. Cantidad de informantes en cada grupo tras la realización del método

K-medias

	LM			L2		
	1	4	TOTAL	2	3	TOTAL
A	3	4	7	8	4	12
B	7	6	13	2	6	8

Como se ve en el Cuadro 7 las informantes con LM español y lengua de reproducción español, se clasifican 3 en el grupo A y 7 en el grupo B. Aquellas informantes con LM español y lengua de reproducción euskera, se han clasificado 4 en el grupo A y 6 en el grupo B. En el caso de las informantes con LM euskera y lengua de reproducción euskera, se han encontrado 8 en el grupo A y 2 en el grupo B. Por último, aquellas informantes con LM euskera y lengua de reproducción euskera, se han catalogado 4 en el grupo A y 6 en el grupo

B². Con este método creamos el gráfico de dispersión de escalamiento multidimensional que se presenta a continuación (ver Figura 3).

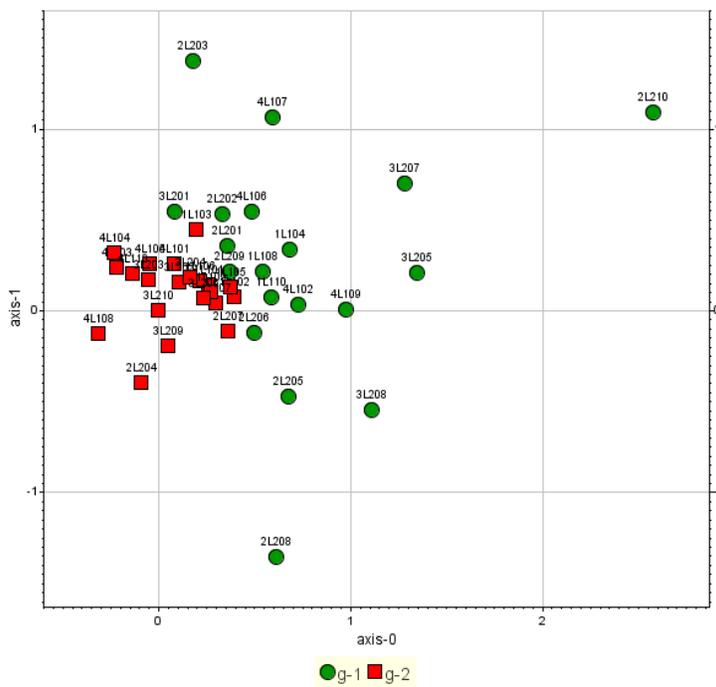


FIGURA 3. Dispersión de los resultados con el método K-medias

En esta figura aparecen los dos grupos con diferente iconicidad y color: uno de ellos está representado por un cuadrado rojo y el otro por un círculo verde. Además, podemos observar cómo el grupo verde es más disperso que el rojo.

3.2.3. Fuzzy C-means (FCM)

Según el método de conglomeración basado en la lógica difusa (FCM) utilizando dos grupos de conglomeración, obtenemos los datos de la tabla 8 donde aquellos elementos con función de pertenencia menor a 0,85 se consideran como elementos del conglomerado, pero de

² Los resultados del análisis KM coinciden plenamente con los resultados del análisis jerárquico (expuesto en el apartado 3.1.).

transición, ya que, una parte de su función de pertenencia está incluida en el otro conglomerado.

CUADRO 8. Clasificación de las informantes según lengua materna y L2 en dos grupos según el método FCM

		LM			L2		
		1	4	TOTAL	2	3	TOTAL
A	$\geq 0,85$	4	6	10	1	6	7
	$< 0,85$	3	0	3	1	1	2
B	$\geq 0,85$	3	3	6	5	3	8
	$< 0,85$	0	1	1	3	0	3

En el Cuadro 8 podemos ver que, en el primero de los grupos (grupo A) y con pertenencia al grupo superior al 0,85, las informantes que respondieron en su lengua materna han sido 10 (4 tenían LM español y 6 euskera). Siguiendo con el grupo A, pero en este caso con las informantes de transición, tenemos 3 que respondieron en su LM (los 3 tenían LM español). Dentro del grupo A, concretamente en las tres últimas columnas, se encuentran las informantes que han respondido a los estímulos en una lengua diferente a su LM. Con una pertenencia $\geq 0,85$ hay una informante de lengua materna español que respondió en euskera y 6 vascoparlantes que respondieron en español. Con una pertenencia inferior a 0,85 hay una informante de lengua materna español que respondió en euskera y 1 vascoparlantes que respondió en español. En lo que respecta al grupo B, las informantes con gran pertenencia al grupo ($\geq 0,85$) son un total de 6 (3 tenían LM español y 3 euskera); y respecto a las informantes de transición en su LM es uno con LM euskera. Dentro del grupo B, en las tres últimas columnas, se encuentran las informantes que han respondido a los estímulos en una

lengua diferente a su LM. Con una pertenencia $\geq 0,85$ hay 5 informantes de lengua materna español que respondieron en euskera y 3 vascoparlantes que respondieron en español. Con una pertenencia inferior a 0,85 hay 3 informantes de lengua materna español que respondieron en euskera y ningún vascoparlante que respondiera en español.

Basándonos en estos datos se ha creado el Gráfico de dispersión que se muestra en la Figura 4 donde podemos ver representados cada uno de los grupos en dos colores y formas geométricas diferentes. El grupo A se muestra por un color grisáceo y círculos, mientras que el grupo B es de color rojo y las informantes están representados por cuadrados. Pero, además de ello, la tonalidad y tamaño de los mismos es diferente en función del grado de pertenencia al grupo. Así, aquellos con mayor tamaño y color más intenso son aquellos con mayor pertenencia al grupo; y cuanto más claros y menor tamaño, tienen menor pertenencia al grupo.

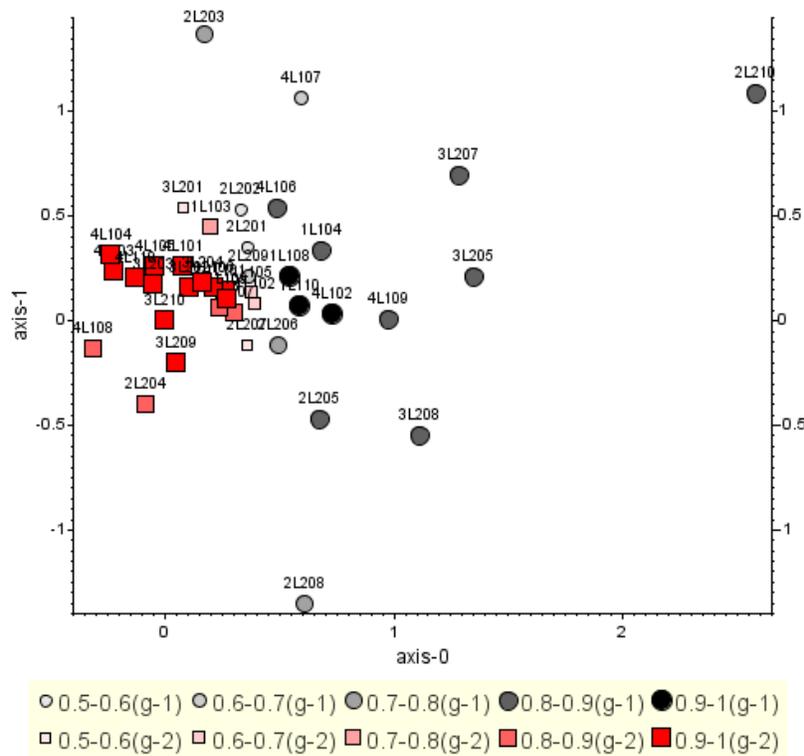


FIGURA 4. Dispersión de los datos tras la aplicación del método FCM

Si comparamos los resultados obtenidos en el análisis de conglomerados jerárquicos y KM con el análisis FCM vemos una diferencia importante en los componentes de cada conglomerado, ya que, en el análisis FCM hay una informante con LM euskera que se clasifica en el otro conglomerado en comparación con los análisis anteriores.

4. CONCLUSIONES Y TRABAJOS FUTUROS

Este estudio se ha llevado a cabo en un territorio en el que conviven dos lenguas: el euskera y el español. Sin embargo, en función de la exposición de las informantes a cada una de las lenguas en contacto, el tiempo de reacción antes los estímulos léxicos se ve influenciado. Esto se ha podido corroborar en este trabajo, donde podemos observar que mediante los tres métodos de conglomeración se crean distintas agrupaciones de los hablantes en función de

las variables objeto de estudio: la lengua materna y la lengua de reproducción.

En los resultados obtenidos se observa que los grupos formados por las informantes son grupos parcialmente heterogéneos. Dentro de esta heterogeneidad se constata que el grupo A, conformado por informantes de lengua materna euskera, muestra mayor homogeneidad en la reproducción léxica en ambas lenguas en comparación con el grupo B formado por hablantes de lengua materna español. Lo que significa que las informantes del grupo B muestran mayor tiempo de reacción en sus producciones léxicas en euskera que en su LM. Mientras que en el caso de las informantes con LM euskera la diferencia entre ambos idiomas no es tan grande.

Por lo tanto, pese a que el contexto es bilingüe, las adquisiciones léxicas no lo son, ya que según cual es la lengua materna y según cual es la lengua de reproducción, varía el tiempo de reacción. Esto es debido a la situación diglósica en la que el español es la lengua hegemónica.

Con todo esto, una de las aportaciones de este estudio es el empleo de los métodos cuantitativos objetivables para el análisis y la clasificación de las informantes aplicado al área de la riqueza léxica; que, como se ha expuesto en la introducción es un campo poco estudiado en el contexto de la lengua vasca.

De cara al futuro vemos oportuno el aumento del corpus y el análisis comparativo entre los resultados obtenidos en función del género de los informantes. Además, los métodos presentados de conglomeración pueden ser aplicables a otros contextos bilingües e incluso expandidos a otras áreas de la lengua.

5. BIBLIOGRAFÍA CITADA

BOERSMA, PAUL & WEENIK, DAVID (2016). Praat: doing phonetics by computer, Version 5.1.

Recuperado de <http://www.praat.org>.

EGUSKIZA, NAIA, DE PABLO, IRATI, & GAMINDE, IÑAKI (2017). Hizkuntz hizkeren sailkapena metodo kuantitatiboen bidez. En ARRIAGA, CRISTINA & ROMERO, ASIER (eds.): *XXIV Jornadas de Investigación en Psicodidáctica*, Bilbao: UPV/EHU, 403-416.

ETXEBARRIA, AINTZANE, GAMINDE, IÑAKI, OLALDE, ANDER, & GAMINDE, URSUA (2016). Hizkuntza aldakortasuna Larrabetzuko aditz morfologian. En IGLESIAS, AITOR, ROMERO, ASIER & ENSUNZA, ARIANE (Eds.): *Linguistic variation in the Basque language and education – II / Euskararen bariazioa eta bariazioaren irakaskuntza – II*, Bilbao: UPV/EHU, 104-119.

GAMINDE, IÑAKI (2010). Bizkaiko Gazteen Prosodiaz: Euskaraz eta Gaztelaniaz. *Mendebalde Kultura Alkartea y Bizkaiko Foru Aldundia*.

GAMINDE, IÑAKI (2011). Noraka Euskal Azentuak?. En AURREKOETXEA, GOTZON & GAMINDE, IÑAKI (Coor.) *Prosodiaz eta Hezkuntzaz I. Jardunaldiak/I. Jornadas sobre Prosodia y Educación*, Bilbao: UPV/EHU, 63-85.

GAMINDE, IÑAKI, & EGUSKIZA, NAIA (2017). Arratiako gazteen aditz laguntzaileen bariazioaz. En IGLESIAS, AITOR & ENSUNZA, ARIANE (Eds.): *Gotzon Aurrekoetxea lagunarterik hara*, Bilbao: UPV/EHU, 141-154.

GAMINDE, IÑAKI, ETXEBARRIA, AINTZANE, EGUSKIZA, NAIA, ROMERO, ASIER & UNAMUNO, LOREA (2016a). Lexikoaren bariazioa eta multzokatze-azterketa, *Euskalingua* 28, 19-3.

GAMINDE, IÑAKI, OLALDE, ANDER, GAMINDE, URSUA & ETXEBARRIA, AINTZANE (2016b): Populazioaren sailkapena lexikoaren arabera, *Ikastorratza, e-Revista de Didáctica*, 16,

68-84.