

**DESCRIPCION LINGUISTICA Y  
SOCIO-LINGUISTICA DEL BILINGUISMO  
ESCOLAR DE EUSKADI-SUR, A TRAVES DEL  
LEXICO DISPONIBLE REFERIDO AL MUNDO  
LABORAL**

**María José Azurmendi Ayerbe**

## INDICE

0. INTRODUCCION.
1. MUNDO CONCEPTUAL DISPONIBLE OBTENIDO EN RELACION AL DOMINIO DE FUNCIONAMIENTO SOCIAL DEL «TRABAJO».
2. «USO LINGUISTICO» DEL LEXICO PRODUCIDO POR LA POBLACION ESTUDIADA, EN RELACION CON EL MUNDO LABORAL.
  - 1) Producción media obtenida respecto a las variables lingüísticas.
  - 2) Correlaciones entre las variables lingüísticas.
3. ESTUDIO GEOLINGUISTICO DEL BILINGUISMO EN NAFARROA.
  - 1) Los grupos o estratos socio-lingüísticos distinguidos.
  - 2) Los subgrupos socio-lingüísticos correspondientes a Nafarroa.

## 0. INTRODUCCION

Este trabajo, realizado en homenaje al euskarólogo D. Angel Irigaray, está basado en el material léxico y sociolingüístico obtenidos primero y tratados después por medio del ordenador, en el proyecto de investigación: «Estudio socio-lingüístico del bilingüismo en la población escolar de Euskadi, a través del análisis de su léxico básico disponible en ambas lenguas», subvencionado en su primera fase por Eusko Jaurlaritza / Gobierno Vasco dentro del concepto de las «becas y ayudas para la investigación en el área vasca», concedido al investigador María José Azurmendi, para su realización en colaboración con los investigadores Angeles Izueta y Francisco José Olarte y el ayudante-colaborador Arantza Egia (además de un equipo de encuestadores).

Las características de este proyecto de investigación: objetivos, métodos, clases de encuestas, muestra, etc., aparecen especificados en la Introducción de la primera fase del proyecto realizado ya, entregado al Gobierno Vasco hace ahora un año (en julio de 1983) y que está en espera de su publicación por parte del Servicio de Publicaciones de la Universidad del País Vasco (en Leioa). Por ello, pensamos no debemos repetirlo aquí.

Solamente, y de modo resumido, decir:

1) Que el material analizado en este trabajo ha sido obtenido en la primavera de 1982, a través de dos tipos de encuestas: la encuesta lingüística y la encuesta socio-lingüística.

A través de la encuesta lingüística se ha obtenido el léxico relacionado con el Centro de Interés (C.I.): «Lana: lan motak, lanbideak, ofizioak... / El trabajo: clases de trabajos, de oficios, de profesiones...», que analizamos en concreto en este trabajo (se trata del C.I. 9, siendo en total 16 los C.I. obtenidos).

A través de la encuesta socio-lingüística se han obtenido gran cantidad de datos sociológico-demográficos de la población a estudiar. Sólo una parte del total de datos utilizamos en este trabajo, los relacionados con:

- Las dos lenguas en contacto: el euskara y el español;
- Los tres grupos de edad y de escolarización: 5.º de EGB, 8.º de EGB y 3.º de BUP + 2.º de FP;
- Los dos tipos o clases de escolarización, en relación con el euskara: las ikastolas y los centros públicos y privados;
- Y los cuatro grupos geolingüísticos: Araba, Bizkaia, Gipuzkoa y Nafarroa (de Euskadi Sur).

2) Que la población-objetivo fue todos los alumnos bilingües de las etapas de 5.º de EGB, 8.º de EGB y 3.º de BUP + 2.º de FP (= Formación Profesional), en toda clase de Centros de enseñanza, de toda Euskadi Sur. Se entendió por «alumno bilingüe» aquel que entendía y hablaba con soltura el euskara (ya que el español lo entiende y habla con soltura toda la población en edad escolar de Euskadi Sur, tanto los monolingües como los bilingües); alcanza, así, a los alumnos que tienen como lengua materna o primera lengua al euskara, como a los que tienen el euskara como segunda lengua, adquirida desde los tres años casi siempre, en las ikastolas, en el método que se llama de «inmersión».

Para poder abarcar la diversidad y amplitud de objetivos sociolingüísticos perseguidos, se ha utilizado una muestra, seleccionada por «muestreo estratificado trietápico», de acuerdo con las siguientes unidades de muestreo por etapa:

- Unidad de la primera etapa: el centro de enseñanza.
- Unidad de la segunda etapa: la unidad o clase escolar.
- Unidad de la tercera etapa: el alumno.

En total, se seleccionaron las encuestas de 4.248 alumnos, de los que se obtuvieron 310 tacos o paquetes completos de encuestas, comprendiendo cada uno dos respuestas (una en euskara y otra en español) de cada alumno de los 16 Centros de Interés obtenidos. Así, para cada C.I. y también para éste referido al mundo laboral, se han utilizado 620 encuestas lingüísticas y socio-lingüísticas distintas; de las 620 encuestas lingüísticas, 310 (la mitad exactamente) para cada una de las lenguas, producidas aproximadamente por 265 alumnos.

## **1 . MUNDO CONCEPTUAL DISPONIBLE OBTENIDO EN RELACION AL DOMINIO DE FUNCIONAMIENTO SOCIAL DEL «TRABAJO»**

A través del léxico obtenido respecto del Centro de Interés (C.I.): «Lana: lan motak, lanbideak, ofizioak... / El trabajo: clases de trabajos, de oficios, de profesiones...». podemos descubrir cuál es el mundo conceptual que la

población estudiada mejor conoce y que mayor interés o expectativas futuras le produce, en relación a este dominio de funcionamiento social del mundo laboral.

Hemos obtenido hasta 1.150 términos de referencia distintos, que suponen una riqueza y heterogeneidad significativas grandes, y sobre todo un conocimiento profundo, por parte de la población estudiada, de la heterogeneidad grande del mundo laboral que les aguarda en la etapa post-escolar.

Damos a continuación el listado de los 73 términos de referencia o clases de trabajo-oficio-profesión, obtenidos de más del 10% poblacional o producidos por más del 10% poblacional, y que por ello pueden ser considerados suficientemente básicos como para que puedan ser incluidos en un diccionario conceptual básico, y también suficientemente representativos de las aspiraciones laborales de la población estudiada.

## CENTRO DE INTERES 9 --- TOTAL

DICIONARIO	OTRA LENGUA	IND.DISP.	FREC.ID.	FREC.REL.	FREC.ARS.	RANG.ASO.	COEF.VAR.	FREC.RA.
MEDIKUA	MEDICO	70	7	26.15	2853.40	10.42	86.12	7-0001
MAISUA	MAESTRO	46	4	16.78	1831.70	7.24	115.06	4-0001
IGELTSARJA	ALRAZIL	44	4	16.44	1794.10	10.40	101.41	4-0002
ZURGINA	CARPINTERO	40	4	15.28	1667.20	10.58	112.09	4-0003
MEKANIKARIA	MECANICO	48	4	17.67	1928.60	11.11	96.01	4-0004
ONDARIA	TENDERO	41	4	14.91	1626.80	12.11	80.15	4-0005
ELEKTRIKARIA	ELECTRICISTA	45	4	16.36	1785.30	14.91	90.76	4-0006
IRAKASLEA	PROFESOR	36	3	14.12	1541.30	10.05	111.49	3-0001
BIHEGARIA	OFICINISTA	33	3	12.15	1325.40	12.42	99.87	3-0002
KALF-GARBITZAILFA	BARRENERO	30	3	11.15	1217.20	15.46	84.21	3-0003
ERIZAINA	ENFERMERO	31	3	11.82	1289.50	16.00	76.05	3-0004
FUTROLARJA	FUTROLISTA	36	3	13.09	1428.30	16.42	78.84	3-0005
ARRANTZALEA	PESCADOR	38	3	13.98	1525.60	16.86	76.49	3-0006
POLIZIA	POLICIA	35	3	12.71	1387.20	18.29	79.21	3-0007
PINTATZAILFA	PINTOR	32	3	12.13	1323.80	19.50	72.99	3-0008
NEKAZARIA	AGRICULTOR	32	3	12.07	1317.60	19.63	76.01	3-0009
ILEAPANTZAILFA	PELUQUERO	30	3	11.25	1222.60	19.45	72.72	3-0010
HAROTZA	HERRERO	25	2	9.36	1021.70	8.37	124.44	2-0001
LANGILEA	OBRERO	29	2	12.40	1353.60	12.45	84.15	2-0002
OKINA	PANADERO	24	2	8.96	978.10	13.24	78.58	2-0003
ARKITIKTOA	ARQUITECTO	28	2	10.27	1121.00	14.13	87.37	2-0004
INGENIARIA	INGENIERO	25	2	10.23	1116.70	14.34	91.91	2-0005
KAMIOLARIA	CAMIONERO	28	2	10.14	1104.00	15.04	86.35	2-0006
ITURGINA	FONTANERO	24	2	8.73	952.80	15.36	73.82	2-0007
HARAKINA	CARNICERO	25	2	9.35	1020.70	15.68	81.35	2-0008
IDAZKARIA	SECRETARIO	26	2	9.88	1072.70	16.40	84.29	2-0009
AROKALUA	AHOGADO	28	2	10.07	1096.80	17.40	79.24	2-0010
SUKALDARJA	COCTINERO	21	2	7.82	853.00	19.11	69.54	2-0011
TAXISTA	TAXISTA	25	2	9.19	1002.40	19.25	72.51	2-0012
ZERBITZARIA	CAMARERO	23	2	8.24	898.70	19.62	70.30	2-0013
ABESIARJA	CANTANTE	20	2	7.46	814.40	20.21	67.69	2-0014
ZAPATARJA	ZAPATERO	25	2	9.01	983.60	20.33	70.17	2-0015
IXOFFERKA	CONDUCTOR	23	2	8.89	970.50	21.01	73.82	2-0016
SALTZAILFA	DEPENDIENTE	22	2	10.70	1169.10	21.54	75.96	2-0017
ANDERIZO	MAESTRA DE IKASTOLA	15	1	5.39	587.20	4.87	113.06	1-0001
BASERRITARJA	CASERO	14	1	5.09	555.40	12.54	71.95	1-0002
SUHILTZAILFA	BOMBERO	15	1	5.63	614.10	14.66	87.15	1-0003
ZARAMARIA	HAJIRERO	13	1	4.95	540.40	17.06	80.19	1-0004
ARRAINSALTZAILFA	PESCADERO	12	1	4.42	482.20	17.14	68.19	1-0005
MAKINARIA	MAQUINISTA	12	1	4.54	495.80	17.85	63.04	1-0006
ARTZAINA	PASTOR	17	1	6.24	680.90	17.96	67.32	1-0007
TABERNARIA	TABERNERO	15	1	5.69	621.00	18.49	63.85	1-0008
PELOTARIA	PELOTARI	11	1	4.10	447.50	18.54	77.63	1-0009
ETXEKOANDREA	AMA DE CASA	12	1	4.32	471.40	18.68	61.21	1-0010
TELEFONARIA	TELEFONISTA	11	1	4.05	442.10	18.94	72.42	1-0011
DELINEATZAILFA	DELINEANTE	12	1	4.46	486.40	19.04	69.52	1-0012
AERONESKA	AZAFATA	18	1	6.44	703.00	19.04	80.97	1-0013
GOZOGILEA	PASTELERO	11	1	4.09	446.50	19.22	84.61	1-0014
MEATZARIA	MINERO	15	1	5.58	608.50	19.63	72.01	1-0015
KAZETARIA	PERIODISTA	17	1	6.30	687.00	20.02	67.99	1-0016
ZUZENDARIA	DIRECTOR	19	1	7.34	801.00	20.24	68.54	1-0017

## CENTRO DE INTERES 9 --- TOTAL

DICCIONARIO	OTRA LENGUA	IND_DISP.	FREC.ID.	FREC.REL.	FREC.ABS.	RANG.ASO.	COEF.VAR.	FREC.RA.
KIMIKARIA	QUIMICO	10	1	3.46	377.60	20.46	64.93	1-0017
POSTARIA	CARTERO	11	1	4.18	455.90	20.75	77.41	1-0018
ATEZAINA	PORTERO	16	1	6.19	675.60	21.14	60.08	1-0019
DENTISTA	DENTISTA	19	1	6.91	754.00	21.60	67.93	1-0020
TOAZLEA	ESCRITOR	19	1	7.02	766.20	22.24	67.84	1-0021
JOSTUNA	SASTRE	13	1	4.80	523.40	22.35	55.41	1-0022
LOREZAINA	JARDINERO	11	1	4.00	436.00	22.42	68.46	1-0023
MARINELA	MARINERO	10	1	3.76	410.80	22.68	60.49	1-0024
ZIKLISTA	CICLISTA	12	1	4.33	472.90	23.21	64.87	1-0025
AKTOREA	ACTOR	17	1	7.41	808.90	23.49	53.30	1-0026
TEKNIKOA	TECNICO	10	1	3.79	413.30	23.51	72.45	1-0027
APAIZA	CIURA	18	1	6.76	737.40	23.62	68.39	1-0028
BANKARIA	EMPLEADO DE BANCA	17	1	6.39	697.60	23.70	67.36	1-0029
ALKATEA	ALCALDE	12	1	4.28	466.60	23.90	70.42	1-0030
ESKULTOREA	ESCULTOR	10	1	3.83	418.40	23.98	58.69	1-0031
RIZARGINA	BARBERO	10	1	3.78	412.30	23.98	62.60	1-0031
PILOTTA	PILOTO	15	1	6.14	670.40	24.47	56.90	1-0032
MUSIKERJA	MUSICO	10	1	3.59	391.60	25.55	60.78	1-0033
KIROLARIA	DEPORTISTA	10	1	3.89	424.10	25.58	54.65	1-0034
JOSKINA	MODISTA	12	1	4.71	514.20	27.24	61.84	1-0035
RELOJARJA	RELOJERO	11	1	3.88	423.00	27.47	54.01	1-0036
AHELIZAINA	GANADERO	11	1	3.88	423.20	28.07	57.37	1-0037

Lo que más sorprende en el listado anterior es la gran dispersión con que se han producido, tanto los 1.150 términos de referencia distintos obtenidos en total, ya que los 73 términos de referencia que han sido producidos por lo menos por el 10% poblacional no suponen más que el 6% del total de términos de referencia, como también la gran dispersión de estos mismos 73 términos de referencia, ya que excepto la profesión de: MEDIKUA/MEDICO, que ha sido producido por más del 70% poblacional (constituyéndose, así, en la profesión de mayor interés para la población estudiada, con mucha diferencia), el resto de estos 73 términos de referencia solamente ha obtenido un consenso de hasta el 40% poblacional, siendo su distribución de:

- 6 términos de referencia: producidos por el 40% poblacional,
- 10 términos de referencia: producidos por el 30% poblacional,
- 17 términos de referencia: producidos por el 20% poblacional,
- 39 términos de referencia: producidos por el 10% poblacional.



## 2. «USO LINGUISTICO» DEL LEXICO PRODUCIDO POR LA POBLACION ESTUDIADA, EN RELACION CON EL MUNDO LABORAL

El «uso lingüístico» del léxico obtenido se ha estudiado a través de cinco tipos de variables lingüísticas distintos, que vienen definidas en el «Diccionario de cuestiones y de modalidades de respuesta» que damos a continuación:

DICCIONARIO DE CUESTIONES Y DE MODALIDADES DE RESPUESTA		
1/	TOTAL DE PERSONAS/TOTAL DE OCURRENCIAS	1PE*
2/	TOTAL DE NOMBRES/TOTAL OCURRENCIAS	2NO*
3/	SUSTANTIVACION/TOTAL DE OCURRENCIAS	2SI*
4/	VERBOS/TOTAL DE OCURRENCIAS	2VE*
5/	SUSTANTIVOS EN UN CASO/TOT. OCURRENCIAS	2CS*
6/	VERBOS+SUSTANTIVOS/TOTAL OCURRENCIAS	2VA*
7/	VERBOS+SUSTANT. EN UN CASO/TOT. OCURRF.	2VC*
8/	ORACIONES/TOTAL OCURRENCIAS	2OR*
9/	EXPRES. FONOLOGICTF. CASO IGUAL/OCURREN.	3IG*
10/	EXPRES. FONOLOGICTE DIFERENTES/OCURRENC.	3DI*
11/	EXPRES. DE OFICIOS RESPEC. CULTURA/OCURP.	5CI*
12/	EXPRES. DE OFICIOS RESPEC. COMERCIO/OCURP.	5CO*
13/	EXPRES. RESPEC. EMPRES. -JEFES/OCURREN.	5EJ*
14/	EXPRES. RESPEC. RELIGIOS=MILITAR/OCURP.	5RM*
15/	EXPRES. RESPEC. UNIVERSITARIOS/OCURREN.	5UN*
16/	EXPRES. RESPEC. AGRIC-GANAD-PESCA/OCURP.	5GP*
17/	EXPRES. RESPEC. AMAS CASA=SERV. DOM/OCUP.	5SD*
18/	EXPRES. OFIC. INDUS. Y ARTESAN./OCURP.	5OF*
19/	EXPRES. RESPEC. ADMON. PUB=PRIV./OCURRENC.	5AD*
20/	RESTO SIGNIFICATIVO/TOTAL OCURRENCIAS	5BR*
21/	EXPRES. INCORRECTAS ORTOGRAF-FONET/OCURP.	4OF*
22/	EXPRES. INCORRECTAS EN -A OMGAMIC/OCURP.	4DE*
23/	EXPRES. INCORRECTAS MORFOLOGICAS/OCURP.	4MO*
24/	EXPRES. CORRECTAS O NORMATIVAS/OCURP.	4PN*
NUMERO DE OBSERVACIONES 250 NUMERO TOTAL DE CUESTIONES 24 NUMERO TOTAL DE MODALIDADES 24		

Resumiendo, estos cinco tipos de variables lingüísticas son:

- El tipo 1: de la cantidad de producción media de léxico;
- El tipo 2: de la clase léxica o grado de lexicalización, respecto del léxico obtenido. Hablamos de «grado de lexicalización» porque una parte del léxico no ha sido palabras, sino verdaderas perífrasis, aunque con el mismo valor de denominaciones. Algunos ejemplos de perífrasis:

- . **2VN** (= verbo + sustantivo o nombre):  
ZUHAITZAK INAUSI / PODAR ARBOLES;
- . **2VC** (= verbo + sustantivo utilizado en un caso (aparte de los casos ergativo, absolutivo y acusativo)):  
FUTBOLEAN JOLASTU / JUGAR A FUTBOL;
- . **2OR** (= oración):  
LANDATZEKO ZULOAK EGIN / HACER AGUJEROS  
PARA PLANTAR.

- El tipo 3: de la distancia interlingüística mantenida entre el euskara y el español. **3IG** (= expresiones fonológicamente casi iguales), se refiere casi exclusivamente a préstamos recientes desde el español al euskara, a veces también a préstamos tradicionales, siempre que se hayan mantenido fonológicamente iguales en la lengua de llegada. Ejemplo: MEDIKUA / MEDICO.

En este C.I. respecto del mundo laboral, prácticamente toda la producción léxica ha sido o **3IG** o **3DI** (= fonológicamente diferente). No se ha dado ni **3PA** (= fonológicamente parecido), que respondería a la mayoría de los préstamos tradicionales, en que se ha dado una adaptación fonológica del préstamo a las características lingüísticas de la lengua de llegada, el euskara, como en:

DENDARIA / TENDERO;

tampoco se ha dado **3NP** (= nombre propio), como en otros C.I. estudiados.

- El tipo 4: de la corrección o normativización lingüísticas.

En este C.I. hemos estudiado estos usos incorrectos o no-normativos: **4OF** (= incorrección ortográfico-fonética); **4DE** (= incorrección debida a la no-utilización de la determinación en los sustantivos y sustantivaciones); y **4MO** (= incorrección morfológica) aplicado a estos dos usos incorrectos fundamentalmente: o a la mala utilización de los sufijos: FUTBOLISTA, en vez de FUTBOLARIA, o a la no-adaptación fonológico-morfológica de final de palabra, como en: MEDIKOA, en vez de MEDIKUA, de aplicación en ambos casos casi exclusivamente respecto del euskara y casi exclusivamente en relación con los préstamos.

Prácticamente no se han dado incorrecciones sintácticas (**4SI**), como en otros C.I., por lo que no hemos incluido esta clase de variante lingüística aquí.

- El tipo 5: de las divisiones significativas o semánticas distinguidas en este C.I., las clases de trabajos-oficios-profesiones distinguidos en el mundo laboral en este trabajo y que por eso constituyen las variantes lingüísticas más específicas en cada C.I. distinto estudiado. Las clases o variantes distinguidas son:

**5CU** = cultura (pintura, deporte, canción...)

**5CO** = comercio

**5EJ** = empresarios, ejecutivos, jefes

**5RM** = religiosos, militares

**5UN** = universitarios (profesiones liberales)

**5GP** = agricultura, ganadería, pesca

**5SD** = amas de casa, servicio doméstico

**5OF** = oficios industriales y artesanos

**5AD** = administración pública y privada

**5RE** = resto de trabajos, no clasificables en las anteriores clases.

### 1) Producción media obtenida respecto de las variables lingüísticas.

El cuadro obtenido del ordenador para ello:

DESCRIPCIÓN SUMARIA DE LAS VARIABLES				
VARIABLE	MEDIA	DESVIACION TÍPICA	MÍNIMO	MÁXIMO
1 / 1PE	5,62	1,40	1,23	25,00
2 / 2BU	27,41	5,32	0,00	100,00
3 / 2SU	0,03	1,12	0,00	7,10
4 / 2VE	0,15	0,62	0,00	5,00
5 / 2CS	0,17	0,54	0,00	20,00
6 / 2VN	0,93	1,85	0,00	100,00
7 / 2VC	0,03	0,11	0,00	1,00
8 / 2OR	0,24	1,85	0,00	100,00
9 / 3TG	16,43	24,74	0,00	78,28
10 / 3BI	27,07	29,44	15,94	100,00
11 / 3CI	15,18	3,52	0,00	30,00
12 / 5CO	11,00	2,71	0,00	22,05
13 / 5FJ	2,00	0,83	0,00	9,09
14 / 5RM	6,96	1,76	0,00	15,00
15 / 5UN	21,07	3,75	0,00	50,00
16 / 5GP	5,01	2,17	0,00	16,18
17 / 5SD	1,75	0,70	0,00	9,00
18 / 5OF	30,00	4,00	13,30	200,00
19 / 5AD	4,55	1,14	0,00	11,74
20 / 5RE	3,52	3,63	0,00	38,66
21 / 4GE	9,25	5,74	0,00	75,00
22 / 4DE	5,15	4,44	0,00	56,33
23 / 4MO	4,03	7,01	0,00	50,00
24 / 4FI	40,71	18,82	0,00	100,00

— Respecto del tipo de variables 2: de la clase léxica o grado de lexicalización, hay un gran predominio de nombres o sustantivos (**2SU**) de hasta el 97,81% de la producción total.

El resto de las clases léxicas distinguidas no llegan al 1% de la producción total. Entre ellas, la producción más abundante, del 0,93% ha sido de **2VN** (perífrasis formada por verbo + nombre o sustantivo); y la clase menos abundante, de sólo el 0,03% **2VC** (= perífrasis formada por verbo + sustantivo en un caso), como en los ejemplos dados más arriba.

— Respecto del tipo de variables 3: de la distancia interlingüística, prácticamente toda la producción léxica obtenida, en este C.I., es o fonológicamente distinta respecto de la otra lengua en contacto (**3DI**), manteniendo una distancia interlingüística grande, que llega hasta el 77,97% de la producción léxica total, o fonológicamente casi igual (**3IG**), suponiendo entonces una distancia interlingüística inexistente, casi exclusivamente en relación con préstamos desde el español al euskara, suficientemente recientes

casi siempre como para que se mantengan fonológicamente casi igual en la lengua de llegada, y que llega en este C.I. al 18,43% de la producción léxica total.

En comparación con lo obtenido en otros C.I., podemos decir que la distancia interlingüística mantenida en éste es bastante grande.

Es casi inexistente la producción respecto de otras variables como **3PA** (fonológicamente parecido) y **3NP** (= nombre propio) en este C.I., como hemos dicho más arriba, ya que no llegan más que al 3,6% de la producción léxica total, por lo que no las hemos incluido en este C.I.

Este es el tipo de variables lingüísticas respecto del que se ha dado mayor dispersión, respecto del que la población estudiada se ha comportado más heterogéneamente, con mucha diferencia, ya que las desviaciones típicas llegan hasta el 24,76% para **3IG** (= fonológicamente casi igual). Esto significa:

. Que los términos de referencia lematizados contienen en realidad una gran variedad léxica interna, una gran variedad de expresiones diferentes;

. Y que esta variedad léxica no debe de interpretarse necesariamente como «riqueza» léxica, sino más bien como un grado grande de vacilación léxica en este caso concreto, y respecto de este C.I., vacilación entre el préstamo reciente y el término más tradicional, como en:

MEDIKUA

/ respecto de MEDICO;

SENDAGILEA

o entre el préstamo reciente y el neologismo, como en:

ZIKLISTA

/ respecto de CICLISTA.

TXIRRINDULARIA

— Respecto del tipo de variables 4: de la corrección o uso normativo lingüísticos, se puede decir que, relativamente, en relación con este C.I. concreto, la producción léxica obtenida ha sido producida bastante correcta o normativamente, puesto que **4BI** (= uso lingüístico correcto o bien) llega hasta el 80,71% de la producción léxica total.

La incorrección ortográfico-fonética (**4OF**) es la más abundante, llegando al 9,25 % de la producción total. Después vendría la incorrección por ausencia de determinación en los sustantivos o sustantivaciones (**4DE**), con el 5,35% Y, por último, la incorrección morfológica (**4MO**), que sólo llega al 4,03% y que se refiere, como hemos dicho más arriba, a la mala utilización de los sufijos y a la mala adaptación morfo-fonológica de los finales en los préstamos.

— Y respecto del tipo de variables 5: de las clases o divisiones significativas distinguidas en este C.I., ha habido un claro predominio de **5OF** (= oficios industriales y artesanos), que llega hasta el 30,09% de la producción léxica total. Viene seguido de **5UN** (= universitarios, profesiones

liberales), llegando al 21,07% de la producción total. Con bastante diferencia, vienen **5CU** (= mundo cultural: arte, deporte, literatura...), que llega al 13,38% de la producción total; y **5CO** (= comercio), con el 11,6%

El resto de clases de trabajos-oficios-profesiones está bastante por debajo del 10% de la producción léxica. Entre ellas, destaca sorpresivamente (?) **5RM** (= religiosos, militares). La clase menos producida ha sido **5EJ** (ejecutivos, empresarios, jefes), también sorpresivamente (?), que no llega más que al 2,06% de la producción léxica total.

En la medida en que estos datos se correspondan con la real distribución laboral que se da en Euskadi Sur, podríamos suponer que ha habido un predominio del «conocimiento» de nuestra situación laboral real por parte de la población estudiada aquí; pero en la medida en que no se correspondan, significaría que la población estudiada ha expresado sobre todo sus «intereses y expectativas» laborales futuras o post-escolares. Así, este tipo de variables lingüísticas nos parece de gran importancia desde el punto de vista psico-sociológico, sobre todo de cara a la planificación.

## 2) Correlaciones (o interdependencia) entre las variables lingüísticas

Damos a continuación el cuadro obtenido por medio del ordenador:

MATRIZ DE CORRELACIONES

	1PE	2NO	2SU	2VF	2CS	2VN	2VC	2OR	3TG	3DT	5CU	5CO	5FJ	5RM	5UN	5GP	5SD	5OF	5AD	5RF	
1PE	1,00																				
2NO	-0,57	1,00																			
2SU	0,64	-0,47	1,00																		
2VF	-0,67	-0,54	0,20	1,00																	
2CS	-0,29	-0,57	0,24	0,20	1,00																
2VN	0,13	-0,67	0,07	0,14	0,05	1,00															
2VC	0,11	-0,42	0,22	0,27	0,19	0,54	1,00														
2OR	0,08	-0,64	0,05	0,05	0,03	0,10	0,10	1,00													
3TG	0,70	-0,57	0,68	0,23	0,30	-0,05	0,21	0,04	1,00												
3DT	-0,70	0,32	-0,69	-0,22	-0,30	0,05	-0,22	-0,06	-1,00	1,00											
5CU	-0,38	0,28	-0,05	-0,13	0,01	-0,40	-0,09	-0,04	-0,13	0,13	1,00										
5CO	0,11	0,11	-0,06	0,04	0,01	-0,10	-0,06	-0,07	-0,06	0,05	-0,09	1,00									
5FJ	-0,13	-0,03	0,08	0,10	0,03	-0,10	-0,10	0,08	0,06	-0,06	0,28	-0,34	1,00								
5RM	-0,26	-0,03	-0,17	-0,24	-0,22	0,20	-0,03	0,08	-0,16	0,16	-0,20	-0,04	-0,02	1,00							
5UN	-0,12	0,24	-0,07	0,03	-0,14	-0,26	-0,16	-0,07	0,01	0,00	0,04	-0,44	0,18	0,06	1,00						
5GP	0,68	-0,16	0,58	0,23	0,16	-0,14	0,24	0,08	0,62	-0,63	-0,05	-0,14	0,15	-0,10	0,05	1,00					
5SD	-0,13	0,07	-0,04	-0,16	-0,07	-0,08	-0,17	0,05	-0,03	0,04	0,01	0,04	-0,04	0,08	0,04	-0,10	1,00				
5OF	0,10	-0,03	-0,06	0,00	-0,02	0,16	0,05	-0,05	-0,16	0,17	-0,47	0,16	-0,41	-0,14	-0,48	-0,21	-0,17	1,00			
5AD	0,36	0,16	-0,07	-0,32	-0,23	-0,08	-0,15	-0,03	-0,23	0,23	-0,22	-0,32	0,11	-0,01	-0,27	-0,27	0,21	-0,22	1,00		
5RF	0,50	-0,50	0,12	0,16	0,25	0,52	0,21	0,12	0,14	-0,15	-0,36	-0,13	-0,01	-0,08	-0,31	-0,10	-0,10	-0,04	-0,20	1,00	
4OF	0,29	-0,42	0,53	0,25	0,32	0,19	0,17	0,09	0,29	-0,29	-0,30	0,18	-0,13	-0,20	-0,24	0,51	-0,04	-0,05	-0,40	0,30	
4DF	0,57	-0,20	0,51	0,14	0,19	-0,09	0,12	0,06	0,77	-0,76	-0,07	-0,25	0,12	-0,10	0,16	0,63	-0,10	-0,22	-0,15	0,08	
4MO	0,78	-0,25	0,43	0,44	0,52	-0,04	0,06	0,03	0,78	-0,77	-0,21	0,16	0,00	-0,18	0,10	0,45	-0,05	-0,06	-0,54	0,09	
4RI	-0,87	0,55	-0,60	-0,57	-0,57	0,00	-0,14	-0,07	-0,92	0,93	0,20	-0,01	-0,01	0,19	0,00	-0,63	0,08	0,12	0,32	-0,17	

4OF	4DF	4MO	4RI	
4OF	1,00			
4DF	0,58	1,00		
473	0,70	0,51	1,00	
4RI	-0,86	-0,85	-0,85	1,00

Hemos subrayado los índices de correlación suficientemente significativos y hemos semi-subrayado los índices de correlación no-claramente significativos, pero que sin embargo nos marcan una tendencia comportamental también clara (los índices que llegan al 0,4, pero que no pasan del 0,5).

Solamente vamos a fijarnos en las correlaciones entre las variables lingüísticas de distintos tipos y no entre las variables internas al mismo tipo, porque es lo que ahora nos interesa más.

— Respecto del tipo de variables lingüísticas 1: de la cantidad de producción léxica, a través de la variable **IPE** (= cantidad de producción léxica media por persona):

La producción léxica más abundante, ordenada de más a menos, viene correlacionada con las variables lingüísticas:

. Sobre todo del tipo 4 (de la corrección lingüística): Positivamente con **4OF** (= incorrección ortográfico-fonética), **4MO** (= incorrección morfológica), y **4DE** (= incorrección respecto de la determinación); y correlacionada negativamente con **4BI** (= uso bien o correcto). Es decir, que a mayor cantidad de producción léxica, menor grado de corrección lingüística en el uso lingüístico de dicho léxico, o peor uso lingüístico; ésta es una de las conclusiones sorprendentes, respecto de este C.I.

. Del tipo 3 (de la distancia interlingüística): a mayor cantidad de producción léxica se da también una menor distancia interlingüística; es decir, la mayor productividad léxica viene acompañada de una mayor influencia del español sobre el euskara, a través de préstamos casi siempre recientes, como veíamos más arriba. ¿Cuál podría ser su interpretación?

. Del tipo 5 (de las clases significativas o semánticas propias de este C.I. relacionado con el mundo laboral): la mayor cantidad de producción léxica, la mayor productividad léxica, se da en relación con **5GP** (= agricultura, ganadería, pesca); y la menor productividad léxica, en relación con **5OF** (= oficios industriales y artesanos), paradójicamente, ya que esta clase significativa ha sido la que ha obtenido la mayor cantidad de producción léxica total, como veíamos más arriba. Esto puede interpretarse como que la población estudiada conoce y ha producido más léxico diferente en relación con la agricultura-ganadería-pesca que en relación con los oficios.

— Respecto del tipo de variables 2 (de la clase léxica o grado de lexicalización), las correlaciones de mayor interés son:

. Con las variables del tipo 3 (de la distancia interlingüística): **2SU** (= sustantivaciones) en correlación positiva con **3IG** (= fonológicamente casi igual), como por ejemplo en:

KANTATZEA, en vez de ABESLARIA / CANTANTE  
ABESTU / CANTAR

. Con las variables del tipo 4 (de la corrección lingüística): sobre todo, respecto de **2SU** (sustantivaciones) otra vez, producido casi exclusivamente en euskara, en que se da la mayor correlación negativa respecto de **4BI** (=

bien o correcto) y, por tanto, donde mayor uso incorrecto se ha dado. La clase de incorrección ligada más fuertemente con la sustantivación (**4SU**) es **4OF**, la ortográfico-fonética.

Con las variables del tipo 5 (clases significativas o de trabajos-oficios-profesiones distinguidos) no existen correlaciones significativas. Sólo **5RE** (= resto significativo, el léxico de difícil clasificación en ninguna de las clases estudiadas) viene correlacionada positivamente con **2VN** (= perífrasis de verbo más nombre o sustantivo), y negativamente con **2NO** (= nombres o sustantivos); esto es lo esperable, ya que casi siempre este resto significativo incluye a trabajos no-normalizados socialmente, de realización más esporádica, de modo que todavía sean denominados a través de una descripción más o menos perífrásica y no a través de una lexicalización acabada de un sustantivo o un verbo.

Con las variables del tipo 3 (de la distancia interlingüística): es donde encontramos las más altas correlaciones y precisamente con las variables del tipo 4 (de la corrección lingüística).

Como conocíamos ya por otros trabajos anteriores, también ahora, y respecto de este C.I. relacionado con el mundo laboral, podemos asegurar que: a mayor distancia interlingüística (**3DI**), a menor influencia léxica del español hacia el euskara, mayor grado de corrección lingüística (**4BI**) en el uso de ese léxico; y, a la inversa, a menor distancia interlingüística, a mayor influencia léxica del español sobre el euskara, menor grado de corrección lingüística, peor uso lingüístico. Así pues, el uso incorrecto viene ligado fundamentalmente a los préstamos, y sobre todo a los préstamos recientes o fonológicamente casi iguales en la lengua de llegada, como veíamos más arriba. ¿Podría esto interpretarse como una gran vacilación, una gran falta de criterio lingüístico en la adopción de los nuevos préstamos en el euskara?

La incorrección lingüística más favorecida en estos préstamos en el euskara es la ortográfico-fonética (**4OF**), después la morfológica (**4MO**) y, por último, la de la determinación (**4DE**), aunque con escasa diferencia..

. Con las variables del tipo 5 (de las clases significativas), sólo respecto de **5GP** (= agricultura-ganadería-pesca) puede decirse que existe correlación significativa. Pero nos parece sorprendente el sentido de esta correlación, ya que esta clase significativa es la única que viene ligada a una distancia interlingüística pequeña, a una dependencia léxico-fonológica grande en el euskara respecto del español, en contra de lo que sería más esperable, dado el carácter tradicional y autóctono que suele concederse a este dominio de la agricultura-ganadería-pesca.

Aunque no puede hablarse de correlación significativa, sino sólo de cierta tendencia comportamental (ya que el índice de correlación llega sólo al 0,23), la clase significativa en que la distancia interlingüística se mantiene más, en donde menor dependencia léxico-fonológica se manifiesta en el euskara, es respecto de la clase significativa **5AD** (= amas de casa y servicio doméstico); es decir, el dominio familiar, el trabajo en casa, es el que se manifiesta más



autóctono léxicamente en el euskara. Esto sí entra dentro de lo que esperábamos.

. Con las variables del tipo 4 (de la corrección lingüística), aparte de lo dicho ya, y por tanto en relación con el tipo de variables 5 (de las clases significativas), nos encontramos con que las únicas correlaciones significativas se dan con **5GP** (= agricultura-ganadería-pesca). En esta clase significativa es donde se daba la mayor dependencia léxica en el euskara respecto del español y, a la vez, donde mayor grado de incorrección lingüística va a darse también.

### **3. ESTUDIO GEOLINGUISTICO DEL BILINGUISMO EN NAFA- RROA**

A través de las técnicas informático-estadísticas del análisis multivariante en componentes principales, se han obtenido unos factores que seleccionan y resumen la variabilidad o heterogeneidad lingüísticas más importantes (por más abundantes) producidas en el «uso lingüístico» del léxico obtenido.

El primer factor, o factor 1, es el más importante porque es el que concentra la mayor cantidad de variabilidad lingüística producida, conteniendo, en este caso concreto, un valor explicativo de hasta el 30,64% del total de dicha variabilidad lingüística. El factor 2 concentra un valor explicativo ya solamente del 12,96% del total; y el resto de factores tienen, cada uno, un valor explicativo menor que 10%

## EDICION DE VALORES PROPIOS

SUMA DE VALORES-PROPIOS ACTIVOS

24,00695801

## HISTOGRAMA DE LOS PRIMEROS VALORES PROPIOS

	VALOR PROPIO	PORCENTAJE	PORCENTAJE ACUMULADO	
1	7,35687828	30,64	30,64	*****
2	5,11087227	12,96	43,60	*****
3	2,21306607	9,22	52,82	*****
4	1,47251770	6,13	58,95	*****
5	1,28777695	5,36	64,32	*****
6	1,14584827	4,77	69,09	*****
7	1,09006021	4,54	73,63	*****
8	0,90095213	4,13	77,76	*****
9	0,90645009	3,78	81,54	*****
10	0,76339698	3,18	84,72	*****

## EDICION SUMARIA DE LOS VALORES PROPIOS DE 11A 24

0,68776029 0,66426808 0,46450186 0,42259206 0,39051652 0,34158844 0,27156794 0,16957325 0,15288454 0,09025598  
 0,00702688 0,00249384 0,00084348 0,00363773

Para la caracterización geolingüística de Nafarroa y de sus dos subgrupos geolingüísticos, vamos a estudiar ahora solamente el factor 1. Este factor 1 y también los demás factores vienen definidos en el siguiente cuadro ofrecido también por el ordenador:

ORDENADAS SOBRE LOS PRIMEROS FACTORES

VARIABLES ACTIVAS	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
1RF	-0,849	0,111	-0,191	-0,056	-0,106	-0,181
2ND	0,511	-0,553	-0,526	-0,026	-0,210	-0,107
2SU	-0,676	-0,090	-0,425	0,114	0,491	-0,232
2VF	-0,420	0,160	0,048	0,440	-0,311	0,221
2CB	-0,411	0,119	0,047	0,410	0,188	-0,167
2VK	-0,138	0,716	0,436	-0,154	-0,122	-0,157
2VC	-0,296	0,327	0,286	0,244	-0,113	0,185
20W	-0,181	0,207	0,448	-0,145	0,451	-0,152
3IG	-0,918	0,251	-0,024	-0,112	0,162	-0,180
4BI	0,921	0,248	0,011	0,113	-0,070	0,279
5TL	0,271	-0,531	0,101	0,518	0,370	-0,001
5CC	-0,007	0,265	-0,665	0,076	0,376	0,268
5FJ	-0,010	-0,406	0,475	0,301	-0,024	0,155
5KK	0,219	0,094	0,216	-0,406	-0,051	0,455
5JH	0,106	-0,602	0,278	-0,186	-0,428	0,076
5GP	-0,652	-0,325	0,105	0,105	-0,301	0,262
5SC	0,120	-0,140	0,014	-0,384	0,545	-0,102
5DF	0,055	0,585	-0,441	-0,046	-0,165	-0,184
5AD	0,398	-0,526	0,335	-0,186	0,437	-0,421
5RE	-0,300	0,554	0,357	-0,041	-0,051	-0,303
4DF	-0,881	0,125	-0,167	-0,109	0,121	-0,155
4BE	-0,756	-0,375	0,104	-0,162	-0,099	-0,156
4MB	-0,819	-0,114	-0,232	-0,004	-0,016	0,114
4FI	-0,935	0,183	0,095	0,107	-0,062	0,014

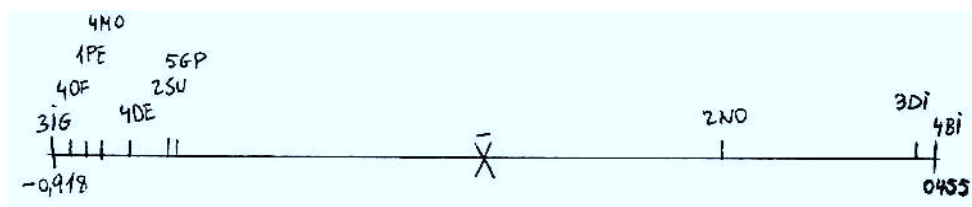
Resumiendo, dicho factor 1, el que incluye el 30,64% de la variabilidad lingüística total producida en el uso lingüístico del léxico obtenido, viene definido sobre todo por las siguientes variables lingüísticas, ordenadas decrecientemente (de más a menos importantes):

— Positivamente, sobre todo, por: **4BI** (= bien o utilización correcta), con un índice de correlación con el factor del 0,955; **3DI** (= fonológicamente diferente, manteniendo una distancia interlingüística máxima), con un índice de correlación del 0,921. Y con mucha menor fuerza: **2NO** (= nombres o sustantivos), con un índice sólo ya del 0,511.

— Y negativamente, sobre todo, por: **3IG** (= fonológicamente casi igual, manteniendo una distancia interlingüística mínima, marcando una dependencia léxica en el euskara respecto del español máxima), con un índice de correlación con el factor de hasta el  $-0,918$ ; **4OF** (= incorrección ortográfico-fonética), con un índice del  $-0,881$ ; **1PE** (= proporción de personas necesarias/cada expresión léxica), con un índice del  $-0,849$ ; **4MO** (= incorrección morfológica), con un índice del  $-0,819$ ; **4DE** (= incorrección por el no-uso del determinante), con un índice del  $-0,756$ . Y ya con menor fuerza, por las variables: **2SU** (= sustantivaciones), con un índice del  $-0,676$ ; y **5GP** (= agricultura-ganadería-pesca), con un índice del  $-0,652$ .

Es decir, que este factor viene definido sobre todo por las variables del tipo 3 de la distancia interlingüística y del tipo 4 de la corrección lingüística, como preveíamos por lo dicho más arriba; se trata de las variables más ligadas

al uso lingüístico. Además, está la variable del tipo 1: de la cantidad de producción léxica; el que **1PE** aparezca en el polo negativo significa que a mayor productividad léxica cuantitativa, peor es dicha producción léxica, como ya esperábamos por lo que conocemos de trabajos anteriores, por inesperable que nos parezca. Por último, los sustantivos aparecen bastante bien utilizados, tanto desde el punto de vista de la distancia interlingüística como de la corrección lingüística; mientras que las sustantivaciones y el léxico en general relacionado con la agricultura-ganadería y pesca, mal utilizados, tanto desde el punto de vista de la distancia interlingüística como de la corrección lingüística. Esto podríamos representarlo gráficamente así:



Respecto de este factor 1 y, por tanto, respecto de las variables que hemos visto lo definen, los distintos grupos y subgrupos sociolingüísticos de interés ahora aparecen representados con los valores o índices siguientes en dicho factor (aparecen también sus valores respecto de los siguientes factores, que no estudiaremos aquí):

## DESCRIPCION LINGUISTICA Y SOCIO-LINGUISTICA DEL BILINGUISMO ESCOLAR

## OBSERVACIONES

	RESOS	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
E1	0.4000	-0.1006	0.0163	-0.0047	0.0091	0.0074	-0.1600
E7	0.4000	-0.0223	-0.3095	-0.0332	-0.1110	0.0024	-0.0006
E2	0.4000	-0.1590	-0.0071	0.0095	-0.1114	-0.0071	0.1044
C5	0.4000	0.1653	0.4666	-0.0004	-0.0001	-0.0004	-0.0032
C7	0.4000	0.4091	-0.0117	0.0011	0.0022	0.0048	0.0004
C2	0.4000	0.1512	-0.1556	0.0002	0.0027	-0.0152	0.0000
E1	0.4000	-0.1505	-0.2995	0.1227	-0.0649	0.0062	-0.1013
EP	0.4000	-0.0971	0.1005	-0.1666	-0.0007	-0.0227	0.1448
C1	0.4000	0.1382	-0.0281	0.2582	0.0021	0.0000	-0.0022
CP	0.4000	0.1381	0.1166	-0.0099	0.0007	-0.1122	0.0034
EA	0.4000	-1.1004	-0.2118	0.6722	0.6221	-0.0053	0.0006
EP	0.4000	-0.1735	-0.3773	-0.0034	-0.0049	-0.0110	0.0006
E6	0.4000	-0.0223	-0.0004	0.0019	-0.0059	0.0009	-0.0071
CA	0.4000	-0.0223	-0.1400	-0.1173	-0.0085	-0.0053	0.1448
CB	0.4000	0.4051	-0.2388	0.1647	0.0033	-0.0069	0.0097
CB	0.4000	0.3766	0.1151	0.1500	-0.0101	-0.0040	-0.0031
CG	0.4000	0.3933	0.0039	-0.1116	0.0083	0.0027	-0.0001
CN	0.4000	0.2866	0.1998	-0.3155	0.1411	-0.1028	0.0085
E41	0.4000	-0.2792	-0.5513	0.4350	-0.5003	-0.2247	-0.3384
E43	0.4000	-1.1334	-0.1900	0.6095	0.7288	-0.0136	1.0026
E41	0.4000	-1.1335	-0.0013	-0.1558	-0.3339	-0.1144	0.1766
E42	0.4000	-1.8005	-0.3608	-0.0144	-0.2290	-0.0175	0.0021
E44	0.4000	-0.0642	-0.5115	0.1449	-0.0021	-0.0085	-0.0054
E45	0.4000	-0.1741	-0.1800	-0.4200	-0.1633	-0.1078	0.1002
E46	0.4000	-0.1740	-0.2822	-0.2600	0.2092	0.1152	0.0029
E47	0.4000	-1.0042	0.1628	-0.0000	-0.0039	-0.0211	-0.1433
E41	0.4000	-0.1735	-0.0054	0.0029	0.0220	0.0015	-0.1031
E42	0.4000	-1.1911	0.2643	-0.1116	0.0600	0.1608	-0.4766
E43	0.4000	-0.6008	-0.0026	-0.0051	-0.1089	-0.1065	-0.0047
E44	0.4000	-0.1748	0.1355	0.4775	0.0076	0.0086	-0.0273
E45	0.4000	-0.5066	-0.0054	-0.1005	-0.0053	0.0020	-0.1400
E46	0.4000	-0.4049	-0.0039	-0.2549	-0.2333	0.0008	-0.0054
E47	0.4000	-0.0647	-0.2778	-0.1116	-0.1666	0.1144	-0.2400
EN1	0.4000	-1.0049	0.0041	-0.1085	-0.0043	0.0031	0.1006
EN2	0.4000	-0.6004	-0.3335	-0.1662	0.0038	-0.1025	0.1085
C41	0.4000	0.1521	-0.5889	0.4200	-0.2884	-0.1025	-0.2773
C43	0.4000	0.4427	-0.2178	0.1332	0.0018	-0.0000	0.0004
C41	0.4000	0.1107	0.8153	0.0015	-0.0115	-0.1071	0.2209
C42	0.4000	0.4039	-0.0031	-0.0024	-0.0014	-0.0193	-0.0043
C44	0.4000	0.3555	0.2773	0.4004	-0.1003	-0.0032	-0.0037
C45	0.4000	0.4000	-0.0056	-0.0070	-0.0053	-0.0036	0.1006
C46	0.4000	0.3570	0.0062	-0.1023	0.0052	0.0017	-0.1006
C47	0.4000	0.3095	0.0087	-0.2200	-0.1442	-0.1013	-0.0045
C41	0.4000	0.4014	0.1334	-0.2996	0.0023	0.0013	0.1000
C42	0.4000	0.3097	-0.0054	0.0044	0.1075	-0.0095	0.0018
C43	0.4000	0.4005	-0.1599	-0.0045	0.1078	-0.0054	0.0032
C34	0.4000	0.2588	0.4094	-0.0088	-0.2109	0.0073	0.0002
C45	0.4000	0.4478	-0.0065	-0.1016	0.1330	0.0054	-0.0025
C46	0.4000	0.4482	0.0028	-0.1028	0.1008	0.0044	-0.1000
C47	0.4000	2.8553	43.5556	-40.3222	-4.6882	-19.2774	-20.0020
CM1	0.4000	0.2644	0.3336	-0.4200	0.0067	-0.0022	0.0018
CM2	0.4000	-0.2649	0.0037	-0.1003	0.2225	-0.0268	0.1600
F	0.4000	-0.0066	-0.1600	0.0014	-0.0035	0.0018	-0.0005
Y	0.4000	0.3082	0.0076	-0.0008	0.0017	-0.0000	0.0004
X	0.4000	-0.1009	-0.3007	-0.0077	0.0000	0.0015	-0.0008
Z	0.4000	-0.0041	-0.1007	-0.0007	-0.0028	0.0000	0.0001
Z	0.4000	-0.8077	1.5010	2.2227	0.1071	0.0019	0.6003
L	0.4000	0.0047	-0.1001	0.0085	-0.0007	0.0047	-0.0053
P	0.4000	-0.0041	0.1001	-0.1022	0.0003	-0.0009	0.0070
R	0.4000	0.3049	-0.2333	0.2002	0.1005	-0.0044	0.3006
R	0.4000	0.0004	-0.1022	0.0000	-0.0002	-0.0044	-0.0005
G	0.4000	0.0006	0.0024	-0.0044	0.0029	0.0054	-0.0028
M	0.4000	-0.1004	0.0079	-0.2065	0.0089	-0.1001	0.1007
X1	0.4000	0.0043	-0.5061	0.4024	-0.0002	-0.0297	-0.3014
X3	0.4000	0.0049	-0.2111	0.2773	0.1007	-0.0000	0.3006
P1	0.4000	-0.2097	0.5043	-0.0018	0.0052	-0.5024	0.2012
R4	0.4000	-0.0038	-0.1041	-0.0043	-0.1003	-0.1007	-0.0023
R4	0.4000	-0.0001	-0.0009	0.3012	-0.1025	-0.0050	-0.0051
R5	0.4000	0.0083	-0.0075	-0.1071	-0.0080	0.0077	0.1003
R6	0.4000	-0.0030	-0.0058	-0.1044	0.1035	-0.2000	-0.1006
R7	0.4000	-0.0024	0.1006	-0.1035	-0.0088	-0.2000	0.0011
R1	0.4000	0.0047	0.1008	-0.1001	-0.0058	0.0080	0.0015
R2	0.4000	-0.1023	-0.0012	0.0003	0.4017	0.0018	-0.1000
R3	0.4000	0.0079	-0.1016	-0.0048	0.0060	-0.0091	0.0006
R4	0.4000	-0.2044	0.3022	0.1002	-0.0075	0.0082	-0.1007
R5	0.4000	0.0044	-0.0041	-0.1013	0.0078	0.0045	0.3035
R6	0.4000	-0.1004	-0.0004	-0.1001	0.0000	0.2000	-0.1000
R7	0.4000	-0.1028	0.0054	0.0002	0.0070	-0.1028	-0.0002
N1	0.4000	-0.2012	0.2034	-0.3038	0.0030	-0.0011	0.0053
N2	0.4000	-0.0097	-0.0097	-0.1082	0.1006	-0.2002	0.1600

## 1) Los grupos o estratos socio-lingüísticos distinguidos

— Los dos grupos, según la lengua utilizada:

**E** (= producción léxica total en euskara) aparece representado en el factor 1 con un índice negativo del  $-0,686$  (recordar que el límite negativo de dicho factor tenía un índice del  $-0,918$ );

**C** (= producción léxica total en castellano o español) aparece representado en el factor con un índice del  $0,382$  (cuando el límite positivo de dicho factor tiene un índice del  $0,955$ ).

Esto significa que la producción lingüística en euskara ha sido, en total, bastante mala, ya que viene representada en el factor con un índice relativo negativo bastante alto. Se trata de una producción lingüística mala respecto de las variables que definen el factor: escasa distancia interlingüística, o dependencia léxica grande, y uso incorrecto; y, por otra parte, la mayoría o casi exclusividad de las sustantivaciones y el léxico relacionado con la agricultura-ganadería-pesca vienen ligadas fundamentalmente al euskara. Aunque parezca paradójico, el euskara viene ligado a una productividad cuantitativa de léxico grande; así, aunque el euskara es malo, la población estudiada produce más léxico en euskara que en español y con bastante diferencia, como si se sintiese más como en su casa.

La producción lingüística en español, a la inversa, es mejor, manteniendo una distancia interlingüística respecto del euskara bastante grande, y un grado de corrección lingüística también bastante grande; por otra parte, la producción de sustantivos viene ligada sobre todo al español, en contradicción esta vez con lo que esperábamos, y a lo que conocemos en otros C.I. diferentes. Pero estos comportamientos son menos radicales en el español que lo que antes hemos visto en el euskara, ya que el índice de **C** es bastante menor que el de **E**.

— Los tres grupos de edad:

**X** (= producción léxica total de 5.º de EGB), aparece representado en el factor con un índice negativo del  $-0,199$ ;

**Y** (= producción léxica total de 8.º de EGB) aparece representado con un índice, positivo esta vez, del  $0,091$ ;

**Z** (= producción léxica total de 3.º de BUP y de 2.º de FP) aparece con un índice negativo de  $-0,857$ .

El único grupo que viene representado positivamente, aunque muy débilmente, es **Y**; éste es, así, el único grupo de edad o de escolarización en que su producción lingüística es algo mejor que la de la media poblacional o total y en donde se da una proporción relativa de sustantivos también mayor.

Los otros dos grupos aparecen negativamente representados, que significa que su producción lingüística es peor que la de la media poblacional. Pero **X** es muy débilmente peor, es casi igual que la de la media total; mientras que **Z** (= 3.º de BUP y 2.º de FP) es fuertemente peor, hasta el punto de constituirse

en el grupo sociolingüístico mejor representado negativamente en el factor y, por tanto, el de peor producción lingüística (donde mayor dependencia lingüística se da, sobre todo, respecto del euskara, y donde mayor incorrección lingüística se da también), aunque sea el grupo que más cantidad media de léxico produzca (como era de esperar).

Además, en **Z** encontramos no sólo la peor productividad lingüística respecto al euskara, sino que también con la peor productividad lingüística en el español, y bastante peor que en la media poblacional, ya que el índice con que aparece **Z** es de  $-0,857$ , mientras que el total de producción en euskara aparecía con el de  $-0,686$ .

Resumiendo, en el paso de 5.º a 8.º de EGB se manifiesta una mejoría en el uso lingüístico que supone pasar del índice negativo al positivo en el factor. Pero en el paso de 8.º de EGB a 3.º de BUP y 2.º de FP se manifiesta un retroceso enorme, un empeoramiento enorme en el uso lingüístico, llegando a unos índices negativos tan fuertes que nos indican una situación de uso lingüístico en ambas lenguas, de bilingüismo, desastrosa. ¿Cómo puede explicarse un cambio de sentido y sobre todo con la brusquedad y radicalidad con que se da entre estas dos etapas escolares?

— Los dos grupos sociolingüísticos dados por las dos clases de escolarización fundamentales en relación con el bilingüismo, según se haya dado la escolarización total o parcial en euskara en las ikastolas o se haya dado la ausencia de dicha escolarización en euskara en los centros públicos y privados:

**I** (= producción léxica total del alumnado de ikastolas y liceos), que aparece débilmente representado positivamente en el factor, con un índice de  $0,047$ ;

**P** (= producción léxica total del alumnado de los centros públicos y privados), representado, también débilmente, pero ahora negativamente, con un índice del  $-0,061$ .

La producción lingüística del alumnado de las ikastolas es muy ligeramente mejor que la de la media poblacional o del total, muestran una situación de bilingüismo mejor, en comparación con la producción de los alumnos de los centros públicos y privados. La diferencia entre estos totales de ambos grupos según la clase de escolarización nos parece que es debido, sobre todo, al uso del euskara, bastante mejor por parte de los alumnos de las ikastolas que de los centros públicos y privados. Lo que nos sorprende es que la distancia relativa entre ambos grupos sociolingüísticos sea tan pequeña; de hecho, son los dos grupos sociolingüísticos en general que más se acercan entre sí y que más se acercan también a la producción media poblacional; es algo así como si la escolarización en las ikastolas no hubiese conseguido unas cotas de bilingüización claramente mejores con respecto a la escolarización en los centros públicos y privados, en el uso lingüístico del léxico de este C.I. relacionado con el mundo laboral, en contra de nuestras expectativas. Quizá sea esto, en parte al menos, porque la adquisición y utilización del léxico de

este C.I. concreto no pertenezca al mundo escolar, sino al de la familia y la calle y, por tanto, donde la influencia de la escolarización pueda ser pequeña.

— Los cuatro grupos geolingüísticos de Euskadi Sur.

Tres de los grupos geolingüísticos aparecen con comportamientos lingüísticos totales casi iguales entre sí: los tres débilmente representados en el factor, con índices muy bajos:

**A** (= producción léxica total de la población de Araba), con un índice positivo del 0,049;

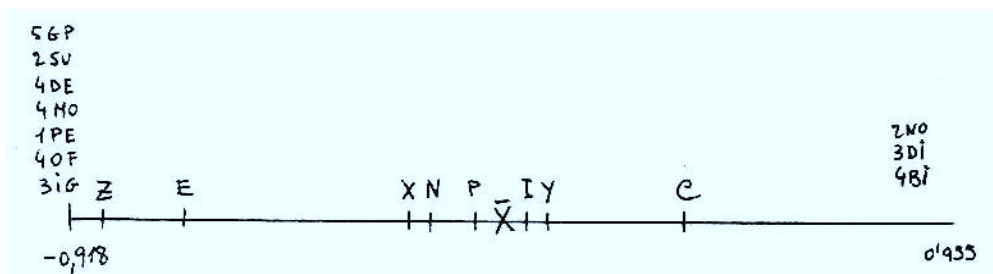
**B** (producción léxica total de Bizkaia), con un índice del 0,004;

**G** (= producción léxica total de Gipuzkoa), con un índice del 0,006.

Mientras que el cuarto grupo: Nafarroa (**N**), es el único grupo geolingüístico que aparece negativamente representado en el factor; su índice es también bastante débil, del  $-0,154$ , pero de todos modos más fuerte que los índices de los otros grupos.

Es decir, la producción lingüística de Nafarroa es: tanto peor que la media poblacional, en contraste con la producción de los otros tres grupos geolingüísticos estudiados, como también el grupo geolingüístico más diferenciado, de comportamiento más específico, con una situación de bilingüismo peor, mientras que los otros tres grupos se nos muestran como casi iguales, con una situación de bilingüismo parecida.

La representación gráfica de todo ello sería así:



## 2) Los subgrupos sociolingüísticos correspondientes a Nafarroa

Hemos subrayado, en el cuadro anterior, los subgrupos correspondientes a N (= Nafarroa), para facilitar su localización. Ahora solamente vamos a analizar algunos de estos subgrupos.

— En relación con el bilingüismo.

El comportamiento claramente más diferenciado en los subgrupos respecto de los cuatro grupos geolingüísticos básicos está en que la producción



lingüística en euskara es muy mala, con índices negativos muy fuertes, mientras que la producción en español es bastante buena, con índices positivos no demasiado altos. En general, pues, puede hablarse de una situación generalizada de bilingüismo muy desequilibrado, en relación a las variables lingüísticas definidoras de este factor.

**EN** (= producción léxica en euskara de Nafarroa) aparece bien representado, negativamente, con un índice del  $-0,886$ . Relativamente, su producción en euskara es mejor que la dada por Araba (**EA**), pero peor que la dada por Bizkaia (**EB**) y por Gipuzkoa (**EG**).

Por una parte esto y por otra parte, que la mejor producción y situación de bilingüismo se den en Gipuzkoa (con todo lo malas que son) nos parece que podría interpretarse como la influencia positiva, respecto de la bilingüización, de la escolarización en las ikastolas, sobre todo a través de una mejora en el uso lingüístico del euskera.

Mientras que **CN** (= producción léxica total en castellano o español, en Nafarroa) es la de peor producción lingüística de los cuatro grupos geolingüísticos, ya que aparece representado en el factor con el índice menor. La mejor producción lingüística en español se da en Araba (**CA**). Entre Gipuzkoa (**CG**) y Bizkaia (**CB**), Gipuzkoa aparece con una mejor producción lingüística, tanto en euskera como en español. Así, Gipuzkoa se muestra como el grupo geolingüístico de mayor grado de bilingüización, de bilingüismo más equilibrado, aunque se trate de un bilingüismo fundamentalmente y en sí muy desequilibrado, siempre a favor del español y en perjuicio del euskara.

— En relación con los dos subgrupos geolingüísticos distinguidos en Nafarroa:

**N1:** Elizondo, Bera, Lesaka, Leitzza, Donestebe, Lekunberri, Goizueta (las poblaciones fronterizas o más cercanas a Gipuzkoa), y

**N2:** Lizarra, Altsasu, Etxarri, Iruñea, Billaba, Zizur-Txiki (las poblaciones más alejadas de Gipuzkoa),

estos subgrupos geolingüísticos aparecen representados negativamente en el factor; esto supone una producción lingüística peor que la de la media poblacional total. Entre ambos grupos, **N1** ( $-0,212$ ) aparece con una producción lingüística peor que **N2** ( $-0,087$ ).

Si queremos saber a cuál de las lenguas puede deberse, sobre todo, este comportamiento distinto y peor, vemos que:

. **CN1** (producción léxica en español, por la zona 1 más cercana a Gipuzkoa) y **CN2** (= producción léxica en español, por la zona 2 más alejada de Gipuzkoa), tienen unos índices positivos, y lo son prácticamente iguales (de  $0,244$  y  $0,249$ ). Además, estos índices son bastante parecidos a los de la producción media total en español (**C**), con un índice del  $0,382$ ; esto significa que la producción en español de ambos subgrupos geolingüísticos navarros es algo peor que la de la media poblacional en su conjunto.

. Pero donde los comportamientos entre ambos subgrupos son claramente diferentes, lo son respecto de su producción léxica en euskara, ya que:

**EN1** (= producción léxica total en euskara, en la zona navarra 1 más cercana a Gipuzkoa) viene representado negativamente con un índice del  $-1,069$ , mientras que

**EN2** (producción léxica total en euskara, en la zona navarra 2 más alejada de Gipuzkoa) viene representado por el factor, también negativamente, con un índice del  $-0,688$ .

Esto significa varias cosas: Que ambos subgrupos geolingüísticos navarros se diferencian entre sí casi exclusivamente en relación al uso del euskara; que en ambos subgrupos la producción de euskara es peor que la media poblacional en general, ya que sus índices son más fuertes que dicha media poblacional respecto del euskara (=E), que tiene un índice del  $-0,686$ , del mismo modo que vemos más arriba que sus producciones en español eran también peores que las de la media poblacional.