

**INTRODUCCIÓN A LOS TRATAMIENTOS Y ANÁLISIS LONGITUDINALES
EN LOS CENSOS DEMOGRÁFICOS**

Jesús Rodríguez Marcos y Enrique Morán



**EUSKAL ESTADISTIKA ERAKUNDEA
INSTITUTO VASCO DE ESTADISTICA**

Donostia-San Sebastián, 1
01010 VITORIA-GASTEIZ
Tel.: 945 01 75 00
Fax.: 945 01 75 01
E-mail: eustat@eustat.es
www.eustat.es

INTRODUCCIÓN A LOS TRATAMIENTOS Y ANÁLISIS
LONGITUDINALES EN LOS CENSOS DEMOGRÁFICOS

Jesús Rodríguez Marcos y Enrique Morán

Indice

INTRODUCCIÓN A LOS TRATAMIENTOS Y ANÁLISIS LONGITUDINALES EN LOS CENSOS DEMOGRÁFICOS	1
INDICE	3
INTRODUCCIÓN.....	4
DEFINICIÓN DE POBLACIONES DE ESTUDIO	5
CALIDAD DE LA INFORMACIÓN DERIVADA DEL ANÁLISIS DE FLUJOS	7
TRATAMIENTOS ESTADÍSTICOS DE LOS MOVIMIENTOS ERRÓNEOS	9
ESTUDIO DE CASOS: LA ACTIVIDAD Y LA LENGUA.....	10
CAMBIOS EN LA RELACIÓN CON LA ACTIVIDAD 1996-2001	11
ESTUDIO DE LA MOVILIDAD LINGÜÍSTICA 1996-2001	17
CONCLUSIONES Y PERSPECTIVAS DE FUTURO	24

Introducción

En orden a mejorar los tratamientos de validación, depuración e imputación de la información de los Censos de Población y Viviendas 2001 –CPV01-, Eustat diseña un proceso de fusión entre los ficheros derivados de la operación citada y los integrados en el Registro de Población de carácter estadístico –RP-. A resultas de estos trabajos se consiguen asignar las claves del citado registro a un 97,5% de las personas de los archivos de los CPV01. En paralelo también se diseñan y ejecutan procesos de fusión de edificios y de huecos (viviendas y locales).

Previamente a estas operaciones, Eustat ya había realizado tratamientos similares de cara a preparar la información de huecos (especialmente viviendas y locales) necesaria para la fase de preimpresión de los Censos de Población y Viviendas de 2001, realizada por el INE.

Definición de poblaciones de estudio

La asignación de las claves del registro permiten asociar la información de los censos del 2001, unidad a unidad, a la contenida en los archivos de la Estadística de Población y Viviendas 1996, ya integrada en el RP.

De la relación entre ambas fuentes de información se derivan tres subgrupos de población que debieran tener la posibilidad de explicar con mayor precisión que a través sólo del procedimiento de los saldos –diferencias de dos stocks poblacionales en dos momentos- los cambios sucedidos en el período intercensal: Altas en el período, bajas y población común o que permanece. Así la población final –en nuestro caso la de 2001-, será el resultado de añadir las altas producidas entre 1996 y 2001 y de restar las bajas del mismo período a la población de partida – 1996-.

Estos subgrupos sólo pueden deducirse o a través de complejos procesos de tratamientos registrales o por derivación a través de un proceso de fusión como el de arriba descrito.

La población que residía en la C.A. de Euskadi en 1996 y que también fue encontrada en 2001, asciende a 1.871.516 personas (el 89,2% de la población de 1996 y el 89,9% de 2001). Entre 1996 y 2001 se produjeron 211.071 Altas –técnicamente población que estaba en los archivos de 2001 y no en los de 1996- y 226.539 Bajas –estaban en 1996 y no en 2001.

Hay que precisar que entre las fechas de referencia censal se producen más movimientos de Alta y Baja que los arriba recogidos. Estos sólo hacen referencia la comparación entre los dos momentos censales citados. Realmente, en los archivos de trabajo se incluyen otras 32.489 bajas adicionales del período, hasta configura un total de registros que ascienden 2.341.615

CUADRO 1 GRUPOS DERIVADOS DE LOS TRATAMIENTOS LONGITUDINALES ENTRE LA ESTADÍSTICA DE POB. Y VIVIENDA 1996 Y LOS CENSOS DE 2001 EN LA C.A. DE EUSKADI. EFECTIVOS OBTENIDOS.

Población 1996	Bajas 1996-2001	Altas 1996-2001	Población 2001	Común 96-01
2098055	226539	211071	2082587	1871516
POB 2001 = POB 96 - BAJAS 96-01 + ALTAS 96-01				2082587
POB 2001 = COMUN 96-01 + ALTAS 96-01				2082587
POB 1996 = COMUN 96-01 + BAJAS 96-01				2098055
POB 1996 = POB 01 - ALTAS 96-01 + BAJAS 96-01				2098055
REGISTROS EN ARCHIVO PARA ANÁLISIS DE FLUJOS				
		REAL	TEÓRICO	
COMUNES 1996-2001		1871516		
ALTAS 1996-2001		211071		
BAJAS CALCULADAS 1996-2001			226539	
BAJAS INCLUIDAS 1996-2001		259028		
TOTAL		2341615	2309126	

Calidad de la información derivada del análisis de flujos

Los análisis de flujos (cambios o transiciones en terminología anglosajona) derivados de estos archivos adolecen de tres fuentes de errores o sesgos: error de cobertura, errores de consistencia y sesgos de tratamiento.

Los censos, como cualquier otro procedimiento de recogida de información, sólo llegan a una parte del total de la población de análisis (se mide la subnumeración y la supernumeración de población y su saldo bruto); en el caso de los archivos preparados para los tratamientos longitudinales, a las carencias de cobertura de ambos censos se suma la ya indicada en el proceso de fusión: un 2,5% de las unidades no pudieron ser asociadas, con lo que pasan a ser Altas, y generan otra tantas Bajas de población de 1996, de las cuales una parte será en realidad población común a ambos censos.

Hasta ahora y a través de las llamadas encuestas de validación o de repetición, se medía la consistencia de las distintas variables obtenidas en una operación previa de recogida. Básicamente todas ellas siguen la metodología desarrollada por Hansen, Hurwits y Bershad en la Oficina del Censo de los Estados Unidos de América. El propio nivel de consistencia de las variables de cada censo determinará necesariamente la calidad de la consistencia y significatividad del movimiento analizado. Aunque parezca obvio decirlo, insistiremos en que bajos niveles de consistencia en variables como la profesión¹, pueden anular la validez de los movimientos.

Además de estos indicadores genéricos de calidad, dentro de los tratamientos de la información censal forman un colectivo sensible de validación y depuración las inconsistencias o incoherencias detectadas en la propia información.

Pero por otro lado, poner en relación dos archivos de información que han sido tratados de forma individual en distintos períodos de tiempo no garantiza la consistencia de la información de cambios que producen: surge un nuevo tipo de información, llámese flujo, cambio o movimiento, que implica la aparición de nuevas inconsistencias o errores. No resulta admisible que una persona en 1996 aparezca como que estaba ocupada y que en 2001 que estaba ocupada y que en 2001 que estaba buscando el primer empleo.

Estas carencias sí que se han tenido en cuenta a la hora de realizar los tratamientos de imputación de no respuesta parcial de los archivos de 2001; además de filtrar los casos manifiestamente erróneos, se han excluido de los procedimientos de imputación longitudinal los cambios muy asociados al tiempo (no se imputa la situación profesional de 2001 en base a la de 1996 para la población de 35 o menos a años, por ejemplo, si no en base a la situación de los jóvenes en 2001), ya que son altamente probables las variaciones en los cinco años del período intercensal, el procedimiento para los casos en que hay que imputar falta de respuesta en 2001, tiende a disminuir el volumen de cambios. Una vía de reducir este efecto consistió en la asignación aleatoria de los

¹ Tomando como indicador el porcentaje de respuestas igualmente clasificadas, en los Censos de 1991, la profesión rondaba el 70% de consistencia.

distintos tipos de movimiento entre 1996 y 2001. No obstante se ha analizado el efecto imputación, tabulando los archivos imputados y sin imputar, dando como resultado variaciones en los principales flujos de análisis de uno a dos puntos porcentuales.

Tratamientos estadísticos de los movimientos erróneos

No obstante, una vez finalizados los urgentes tratamientos de 2001, se plantea el análisis, validación y depuración de los movimientos definidos como erróneos. Aunque se pueden definir con cierta facilidad cambios imposibles (no se puede cambiar de año de nacimiento, de municipio de nacimiento –si no es por anexión o desanexión-, de lengua materna, no se puede pasar de casado a soltero, etc.), resulta problemático tratar los cambios muy improbables pero posibles (cambio de sexo, de buenos niveles de conocimiento de una lengua a la pérdida, de inválido permanente para el trabajo a ser ocupado, jubilado de más de 65 años y ocupado 5 años después, etc.)

Estas dudas obligarán a analizar los grupos de población o los ámbitos espaciales en donde estas variaciones pueden aumentar su probabilidad de verosimilitud: la edad, la actividad y profesión en ambos momentos e incluso anteriores –para el caso de los jubilados- pueden servir para concretar la regla de admisión del cambio. Los casos muy especiales, como el de cambio de sexo, probablemente requieran un estudio previo de fuentes externas.

El segundo problema que se plantea deriva de la dirección temporal en que se deben producir las correcciones. Se puede corregir la información de origen, la de destino o ambas. Previamente hay que aceptar que se puedan modificar los archivos de uno de los censos o de los dos. Esto podría derivar en la posibilidad de obtener cifras distintas a las publicadas. Parece lógico pensar que se van a mejorar ambos ficheros con estos tratamientos adicionales.

Esta circunstancia se acentuaría si se decide asignar valores no imputados en un censo para un individuo, en otro censo en donde sí hubieran sido imputados. Eustat ya desde la EPV del 1996 marca los registros imputados según el origen de la imputación.

Este problema, al igual que el anterior, debiera resolverse a través de tratamientos en dos direcciones: análisis de la coherencia interna de cada censo enfrentada al otro, por un lado, e influencia en la intensidad de los flujos, por otra.

Estudio de casos: la actividad y la lengua

Recordaremos que existe una amplia tradición en las encuestas continuas o de panel y especialmente en las de fuerza de trabajo de este tipo de análisis. Frente a la ventaja de la calidad de la información se encuentran las propias limitaciones muestrales, acentuadas al ser los propios movimientos submuestras.

CUADRO 2 POSIBILIDADES DE ANÁLISIS CON TABLAS LONGITUDINALES

A) Análisis de cambios y de composición según origen

C A M B I O S PERÍODO n PERÍODO n+1

C O M P O S I C I Ó N P O B L A C I Ó N n

		NO 2001	2001				TOTAL 1996	
			V1	V2	V3	V4		
C O M P. P. n+1	NO 1996		A _{1,2001}	A _{2,2001}	A _{3,2001}	A _{4,2001}	A ₂₀₀₁	
	1996	V1	B _{1,2001}	V ₁₁	V ₁₂	V ₁₃	V ₁₄	T _{V1,1996}
		V2	B _{2,2001}	V ₂₁	V ₂₂	V ₂₃	V ₂₄	T _{V2,1996}
		V3	B _{3,2001}	V ₃₁	V ₃₂	V ₃₃	V ₃₄	T _{V3,1996}
		V4	B _{4,2001}	V ₄₁	V ₄₂	V ₄₃	V ₄₄	T _{V4,1996}
	TOTAL 2001	B ₂₀₀₁	T _{V1,2001}	T _{V2,2001}	T _{V3,2001}	T _{V4,2001}		

ALTAS 2001

BAJAS 2001

$$V_{ij} = N(1996 = V_i \wedge 2001 = V_j) \quad \forall i, j = 1, \dots, 4$$

$$B_{i,2001} = N(1996 = V_i \wedge 2001 = \emptyset) \quad B_{2001} = \sum_{i=1}^4 B_{i,2001}$$

$$A_{j,2001} = N(1996 = \emptyset \wedge 2001 = V_j) \quad A_{2001} = \sum_{j=1}^4 A_{j,2001}$$

$$T_{V_j,2001} = \sum_{i=1}^4 V_{ij} + A_{j,2001}$$

$$T_{V_j,1996} = \sum_{j=1}^4 V_{ij} + B_{j,2001}$$

$$T_{2001} = \sum_{j=1}^4 T_{V_j,2001} + A_{2001}$$

$$T_{1996} = \sum_{i=1}^4 T_{V_i,1996} + B_{2001}$$

B) Análisis de Saldos

$$SV1 \text{ 1996-2001} = A_{1,2001} - B_{1,2001}$$

C) Análisis de crecimiento de poblaciones

$$TV1,2001 - TV1,1996 = (A_{1,2001} - B_{1,2001}) + (V_{21} + V_{31} + V_{41}) - (V_{12} + V_{13} + V_{14})$$

UNIDADES: PERSONAS (UN MOVIMIENTO)
MOVIMIENTOS (CAMBIOS EN MÁS DE UN PERÍODO)

PERÍODOS: UNO
VARIOS

En segundo lugar, y según para qué características puede ser crucial, el área espacial de referencia que se va a tomar resulta ser el de toda la C. A. de Euskadi. Resulta plausible pensar que en muchos casos tanto el tipo como el volumen de los flujos puede estar muy influenciado por el ámbito geográfico que se tome, tanto por sus características específicas como por el efecto de los propios cambios migratorios en el período intercensal. Esta circunstancia en los últimos tiempos se ha producido no sólo por el aumento de población extranjera, sino por el gran incremento de variaciones residenciales intermunicipales. Tal es así que las variaciones detectadas en el censo lingüístico entre 1996-2001 a nivel de municipio no pueden entenderse sin analizar previamente los movimientos migratorios.

Hay que decir que introducir la variable espacial en estos análisis, esto es, estudiar varias zonas a la vez –por ejemplo las provincias o territorios de una Comunidad Autónoma- precisaría de una determinación o tipologización de los movimientos posibles: altas por origen y bajas en el período, cambios internos a la zona definida de los sedentarios y cambios en función del origen.

Aunque sea obvio, hay que advertir que el interés del análisis de flujos deriva de la necesidad de medir el efecto del tiempo (de un período) en las características de las personas (o de otras unidades censales como las familias, los edificios, etc.) A su vez, con respecto a las personas, resulta necesario tomar como central una característica asociada al tiempo: la edad o la generación.

No se puede olvidar, también, que en un período de cinco años puede ser suficiente para que en determinados aspectos de las personas se puedan producir varios cambios (no solo de características o situaciones personales, sino de alta y baja en el territorio de análisis), y que los censos sólo nos darían el último. Esta circunstancia, en algunos casos podría producir falsas inconsistencias o flujos muy poco probables.

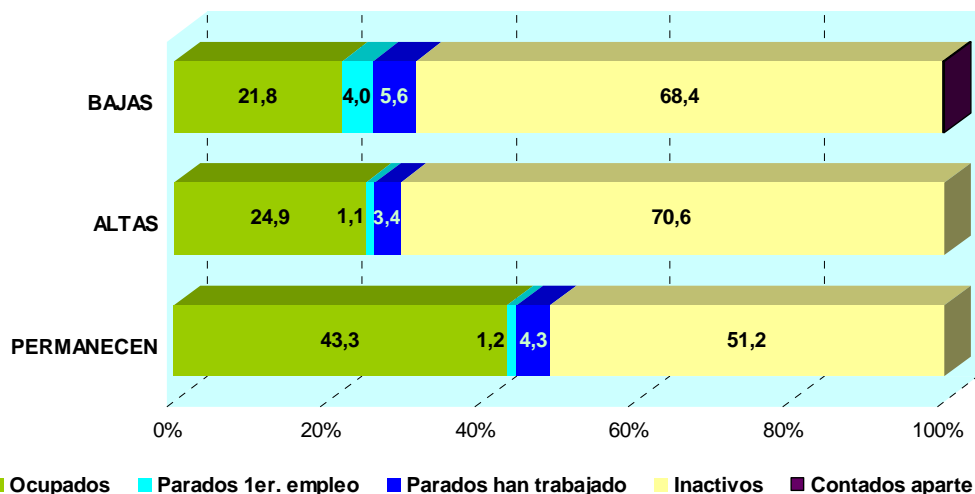
Cambios en la relación con la actividad 1996-2001

La primera aproximación se hará con la variable relación con la actividad con cinco estados: ocupados, parados que buscan 1er. empleo, parados que han trabajado, inactivos y población contada aparte –servicio militar- en 1996.

Para la población que permanece o común a ambos censos el cruce de su relación con la actividad en ambos momentos genera dos inconsistencias o errores: ocupados en 1996 que declaran que en 2001 buscan el 1er. empleo, y parados que han trabajado en 1996 y en 2001 dicen que también buscan el 1er. empleo.

Si se decide corregir el valor de 1996, las poblaciones de partida apenas se verían afectadas (un 0,3% del total de ocupados de 1996 y un 1,1% de los parados que han trabajado, de la población que permanece). Si se toma la dirección contraria, podría afectar a un 13% de los parados que permanecen entre 1996-2001 y un 12% al total publicado de 2001.

Gráfico 1. Distribución de la relación con la actividad según población de alta, de baja y que permanece en el período 1996-2001. C.A. de Euskadi.%.



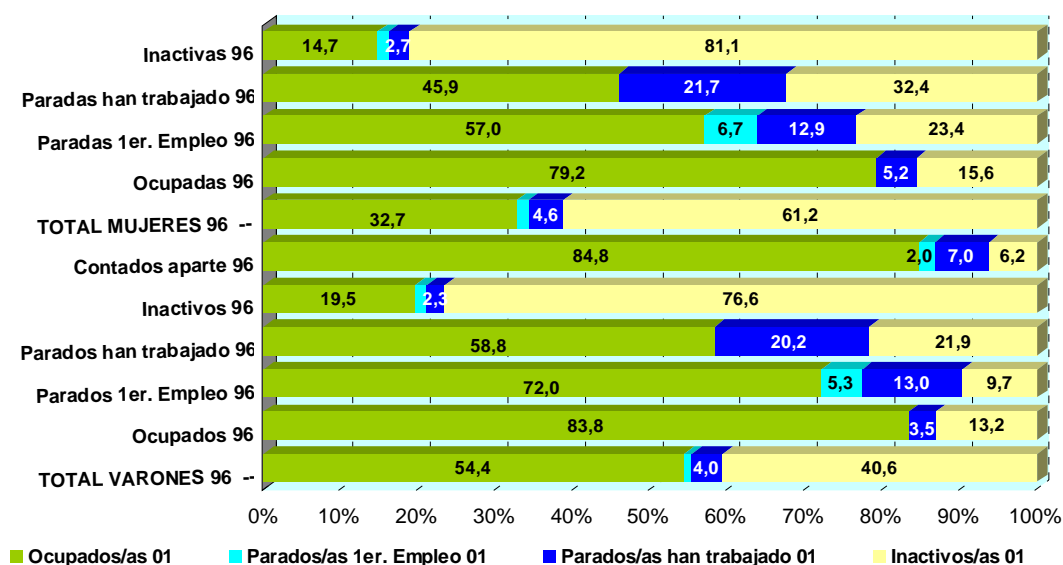
Fuente: Eustat, EPV96-CPV01

La influencia de la inconsistencia detectada obliga a definir, como se comentó más arriba, reglas de validación y de depuración complejas, entre las que no habría que descartar correcciones de flujos a priori consistentes.

Si se pasa a una primera revisión bruta de los resultados, en este caso realizando una simple distribución de la variable en estudio según el tipo de población, empiezan a surgir una serie interesante de explicaciones y aún mayor de hipótesis o preguntas.

Así, aunque cabría esperar una mayor entrada de activos vía altas (inmigración por razones laborales) y una mayor salida de inactivos por vía de las bajas (mortalidad, emigración de retorno, etc.), parece darse un equilibrio y casi un saldo nulo de este colectivo calculado a través de esos dos movimientos. El estudio longitudinal a través de los censos permitiría una mayor determinación y precisión en su análisis (¿Cómo afectan nacimientos y defunciones a la composición de la población según la actividad? ¿Llegan más personas mayores –padres- a la Comunidad que emigrantes de retorno salen de ella?, ¿En el período se ha producido inmigración familiar o individual? ¿Cuáles son las edades del primer colectivo y su situación de jubilación o dependencia?, etc.).

Gráfico 2. Distribución de la población que permanece entre 1996-2001 según los cambios en relación con la actividad y el sexo. C.A. de Euskadi.%



Fuente: Eustat, EPV96-CPV01

Una vez analizados los cambios que han introducido en el stock de población de 1996 los movimientos de altas y de bajas de período, se puede pasar a revisar los movimientos producidos en la características de los individuos que han permanecido (o que al menos estaban al comienzo y al final, esto es, que se censaron en 1996 y en 2001).

La primera cuestión a dilucidar pasa por decidir la dirección temporal del análisis: se puede estudiar qué cambios ha sufrido la población de 1996 según alguna de sus características o se puede analizar la situación de un colectivo concreto en 2001 según su composición en relación a la procedencia.

Así, y volviendo al ejemplo de la población inactiva, resulta viable conocer la composición de ese colectivo en función del origen (no solo geográfico, sino también grupal): saldo de altas y de bajas en el período por un lado y cambios de actividad en los que permanecen –los que pasan de ser activos a inactivos-.

El primer caso, que se puede considerarse más propiamente como análisis longitudinal, ofrece la posibilidad de apreciar los cambios en el tiempo o la permanencia en una misma situación. Así tenemos que el 82,3% de las personas que eran ocupados en 1996, cinco años más tarde permanecían en esa situación, hasta un 6,1% de los parados que buscan el primer empleo siguen haciéndolo y algo más de uno de cada cinco parados que dijeron en 1996 que habían trabajado, dieron la misma contestación en 2001. Como ya se advirtió, la lectura de estos movimientos (o permanencias) debe hacerse con precaución, dado que para la actividad cinco años resulta ser un período largo y pueden haberse producido variaciones intermedias. Aún así puede resultar interesante, dado que los censos lo permiten, precisar las características de los colectivos determinados por sus cambios o por no haber cambiado en los dos momentos en que se han tomado en cuenta.

Como se puede apreciar en el Gráfico 2 las tasas de permanencia en la ocupación resultan ligeramente más bajas para las mujeres que para los varones: un 79,2% frente al 83,4% de los varones, y que también pasan en menor media al desempleo –un 3,5% de los ocupados varones frente a un 5,2% de las mujeres que se definían como ocupadas en 1996-. La tasa de permanencia en la búsqueda del primer empleo resulta próxima en ambos sexos –un 5,3% en varones y 6,7% en mujeres-, dado que cabe esperar una fuerte asociación de esa situación a las edades más jóvenes y una posible mayor igualdad en el acceso al mercado laboral.

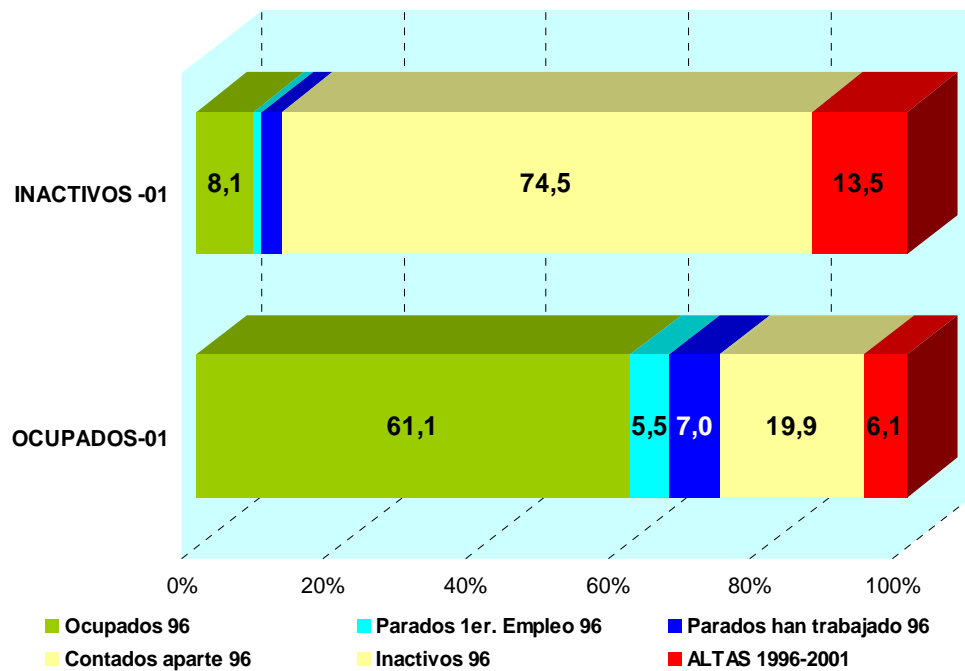
Si se toma el colectivo de los parados que han trabajado y aunque también se aprecia una tasa de permanencia muy similar por sexos -20,2% de los varones y el 21,7% de las mujeres-, si que resulta significativo el mayor paso a la inactividad de las mujeres que en 1996 se definían como paradas que habían trabajado –el 32,4%- frente a los varones en la misma situación –el 21,9%-.

En el Gráfico 3 podemos apreciar el análisis de la composición por procedencia del colectivo de ocupados y de inactivos. Se podría hablar de una tasa de renovación bruta del mercado laboral del 38,9%: uno de cada cinco ocupados –el 19,9%- viene de la inactividad (probablemente jóvenes y/o mujeres), un 12,5% del desempleo y un 6,1% de fuera de la C. A. de Euskadi.

Se podría hablar también del saldo de ocupados por altas y bajas entre 1996 y 2001: frente a 52.551 altas de ocupados se contabilizaron 45.886 bajas, con lo que se obtiene un saldo positivo de casi 7.000 ocupados entre 1996 y 2001. Habría que, no obstante, descontar las bajas por defunción (o disponer del fichero de emigrantes de la C. A. de Euskadi) para construir un ‘saldo migratorio laboral’.

Con respecto a la población inactiva resulta reseñable que hasta el 13,5% viene de altas en el período. Lo mismo que el caso anterior resultaría preciso determinar la razón del alta: los nacimientos o la inmigración. Se podrían calcular, incluso saldos de jubilación, etc.

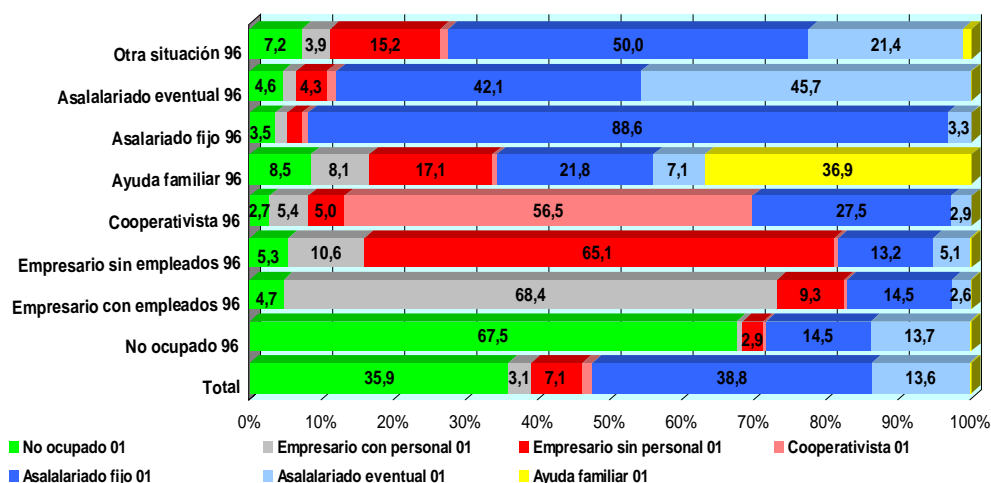
**Gráfico 3. Población ocupada e inactiva de 2001 según procedencia.
C.A. de Euskadi.%.**



Fuente: Eustat, EPV96-CPV01

En el Gráfico 4 se puede empezar a apreciar la potencialidad y también las limitaciones del análisis longitudinal aplicado a los censos. Por un lado pueden servir para identificar subgrupos muy específicos, que raramente pueden reconstruir

Gráfico 4. Distribución de la población que permanece entre 1996-2001 según la situación profesional en 1996 y 2001. C.A. de Euskadi.%



Fuente: Eustat, EPV96-CPV01

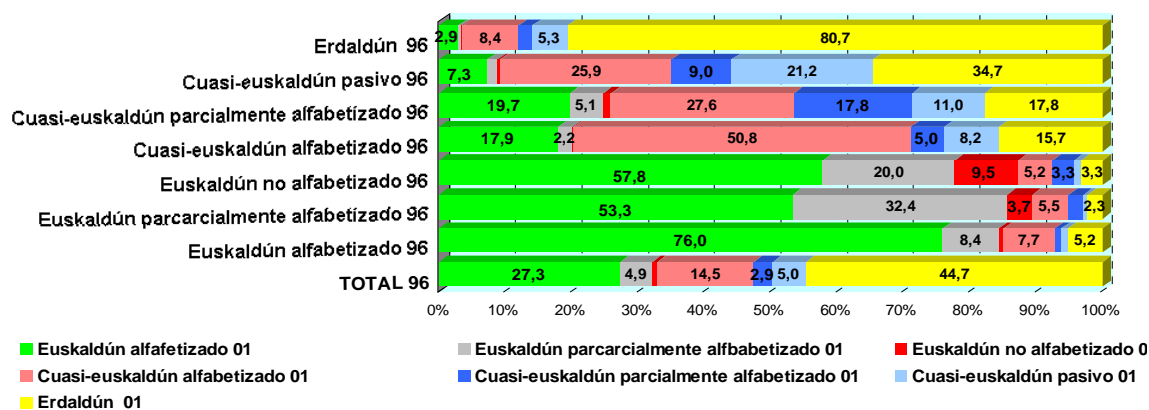
Destaca por su permanencia, que parece tener asociada al nombre, el grupo de asalariados fijos, que con respecto a 1996 se mantienen en un 88,6% de los casos, los empresarios con empleados en un 68,4% y los autónomos que repiten en un 65,1%. El resto de grupos presentan una alta movilidad: se mantienen en su situación un 56,5% de los cooperativistas, pero hasta un 27,5% pasan a ser asalariados fijos. Las personas que trabajan sin contrato en explotaciones o empresas familiares –ayudas familiares– repiten en un 36,9% de los casos, un 21,8 pasan a asalariados fijos y un 7,1% a eventuales; resulta subrayable el 17,1% que pasan a autónomos y un 8,1% a empresarios con trabajadores.

No obstante el dato más sorprendente lo constituye el 45,7% de asalariados eventuales en 1996 que seguían en esa situación en 2001. De este grupo sólo el 42,1% pasó a ser asalariado fijo y un 4,3% decidió montar su propio negocio (aunque sin empleados). Del total del colectivo de eventuales de 2001 llegan a representar el 48,4%.

Si se estudiase este colectivo teniendo en cuenta el sexo se vería como de las mujeres eventuales de 1996, el 50% permanecía en esta situación en 2001, frente a un 42,8% de los varones. El 46,8% de estos pasan a fijos por el 35,7% de las mujeres. Un 7,2% pasan a empresarios o autónomos frente a un 4,3% de las mujeres eventuales. Destacan estas últimas por pasar a situaciones de inactividad o de paro –un 8,5%– frente al 1,9% de los varones.

Estudio de la movilidad lingüística 1996-2001

Gráfico 5. Distribución de la población de 2 y más años que permanece entre 1996-2001 según el nivel global de euskera en 1996 y 2001. Edad a 31/12/2000. C.A. de Euskadi.%



Fuente: Eustat, EPV96-CPV01

La movilidad lingüística en las comunidades con lengua propia aparece como uno de los objetos de análisis más adecuados a la metodología longitudinal, ya que cabe esperar que los cambios en esa área temática tengan cadencias más largas que las que se pueden dar en la actividad, por seguir con un ejemplo ya expuesto.

Tres son las características que se incluyen en el censo lingüístico de la C. A. de Euskadi: nivel global de Euskera (conocimiento o competencia lingüística), lengua/s materna/s y lengua/s hablada/s en casa. Se genera o extrae la primera de 4 ítems (nivel de comprensión, habla, lectura y escritura) y las otras de dos preguntas directas.

En el Gráfico 5 se pueden apreciar con respecto al conocimiento el nivel de estabilidad de los distintos grupos y la dirección de los cambios. Con respecto al primer apartado se puede apreciar la gran estabilidad del grupo de erdaldunes –población que no conoce el euskera- con un 80,7% de permanencia, seguido del de euskaldunes alfabetizados –un 76%- . Los grupos frontera entre los dos primeros colectivos citados, sobre todo los de euskaldunes no alfabetizados, los parcialmente alfabetizados, así como los cuasi-euskaldunes alfabetizados poseen índices de permanencia situados cerca del 50%.

Tanto entre las personas con esos grados de conocimiento como en el resto de cuasi-euskaldunes, se concentran las mayores variaciones. A pesar de disponer de las tendencias generales, ofrecidas por las permanencias y los cambios, parece necesario indagar sobre la explicación de las variaciones. Dado que se conoce el fuerte impacto en el nivel de conocimiento de la lengua propia del sistema educativo, parece obligado recurrir al estudio de las ganancias y pérdidas del idioma en función de la edad.

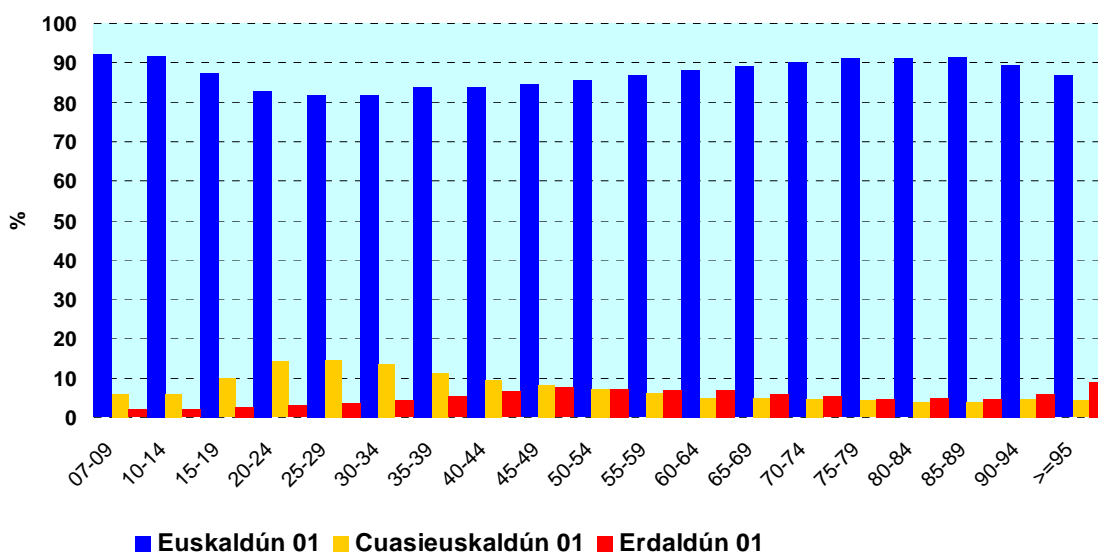
Este tipo de hipótesis requiere la subdivisión en grupos detallados de análisis, lo que de alguna manera puede complicar el trabajo, además de ampliarlo.

En el Gráfico 6 se presenta la situación de los euskaldunes en 1996 según su situación en 2001 con respecto al conocimiento del euskera (agregado en tres categorías). El primer dato a subrayar deriva de que ya se empiezan a detectar subgrupos con muy pocos efectivos: jóvenes euskaldunes en 1996, de 7 a 9 años en 2001, que pasan a erdaldunes o cuasieuskaldunes en 2001 (menos de 500 individuos) o personas en la misma situación de 90 y más años (poco más de 200). Dado que, como ya se comentó, todas las variables están afectadas por errores de consistencia, cabe pensar que algunas evoluciones con pocos efectivos pueden estar escondiendo este tipo de casos.

Por otro lado se aprecia el fuerte nivel de permanencia en general en todos los grupos de edad, que en todos los casos supera el 80%. En los grupos de 15 a 59 años se concentran las mayores pérdidas. Se podrían hacer dos grupos: de 15 a 34 años –en la mayoría de los casos el euskera se ha adquirido a través de la escuela- y el de 35 a 59, cuyas generaciones han tenido un menor acceso institucional a la lengua vernácula.

En el primer caso se aprecia una pérdida paulatina (parece que se pasa en mayor medida a cuasieuskaldún que a erdaldún), mientras que en el segundo el paso parece más indiscriminado, incluso tiende, con la edad, a aproximarse más a la pérdida total. No obstante en estos últimos grupos aumenta la permanencia, fruto, probablemente, de la adquisición de la lengua por vía familiar y de uso, más que a través de la escuela, con lo que ‘padecen’ en menor medida el efecto olvido.

Gráfico 6. Población de 2 y más años euskaldún en 1996 por nivel de euskera en 2001. C.A. de Euskadi. Edad a 31/12/2000. %



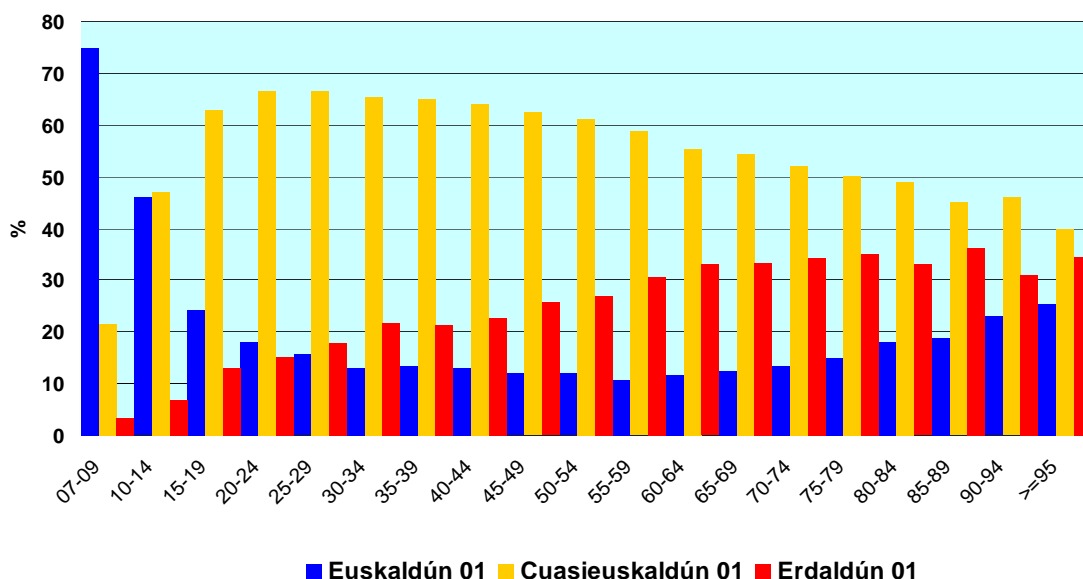
Fuente: Eustat, EPV96-CPV01

No habría que descartar, en los análisis de las pérdidas (ganancias en su caso), precisar otras posibles causas del cambio: cambios familiares, cambios de zona lingüística, etc.

La inestabilidad que se apreció en el colectivo de cuasieuskaldunes (personas que entienden bien o con dificultad el euskera y lo hablan con dificultad; en el caso de los llamados pasivos, no lo hablan aunque entiendan), se refleja en los fuertes cambios con una apreciable asociación a la edad.

De 7 a 9 años tres de cada cuatro cuasieuskaldunes pasan a considerarse euskaldunes, y se mantienen en la situación previa uno de cada cinco. De 10 a 14 años se reparten casi por igual los que mejoran como los que no varían. De 15 a 19 años todavía se aprecia un grupo de mejora importante (uno de cada cuatro pasan a euskaldunes), aunque ya comienza a apreciarse una tendencia creciente de pérdida completa (entre el 20 y el 35% del resto de grupos de edad).

Gráfico 7. Población de 2 y más años cuasieuskaldún en 1996 por nivel de euskera en 2001. C.A. de Euskadi. Edad a 31/12/2000. %

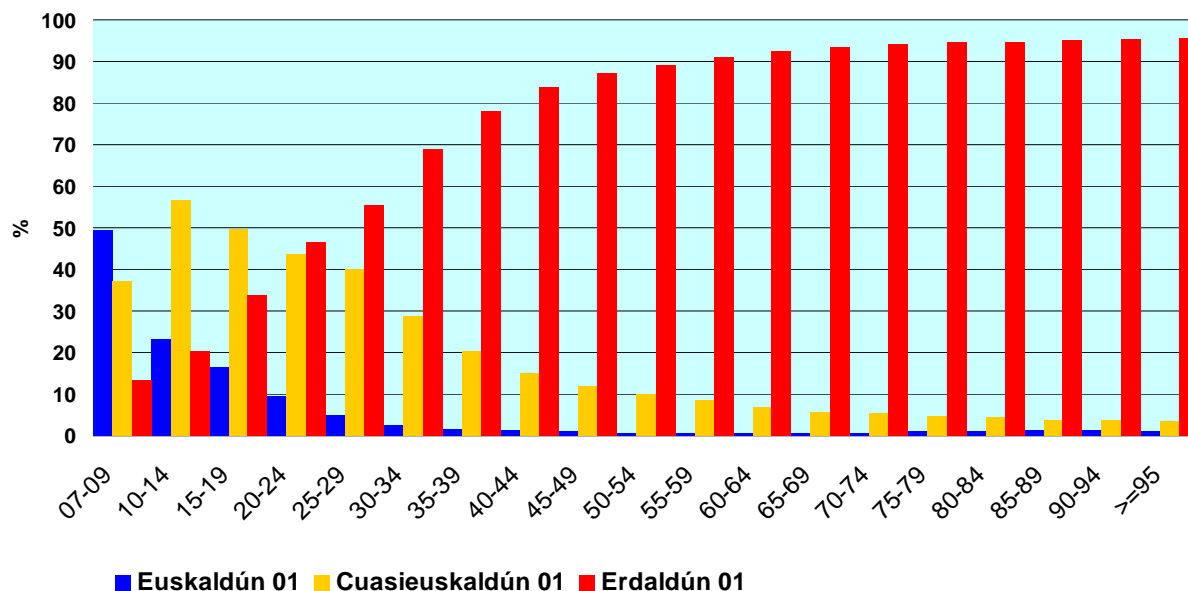


Fuente: Eustat, EPV96-CPV01

Resulta interesante la estabilización de los 30 a los 70 años de los cuasieuskaldunes que pasan a euskaldunes. De entre las posibles explicaciones (excluidas la del propio aprendizaje), cabrían plantearse dos: la falta de consistencia del grupo y el 're-aprendizaje' por cambios familiares, básicamente paternidad/maternidad y convivencia con nietos, que en muchos casos conlleva un aumento, cuando no una vuelta al uso del euskera.

Queda, para completar el análisis de los cambios de los grupos de población según su conocimiento del euskera, revisar la evolución en los cinco años intercensales de la población que desconocía el euskera en 1996, los llamados erdaldunes.

Gráfico 8. Población de 2 y más años erdaldún en 1996 por nivel de euskera en 2001.
C.A. de Euskadi. Edad a 31/12/2000. %



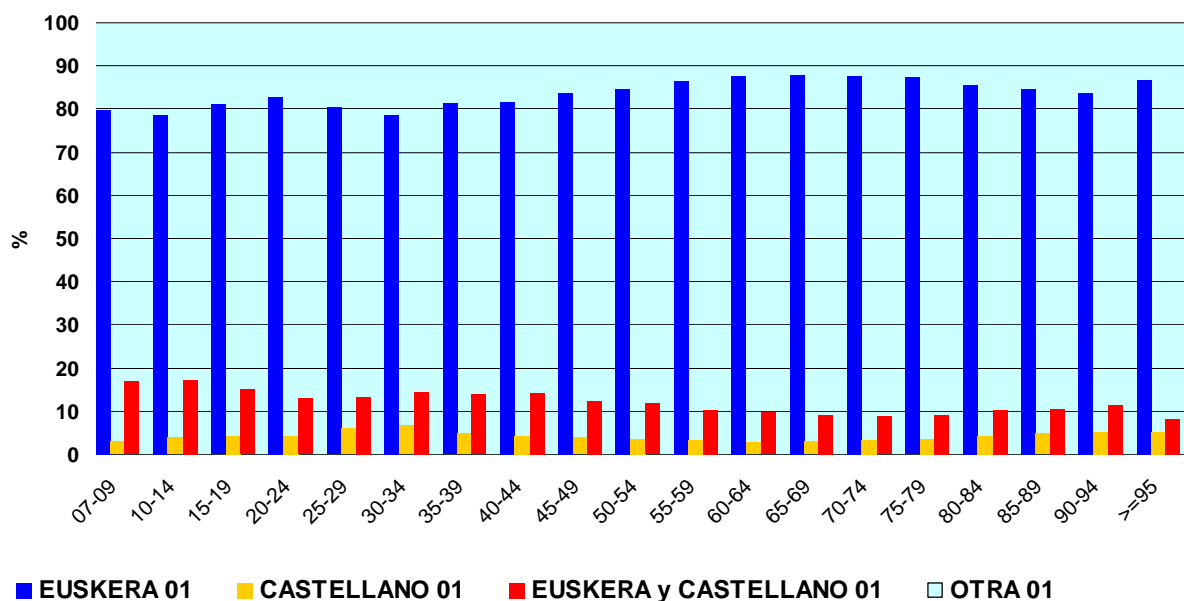
Fuente: Eustat, EPV96-CPV01

Quizá sea el grupo más homogéneo en cuanto al comportamiento teniendo en cuenta la edad, entre otras razones porque cabe pensar que tendrá una fuerte asociación con tener como lengua materna una distinta del euskera y los procesos, en este caso sólo de aprendizaje, estén determinados en su mayoría por la escuela o el euskaltegi (centro de enseñanza del euskera para adultos).

Hasta los 40 años, aparece visible el proceso de aprendizaje, aunque solo de 7 a 9 años se aprecia un cierto peso de los euskaldunes (hasta del 49%). Habría que plantearse, no obstante, la significatividad y calidad de la información tanto del 1996 como de 2001 para los menores de 10 años, ya que resulta problemático asignar niveles de conocimiento de cualquier lengua a niños de 1 a 4 años. Por otro lado, también puede ser cuestionable el paso en 5 años de erdaldun a euskaldun sin pasar por el nivel intermedio de cuasieuskaldun.

Las frecuencias mínimas que aparecen de euskaldunizados o semi-euskaldunizados de los 50 años también merecen un estudio aparte de cara a descartar errores de consistencia.

Gráfico 9. Población de 2 y más años que hablaba euskera en casa en 1996 por lengua hablada en 2001. C.A. de Euskadi. Edad a 31/12/2000. %



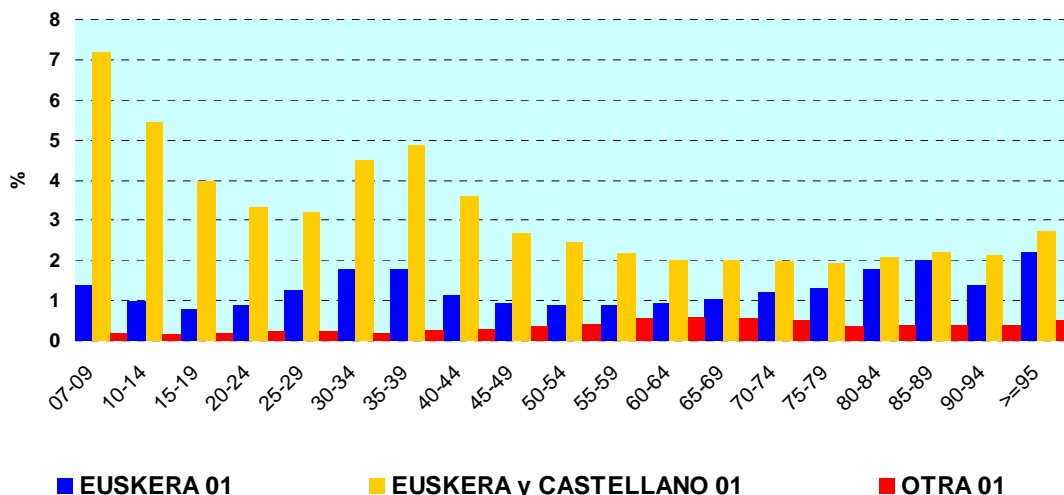
Fuente: Eustat, EPV96-CPV01

Para cerrar esta introducción al análisis de la movilidad del euskera se presenta la situación de las personas que dijeron en 1996 que en su casa sólo utilizaban el euskera. Aunque se produce en todos los grupos de edad un fuerte mantenimiento de los hábitos lingüísticos, sí que aparece un porcentaje sensible de personas que pasan a utilizar las dos lenguas, destacando el grupo de 7 a 19 años con porcentajes que van del 15 y el 17%.

Los casos en que se cambia de lengua, pasando a usar el castellano no superan nunca del 6% y se concentran más entre 25 y 44 años, período en que se suelen formar las familias y en que los matrimonios mixtos, desde el punto de vista lingüístico, pueden modificar los hábitos de las familias de partida.

Dada la alta permanencia en el mantenimiento del castellano como lengua de uso en casa en los hogares vascos (ronda el 95% de los casos que en 1996 ya usaban el castellano), se ha presentado en el Gráfico 10 sólo las variaciones. Aunque los efectivos de las transiciones son mínimos, se aprecian dos puntas de migración al euskera o al hábito de usar euskera y castellano con relación a la edad: con los menores (7 a 9 años) y en el caso de la población de 35 a 39 años. El matrimonio y la llegada de los hijos, así como su posible euskaldunización o refuerzo de la lengua materna en la escuela, parecen estar muy asociada al cambio de hábitos lingüísticos. No cabría otra explicación para el caso del cambio de lengua entre las generaciones más longevas, que la del contacto con los propios nietos.

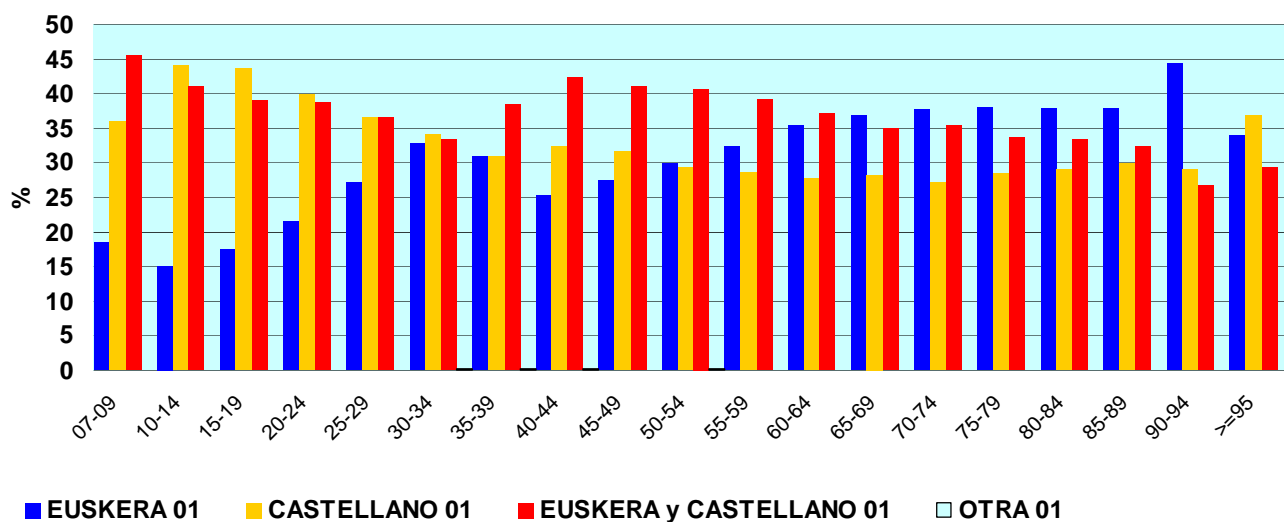
Gráfico 10. Población de 2 y más años que hablaba castellano en casa en 1996 por lengua hablada en 2001. C.A. de Euskadi. Edad a 31/12/2000. %



Fuente: Eustat, EPV96-CPV01

Como se ha visto en otros casos de colectivos sin una definición precisa y objetiva u objetivable, las posibilidades de variación entre dos períodos tienden a aumentar en intensidad y en posibles direcciones. Frente a la gran estabilidad de los que hablan sólo euskera en casa o sólo castellano, aparecen en el Gráfico 11 fuertes cambios por edad en las dos direcciones posibles, si excluimos la migración a otra lengua no oficial; apenas se produce en unos centenares de casos entre los que hablan euskera o euskera y castellano y sí aparece como un poco más frecuente entre los que en 1996 hablaban castellano, alcanzando a algo más de 4.000 personas. Porcentualmente, estos últimos, se sitúan por debajo del 0,6% de cualquier grupo de edad.

Gráfico 11. Población de 2 y más años que hablaba euskera y castellano en casa en 1996 por lengua hablada en 2001. C.A. de Euskadi. Edad a 31/12/2000. %



Fuente: Eustat, EPV96-CPV01

Si se excluyen los grupos de edad extremos, la población que se mantiene en el uso de ambas lenguas se sitúa entre el 30 y el 40%. Más atenuados vuelven a aparecer los dos picos de paso al euskera como lengua única y mantenimiento del bilingüismo entre los más jóvenes y entre las personas de 30 a 34 años. A partir de esa edad cae el monolingüismo en euskera, aunque se mantiene alto el uso de ambas y a partir de los 45 años vuelve a incrementarse el cambio al euskera.

Conclusiones y perspectivas de futuro

La posibilidad que surgió de la necesidad de mejorar los tratamientos de los Censos de 2001, que adolecieron de una limitada cobertura y/o falta de respuesta sensible, a través del anterior censo disponible (la Estadística de Población y Viviendas de 1996), ha potenciado las oportunidades de explotación y difusión de la información censal con una nueva perspectiva: el análisis longitudinal.

La aproximación realizada en Eustat a este tipo de análisis ha dejado al descubierto la necesidad del diseño de una metodología o tratamientos que incluya, al menos, los siguientes puntos:

- Integración de los tratamientos longitudinales en el diseño de las bases de datos censales. Diseño de herramientas específicas.
- Posibilidad de identificar la procedencia u origen de las unidades en los procesos de fusión o asociación de ficheros. En su caso recoger las referencias temporales. Introducir marcadores de nivel de aproximación en las fusiones, en orden a mejorar el análisis de la calidad de los flujos.
- Interés de tener identificadas bajas y altas en los archivos con relación al Movimiento Natural de la Población.
- Desarrollo y fijación de una terminología adaptada y clara, así como definición de las posibles vías de análisis.
- Establecer indicadores de calidad de la nueva información: cobertura de fusión (por tipos de población), consistencia, etc.
- Implementación de tratamientos estadísticos de validación, depuración e imputación de la nueva información (cambios o 'permanencias' indebidas) en paralelo a los tratamientos internos al último censo disponible. Introducir la 'trazabilidad' de los tratamientos en las unidades censales.
- Diseño de un plan de explicación y difusión de la nueva información.

Dado que el futuro de los censos demográficos parece que pasará principalmente por tratamientos de información administrativa (no se descartan microcensos en cuanto a recogida de información en subgrupos de población o temáticas específicas), la perspectiva longitudinal toma aún más sentido, no solo en orden a mantener la información de los individuos un mayor período de tiempo, sino con el fin de poder mejorar la calidad estadística de las propias fuentes. Como se sabe, casi todas las fuentes registrales adolecen de buenas herramientas y recursos para mantener constantemente actualizadas sus informaciones.

La propia posibilidad del acceso más frecuente a este tipo de fuentes podría mejorar y multiplicar, una vez mecanizados los distintos tratamientos necesarios, las unidades de análisis más interesantes, los flujos.

Por último hay que subrayar, con relación al último punto de la lista arriba adjuntado, que resulta muy necesario diseñar formatos de difusión accesibles a los usuarios de información censal. La posibilidad de fijar indicadores de cambio o de permanencia simples, que puedan trasladarse al ámbito espacial –no se puede olvidar que esta ha sido y es la dimensión fundamental de los censos- y grupal, puede ser una vía más en este sentido.