

Carles de Rosselló i Peralta¹ i Albert Fabà Prats²

Consorci per a la Normalització de la Llengua Catalana (CPNL)^{1,2}, Universitat de Barcelona (UB)¹

carlesrossello@gmail.com

Recepció: 06/03/2017, acceptació: 31/05/2017

Resum: En aquest article abordem, en primer lloc, els principals punts forts i dèbils de les dues modalitats amb què es recullen els usos lingüístics: a través del testimoni que reporten els informants (dades declarades) o bé mitjançant l'observació que fa un investigador (dades observades). En la segona part de l'article plantejarem si és possible desenvolupar una forma de comparar les dades declarades i les observades. Ho fem a partir de l'anàlisi de les dades de l'Enquesta d'Usos Lingüístics de la Població i de l'Ofercat relacionades amb els usos al comerç.

Mots clau: dades declarades, dades observades, usos lingüístics.

Datos declarados y observados, pros y contras

Resumen: En este artículo abordamos, en primer lugar, las principales fortalezas y debilidades de las dos modalidades que se usan para obtener datos sobre los usos lingüísticos: el testimonio que reportan los informantes (datos declarados) y la observación que hace un investigador (datos observados). En la segunda parte del artículo planteamos si es posible desarrollar una manera de comparar los datos declarados y los datos observados. Lo hacemos a partir del análisis de los datos de la Encuesta de Usos Lingüísticos de la Población y del Ofercat, relacionadas con los usos lingüísticos en el sector comercial.

Palabras clave: datos declarados, datos observados, usos lingüísticos.

Declared data and observed data, pros and cons

Abstract: In this article we address, first of all, the main strengths and weaknesses of the two modalities with which linguistic uses are collected: through the testimony reported by the informants (declared data) or by the observation made by a researcher (observed data). In the second part of the article we consider whether it is possible to develop a way of comparing declared and observed data. We do this based on the analysis of the data from the Survey on Language Use of the Population and the Ofercat, related to the linguistic uses in the stores.

Keywords: declared data, observed data, linguistic uses.

Els usos lingüístics es recullen a partir de dues modalitats: o bé a través del testimoni que reporten els informants (dades declarades) o bé mitjançant l'observació que fa un investigador (dades observades). Abordem en aquesta primera part de l'article les dades declarades, pròpies principalment de recerques macrosociolingüístiques, com ara enquestes i censos. L'avantatge fonamental d'aquest tipus de dades és la quantitat d'informació que s'obté de grans poblacions —a vegades de tota la ciutadania—, cosa que seria impossible d'atènyer per mètodes qualitatius. Per si sol aquest és un element amb prou de pes per suplir les nombroses mancances que presenten les recerques que opten per aquest tipus de dades.

I quins són aquests punts dèbils? La primera qüestió que cal destacar és que sovint a les enquestes se'ls fa dir allò que no pregunten. Miquel Pueyo (2007: 45) recorda l'episodi en què l'Institut d'Estudis Catalans, així com força articulistes, van alertar del «retrocés inquietant de l'ús del català» arran de la publicació d'uns estudis —previsiblement l'Enquesta de la Regió Metropolitana de Barcelona de l'any 2000— quan en aquesta enquesta no hi havia ni una sola pregunta sobre l'ús social.¹ En el cas dels censos, ocorre una cosa similar: es pregunta sobre el coneixement i no sobre l'ús, de manera que el percentatge de població que declara saber parlar un idioma no vol dir que realment el parli, sinó simplement que manifesta la capacitat de parlar-lo (cf. Junyent 1999). Un segon aspecte bàsic és que allò que diuen que fan els enquestats no ha de ser necessàriament el que realment passa (aquest punt el tractarem amb més detall quan ens referim a les dades observades). Un tercer element és que quan es demana a una persona que valori les seves capacitats lingüístiques no tothom té el mateix grau d'exigència a l'hora d'autoavaluar-se. El fet que no hi hagi una prova estàndard per a tots els enquestats resta credibilitat a les dades recollides. Vinculat a aquesta objecció, en el cas específic del cens es pregunta si la persona sap parlar (o si entén, o sap escriure o sap llegir) en una llengua, sense explicitar què implica aquest saber (en el cas de parlar, significa dir una simple salutació, mantenir una conversa senzilla o fer una conferència davant d'un auditori?). Bourhis i Sachdev (1984) apunten que el context sociolingüístic i la percepció subjectiva de la vitalitat etnolingüística també repercuteixen en l'autoavaluació de l'ús lingüístic. Finalment, les opcions de resposta també determinen l'anàlisi i les conclusions posteriors. En els censos, en què la dinàmica de contesta és l'autoavaluació, la resposta és binària (sé parlar aquesta llengua / no la sé parlar) per tal de minimitzar tan com sigui possible el temps i la dificultat de la resposta. Això comporta que les dades siguin de traç gruixut: no poden matisar el ventall competencial que presenten els bilingües, des dels incipients fins als ambilingües (Beardsmore 1986). Per contra, altres qüestionaris també de caire macrosociolingüístic, com ara l'Enquesta d'Usos Lingüístics de la Població (EULP), tenen un rang de respostes en les seves preguntes molt més exhaustiu que permet afinar les anàlisis amb molta precisió. Això sí, amb un equip d'enquestadors professionals instruïts i uns protocols ben establerts.

Les prevencions metodològiques, doncs, a l'hora de recollir i tractar dades sobre els usos lingüístics estan plenament fonamentades. En aquest sentit, les dades del cens 2001 suposen un bon exemple de recursos públics esmerçats balderament precisament per la manca de cura en algunes de les qüestions esmentades. De fet, la Xarxa CRUSCAT va acabar recomanant que no s'utilitzessin les dades per fer comparacions intercensals perquè «no hi ha cap certesa que aquesta comparació sigui metodològicament lícita» (Vila 2005).

¹ L'Enquesta preguntava per l'ús familiar, la identitat lingüística i el coneixement de la llengua.

A pesar de les flaques que acabem d'apuntar, Modest Reixach, un dels pioners de la demolingüística a Catalunya, defensava en una entrevista (Boix-Fuster 2012: 241) la validesa de les dades declarades a partir dels censos:

«I això per dues raons, com a mínim. Si no es demostra el contrari, hem de suposar, d'entrada, que les respostes dels informants o censats corresponen a la realitat tal com ells la perceben. Això no dispensa d'analitzar la possible existència de factors causants de desviacions o biaixos. Una altra raó a favor de la fiabilitat dels censos es podria formular així: les desviacions produïdes per circumstàncies purament aleatòries tendeixen a compensar-se quan la població censada és molt nombrosa.»

La segona modalitat de recollida de dades són les observades, és a dir, aquelles que l'investigador obté mitjançant estudis de caràcter etnogràfic, basades en l'observació directa —sigui participant o no— i sovint complementades amb tècniques procedents de l'antropologia i la sociologia, com ara les entrevistes o l'enregistrament dels informants. La qualitat de les dades obtingudes és un dels grans arguments que avalen l'ús d'etnografies. Milroy (1987: 78) i Milroy i Gordon (2003: 68) n'apunten d'altres, com ara que la participació en les comunitats de pràctiques possibilita explicar les normes comunicatives i socials que regeixen els informants, i també per què la manera de parlar d'una persona ocupa un lloc determinat dins de l'estructura social.

Unes línies enrere, quan parlàvem de les dades declarades, hem apuntat que sovint el que l'informant diu que passa, en realitat succeeix d'una altra manera. L'observació del comportament humà, doncs, permet contrastar la percepció de la persona. Vol dir això que hi ha una voluntat de mentir? No necessàriament; o com a mínim, no en tots els casos. Gumperz (1982: 61) ja adverteix que molts individus no són conscients de la seva tria lingüística perquè amb l'ús d'alternances de codi el que busquen, bàsicament, és aconseguir un determinat efecte comunicatiu i no presten atenció als codis emprats. De manera similar, Rosselló i Ginebra (2014), en un estudi amb alumnes de 10-11 anys, van comprovar com un exercici aparentment tan senzill com és identificar el codi de relació amb una altra persona, generava una distància considerable entre els usos observats i els usos declarats. Per exemple, entre els joves de L1 castellana, la diferència entre les tries lingüístiques que els investigadors van observar i les dades declarades que els informants van respondre a través dels qüestionaris arribava al 42%. Les interaccions que duïen a més confusió eren aquelles en què es declarava que la conversa es resolvia utilitzant les dues llengües però l'observació ho desmentia i ens aproximava a l'ús del català o del castellà com a codi de relació exclusiu o molt majoritari.

Els mètodes, el temps i l'energia que cal emprar per obtenir dades observades, doncs, fa que resultin molt difícil d'aplicar a grans col·lectius (*cf.* amb els mesuraments de l'ús de les llengües al carrer que es duen a terme al País Basc i Navarra i que, en l'edició de 2016, va suposar l'observació de més de mig milió de persones).² No és la manca de representativitat, però, l'única problemàtica relacionada amb les dades observades. A pesar de l'enorme dispersió de camps i interessos que comprèn la sociolingüística, els investigadors basen els seus avenços, hipòtesis i teories en un punt en comú: les dades. Amb freqüència aquestes dades són producte de l'observació però això no les converteix automàticament en la «realitat». Com apunta Johnstone (2000:22), les observacions sovint estan impregnades de teoria i en conseqüència l'investigador espera que passi allò que se suposa que ha de passar o allò que vol que succeeixi. Tot plegat provoca que científics socials desconfiïn de les observacions sistemàtiques com una eina vàlida per aproximar-se a la realitat i per falsar hipòtesis i teories.

²Torres (2011) apunta que un dels reptes de la demolingüística és augmentar el nombre de dades observades.

Cameron et al. (1992: 6-10) plantegen dues vies per afrontar aquesta situació: el *relativisme* i el *realisme*. Els teòrics socials adscrits a la primera visió consideren que no hi ha cap més realitat social que l'experiència subjectiva de la persona, que tot es produeix al nivell micro de l'organització social i que, per contra, els conceptes macro com classe o gènere no tenen poder explicatori. Els relativistes descriuen actuacions i interaccions particulars que fa la persona en la seva quotidianitat, però eviten fer extrapolacions més enllà de l'àmbit individual encara que hi hagi casos amb característiques semblants. Òbviament, optar per aquesta opció ens aparta de qualsevol perspectiva científica, atès que la ciència intenta cercar les regularitats entre els fets particulars, malgrat les dificultats d'aquest objectiu i la prudència i la perspectiva crítica amb què s'hagin d'afrontar aquestes regularitats. Cameron et al. opten pel realisme perquè malgrat que es poden produir interpretacions divergents davant d'un fenomen, en un cert nivell les persones comparteixen la mateixa experiència i això permet generalitzar interpretacions més enllà estrictament de l'individu. Això és especialment comprovable en sociolingüística amb corrents com ara la sociologia de la llengua o el variacionisme, que aspiren a donar compte del comportament lingüístic de la població a partir de mostres més o menys grans.

Un altre aspecte que cal tenir present per a l'èxit d'una recerca etnogràfica és que aquesta depèn en bona mesura de les decisions i les actituds que assumeixi l'etnògraf durant la investigació, especialment durant les primeres etapes. Així, qüestions com ara el procés que cal seguir per introduir-se a la comunitat, el rol que l'investigador ha de desenvolupar o quin tipus d'informació s'està disposat a compartir amb els informants tenen una gran transcendència. L'habilitat de l'analista per resoldre aquests interrogants i altres circumstàncies que van sorgint en el transcurs de l'observació i que no es poden preveure a priori són clau, ja que en el pitjor dels casos poden vedar l'accés a la comunitat que es vol estudiar de manera definitiva.

El darrer aspecte que volem destacar és la procedència de l'investigador. Si aquest és d'origen forà, tindrà una visió sobre el col·lectiu que li permetrà qüestionar aspectes que passarien inadvertits per un autòcton, precisament perquè aquest ja comparteix moltes de les característiques sociolingüístiques de la comunitat. Per contra, el forà pot tenir mancances lingüístiques que li dificultin l'accés als informants. I encara que no fos així, un bon domini de la llengua i un projecte d'investigació sòlid no converteixen l'observador en un membre més de la comunitat ni asseguren l'èxit de l'empresa. Briggs (1986), per exemple, relata com va haver d'adquirir les normes pragmàtiques que regien les interaccions diàries dels habitants de la població que estudiava després que fracassés diversos cops en els seus intents de recollir dades. Les dificultats de l'observació participant no s'atenuen si l'investigador forma part de la mateixa comunitat, simplement n'apareixen unes altres. En el seu estudi sobre el col·lectiu nipó a Catalunya, Fukuda (2009) esmenta les dificultats per accedir a les xarxes personals dels seus compatriotes precisament pel fet de compartir nacionalitat. Les normes de conducta intragrupal, sobretot en xarxes molt denses, aconsellen mantenir en privat informacions que en mans d'un altre japonès podrien afectar la vida laboral i familiar de l'informant.

1. COMPARAR DADES DECLARADES I OBSERVADES?

Fet i fet, doncs, sembla clar que seria interessant poder comprovar (i no només amb un estudi de cas de tipus etnogràfic) si les dades obtingudes mitjançant els dos mètodes difereixen (i si ho fan molt o poc). A Catalunya, un dels estudis de dades declarades més rigorós des del punt de vista sociolingüístic és la sèrie de l'Enquesta d'Usos Lingüístics de la Població (EULP). Seria lògic pensar en la seva utilització com a font principal, pel que fa a aquest tipus de dades.

Ara bé, quan ens hi acarem ens adonem que no totes les preguntes que s'hi fan són aptes per a una comprovació d'aquesta mena. Fixem-nos, per exemple, en dues de les que s'hi consideren més importants: la llengua inicial i la llengua d'identificació (Rosselló 2015). La pregunta de la llengua inicial es refereix a la llengua que la persona enquestada diu haver parlat primer a casa seva, mentre que la llengua d'identificació respon a la pregunta «quina és la seva llengua?».

És ben clar que cap de les dues preguntes no ens pot resultar útil per resoldre el problema de saber si hi ha divergències, o no, entre els valors declarats i observats d'una determinada variable sociolingüística. D'una banda, perquè quan preguntem la llengua inicial (a persones de 15 anys o més) difícilment la podrem contrastar amb un estudi observacional de la llengua que van parlar els informants, de petits, atès que aquest fet lingüístic es va produir fa molts anys i ja no s'esdevindrà un altre cop. De l'altra, perquè la llengua d'identificació no és, pròpiament, un ús lingüístic, sinó una opinió lingüística, subjectiva per tant, que s'acaba un cop s'ha expressat aquesta opinió mitjançant una declaració. En aquest cas el fet declarat i observat són pròpiament el mateix.

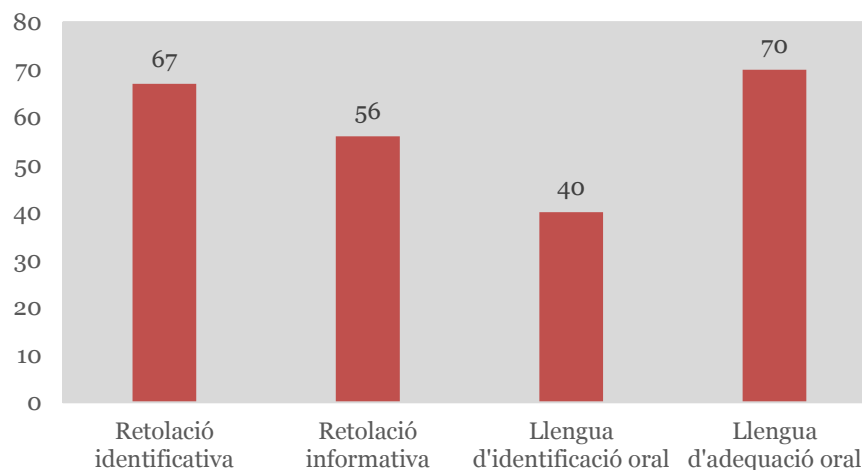
Sembla clar, doncs, que hauríem de cercar alguna pregunta o preguntes de l'EULP en què es puguin fer, alhora, observacions relativament properes en el temps, que ens puguin servir d'ajut a l'objectiu que plantejem. Una d'aquestes preguntes seria la que fa referència a l'ús de les llengües als petits comerços. El motiu d'aquesta tria és que disposem d'estudis observacionals sobre l'ús oral als comerços, amb una mostra prou àmplia i representativa que potser podríem acarar amb els resultats de l'EULP.

Aquests estudis es refereixen a l'oferta lingüística en una determinada ciutat, fets amb el mètode anomenat *Ofercat*. Així, s'observen quatre elements d'una mostra representativa de comerços de la ciutat, dos relacionats amb la retolació (el rètol principal i la llengua de l'aparador) i altres dos amb l'ús oral d'atenció al públic (la llengua en què s'efectua la primera atenció a l'usuari, d'una banda, i la llengua d'adequació, en cas que l'usuari/observador parli en català).

Durant l'any 2012 es van fer observacions mitjançant l'*Ofercat* a una mostra de més de 4.000 comerços de la ciutat de Barcelona (Fabà, Torrijos, Solé 2014). D'altra banda, tenim una submostra de l'EULP 2013 –de pràcticament 1.000 casos–, també de la ciutat de Barcelona, en què es va preguntar la llengua que els usuaris utilitzaven en el comerç.

El resultat d'aquestes observacions d'*Ofercat* s'exposen en el gràfic següent, mitjançant un índex ponderat que fa referència exclusivament al català:

FIGURA 1. Índex *Ofercat* als comerços de Barcelona. Any 2012



Font: *Ofercat*

Tot i això, ens trobem amb un inconvenient difícil de solucionar. No estem parlant de la mateixa situació.

A l'EULP ens situem en el punt de vista de l'usuari i li preguntem en quina llengua parla quan va al petit comerç. No sabem, per tant, qüestions tan cabdals per al desenllaç de la interacció lingüística com ara quina és la persona que ha iniciat la interacció, ni quina llengua ha utilitzat aquesta persona per iniciar-la. Diversos estudis (Fabà 2003 i 2005) indiquen que aquests dos elements (conjuntament amb el grau de coneixement lingüístic que tenen els dos interlocutors) determinen en bona part quina és la llengua que predomina en la interlocució. Això no obstant, en el cas de l'EULP només en sabem el resultat final, és a dir, la llengua que ha utilitzat l'usuari en la interlocució, i no queda constància de les circumstàncies en què s'ha produït la conversa.

Per contra, a l'Ofercat, el punt de vista que s'analitza és el del comerciant. En un primer moment, recreant una situació (certament artificial) mitjançant la qual es facilita que sigui el comerciant el que iniciï el primer torn de la interlocució (els observadors estan ensinistrats, mitjançant diverses tècniques, per tal que la interlocució s'esdevingui d'aquesta manera). En un segon moment ja entra en acció l'usuari/observador: això sí, parlant en català i observant, per tant, en quina llengua és correspost.

No sembla, per tant, que en puguem treure l'entrellat. Podríem recrear un estudi que ens facilités l'objectiu que busquem? Pensem que sí, que és possible, també aprofitant la sèrie d'observacions d'Ofercat que es produeixen regularment. Una possibilitat seria centrar-ho en la nova observació que s'haurà d'esdevenir a la ciutat de Barcelona.

En aquest cas es podria fer una enquesta prèvia als comerços que s'observaran, posteriorment, prou lluny en el temps per tal que els comerciants no puguin associar ambdós esdeveniments, però prou propera perquè la comparació ens resulti útil. Una enquesta amb diverses preguntes en què, entre d'altres, es preguntaria en quina llengua atendria els clients si aquests se li adrecessin en català (la segona variable oral que s'observa a l'Ofercat).

I amb quina llengua s'hauria de portar a terme aquest qüestionari? Sembla que el castellà. Sí, perquè si es preguntés en català estariem reproduint la situació que, teòricament, volem preguntar posteriorment, la qual cosa podria produir un biaix que no ens interessa (contestar *allò que ha fet*, en aquell moment, i no *allò que pensa que fa o farà*, en aquesta situació).

Si en algun moment es pot esdevenir l'esmentat estudi potser podrem respondre, ni que sigui parcialment, la pregunta que ens plantejàvem a l'inici d'aquest apartat: podem desenvolupar una forma de comparar les dades declarades i les observades (en aquest cas en l'àmbit del comerç)?

2. CONCLUSIÓ

En aquest article hem resseguit les dues modalitats de recollida de dades –la declarada i l'observacional– a partir dels quals els estudis sociolingüístics obtenen informació sobre els usos i la competència lingüístics. A través de l'anàlisi dels avantatges i inconvenients principals, hem observat com les tècniques associades a les dades declarades permeten obtenir-ne un volum considerable, però des d'un punt de vista més qualitatiu poden presentar biaixos significatius, com ara una interpretació que s'extralimiti d'allò que es pregunta o que els enquestats afirmen que tenen un capteniment que realment no es produeix. Les dades observades, per la seva part, permeten superar alguns d'aquests esculls però les anàlisis s'han de circumscriure en l'àmbit micro atès que, per norma general, el volum de dades que s'obté a partir d'aquesta modalitat és petit. En la segona part de l'article hem exemplificat les dificultats que hi ha per comparar dades declarades i observades a partir dels usos lingüístics en comerços.

3. REFERÈNCIES BIBLIOGRÀFIQUES

- BAETENS BEARDSMORE, H. (1986). *Bilingualism: Basic Principles*. Clevedon: Multilingual Matters, 2a edició. Traducció catalana: (1989) *Principis bàsics del bilingüisme*, Barcelona: La Magrana.
- BOIX-FUSTER, EMILI (2012). «Entrevista a Modest Reixach (1931-2011)». *Treballs de Sociolingüística Catalana*, 29, 233-247.
- BOURHIS, R. Y.; SACHDEV, I. (1984). «Vitality perceptions and language attitudes: some Canadian data». *Journal of Language and Social Psychology*, 3, núm. 2, 97-126
- BRIGGS, CHARLES L. (1986). *Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge: Cambridge University Press.
- CAMERON, D; FRAZER, E.; HARVEY, P.; RAMPTON, M.B.H.; RICHARDSON, K. (eds.). (1992). *Researching Language: Issues of Power and Method*. Londres: Routledge.
- FABÀ, ALBERT (2003). «L'ús interpersonal del català i altres variables sociolingüístiques. Assaig d'un model interpretatiu. El cas de santa Coloma de Gramenet». *Revista de Llengua i Dret*, 40, 185-229.
- (2005). «Usos lingüístics i grups identitaris. Preferències i opinions lingüístiques». A: *Estadística sobre els usos lingüístics a Catalunya 2003. Llengua i societat a Catalunya en els inicis del segle XXI*. Barcelona: Generalitat de Catalunya, Departament de la Presidència, 171-210.
- FABÀ, ALBERT; TORRIJOS, ANNA; SOLÉ, JOAN (2014). «Usos lingüístics als establiments comercials de Barcelona (Ofertat 2012)». *Revista de Llengua i Ús*, 54.
- FUKUDA, MAKIKO (2009). *Els japonesos a Catalunya i la llengua catalana. Comunitat, llengües i ideologies*. Tesi doctoral. Universitat de Barcelona.
- GUMPERZ, JOHN J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- JOHNSTONE, BARBARA (2000). *Qualitative Methods in Sociolinguistics*. Nova York: Oxford University Press.
- JUNYENT, CARMÉ (1999). «El català, una llengua en perill d'extinció?». *Revista d'Igualada*, 1, 27-38.
- MILROY, LESLEY (1987). *Observing and Analyzing Natural Language. A Critical Account of Sociolinguistic Method*. 2a ed. Oxford: Blackwell.
- MILROY, LESLEY; GORDON, MATTHEW (2003). *Sociolinguistics. Method and Interpretation*. Oxford: Blackwell.
- PUEYO, MIQUEL (2007). *El fantasma de la mort del català*. Barcelona: Proa.
- ROSSELLÓ, CARLES DE; GINEBRA, DAVID (2014). «Tries lingüístiques: vuit anys després. L'evolució dels usos lingüístics des de P3 fins a 6è de primària». *Treballs de Sociolingüística Catalana*, 24, 267-280.
- ROSSELLÓ, CARLES DE (2015). «Llengua inicial, llengua d'identificació i llengua habitual». Anàlisi de l'Enquesta d'usos lingüístics de la població 2013. Resum dels factors clau. Barcelona: Generalitat de Catalunya, Departament de Cultura, 15-17.
- SOZIOLINGUISTIKA KLUSTERRA (2017). VII medicion del uso de las lenguas en la calle. Euskal Herria, 2016.
- TORRES, JOAQUIM (2011). «La demolingüística en els territoris de llengua catalana». *Treballs de Sociolingüística Catalana*, 21, 183-192.
- VILA, F. XAVIER (coord). (2005). *Les dades lingüístiques del cens de l'any 2001: alguns dubtes i limitacions*. Barcelona: Institut d'Estudis Catalans.