

PROBLEMAS RELATIVOS ÓS PROCESOS DE RECOLLIDA E TRANSCRICIÓN DE MOSTRAS ORAIS*¹

XOSÉ SOTO ANDIÓ

Resumo

Este artigo aborda unha serie de problemas de carácter empírico e metodolóxico, relativos ós procesos de recollida de mostras orais e a súa posterior transcripción. A partir de aquí trátanse cuestións que atinxen á selección e localización de informantes, á entrevista (o lugar e o momento da entrevista, formas de realizala...), ás mostras (selección, organización e rexistro das mostras, transcripción do material, tipos de transcripción...), etc. Todo isto encamiñado a contrastar as nosas experiencias neste eido coas doutros investigadores, para así discutir as dificultades metodolóxicas desta actividade e contribuír á procura de mellores solucións.

Preámbulo

Son cada vez máis frecuentes os traballos que centran as súas análises ou na oralidade ou na lingua escrita xunto á oral e á conversacional (Biber *et al.* 1999) ou no contraste de ambas (Jahandarie 1999). Este tipo de estudos require un labor previo baseado en *investigacións empíricas* consistentes na recompilación de mostras orais para a posterior elaboración de cörpera lingüísticos, investigacións que pasan en liñas xerais por dúas fases. A primeira fase ocúpase da *recollida* de mostras, que máis tarde ha dar lugar ó corpus de oralidade. A segunda é a fase de *transcripción* desas mostras. Ambos procesos comprenden unha serie de pasos a seguir e presentan algúns problemas cuxa solución ha de estar acorde cos fins que persiga a investigación que esteamos a desenvolver (De Kock 2001). Estas fases e subfases son as que aborda o artigo que arestora expoñemos, partindo sempre das nosas experiencias quer como transcritores durante

¹ Este traballo é unha ampliación da comunicación presentada no *36 Colloque international de la Societas Linguistica Europaea: "Linguistique et corpus"*, organizado pola *ENS de Lettres et Sciences Humaines* da Universidade de Lyon entre o 4 e o 7 de setembro de 2003.

case dous anos no *Consello da Cultura Galega* de Santiago de Compostela, quer na realización de diversos traballos de campo –entrevistas, gravacións- e transcripcións de textos, que foron realizados para as nosas teses de licenciatura e doutoramento, orientadas a estudos de carácter dialectal, sociolingüístico e gramatical.

A partir de aquí imos expoñer mediante un procedemento indutivo algunhas propostas e problemas derivados desta actividade consistente na recolleita e transcripción de textos orais en Galicia (vid. un antecedente en Sánchez Rodríguez 1986).

Centraremos a nosa atención en temas coma os seguintes: selección e localización de informantes, entrevista, material utilizado, selección e transcripción de mostras, principais convencións de transcripción, rexistro de mostras, etc. A posterior manipulación dos textos mediante programas informáticos queda á marxe desta exposición.

1. Recollida de mostras

Resultan cada vez máis frecuentes en todas as linguas os traballos de recolleita de mostras orais. En linguas minorizadas coma o galego este labor é aínda máis valioso se cabe polas seguintes razóns:

- i) Seren linguas nas que o número de falantes pode estancarse ou retroceder.
- ii) En caso de retroceso dispoñermos da maior cantidade posible de testemuños orais, de gravacións realizadas a individuos que posúan unha fala enxebre e pura, tal e como se usa na vida cotiá (McEnery & Wilson 2001: 104), ou que representan profesións, oficios, etc. que esmorecen ou que xa desapareceron (canteiros, carreiros, leiteiros, zapateiros, afiadores, ferreiros...).
- iii) Crear importantes cörpera de textos orais que sirvan de base a traballos de investigación sobre o galego así coma ó meirande desenvolvemento dunha lingüística propia que ata hai poucas décadas se limitaba a unhas cantas manifestacións esporádicas.

Dende estes presupostos, foron moitos os estudos que en Galicia, nos últimos vinte e cinco anos, se teñen centrado en traballos de campo consistentes na recollida de mostras orais de diferentes xeracións, tanto no ámbito rural, preferentemente, coma no urbano. A partir de aquí, podemos inferir que a recolleita de material que ha servir de base a calquera corpus de textos orais esixe unha *dependencia empírica* que consta de varias fases, subfases e procedementos, que deseguido especificamos.

1.1. Selección e localización de informantes

Para a constitución de calquera corpus de oralidade é preciso realizar unha determinada *operación selectiva* pola que acoutamos un universo empírico na realidade que se nos presenta con características e limitacións de seu. Nos traballos de campo, a selección de informantes que posteriormente vai integrar o corpus pódese levar a cabo comezando polo *censo* ou padrón de habitantes da zona previamente escollida. Comprobamos pola nosa experiencia que resulta de moita utilidade a este respecto un bo *coñecemento xeográfico* da mesma. Agora ben, o problema que xorde seleccionando os informantes a través dun censo municipal é que os resultados non sexan sempre os devecidos, dado que é posible que o investigador chegue a un determinado lugar e a persoa ou persoas ás que ía entrevistar non se atopen nese momento no seu domicilio ou simplemente non desexen conversar. Para sortearmos esta dificultade podemos acordar co informante, por exemplo a través de chamada telefónica previa, una hora de *cita*. A pexa é que mediante este procedemento no se xera confianza dabondo no entrevistado e moitos non se amosan dispostos a falar con descoñecidos, a ceder parte do seu tempo de modo consciente e por tanto a encontrar con facilidade un oco nas súas ocupacións.

Todo isto acaba por obrigar ó investigador-entrevistador a optar polo que os sociólogos denominan *redes sociais*, adoptando o procedemento en cadea ou de bóla de neve, é dicir, situados nun lugar visitar calquera posible informante e que sexa este o que nos encarreire cara a outros falantes (veciños, amigos, parentes...). En ocasións o entrevistado que nos conduce a outros informantes (amigos ou achegados) actúa de *mediador* na presentación e *participa libremente* na entrevista. O feito de irmos de parte doutra persoa amiga ou coñecida facilita bastante o traballo, facendo que o entrevistado incremente o seu grao de confianza en nós por lle resultarmos menos estraños. E aínda moito máis o facilita a condición de *naturais da zona* ou o ter ascendentes e parentes na mesma que o entrevistado identifique. Vexamos algún exemplo:

X- [en voz alta] Manolo.

M- ¿Que pasou?

X- A ver, hai aquí un mozo que seica quere falar contigo.

M- ¿E logo que quere?

X- Quere falar cun canteiro pa un traballo que está facendo ou non sei que...

ENTREV: Ola, boas tardes, é que quería falar con vostede do traballo dos canteiros...

X- A ver, pasade, pasade pa dentro.

ENTREV: Boas tardes, ¿como está?

L- Ola, ¿que é o que querías?

ENTREV: Era para facerlle unha entrevista e que me falase sobre a vida de antes. Eu son o neto do señor José, o cocineiro de Covas, ¿non se acorda del?

L- Ai, xa m' acordo, ¿pero xa morreu, non...?

1.2. Entrevista

O termo *entrevista* defínese, en xeral, da seguinte maneira:

1. “Encontro de dúas ou máis persoas nun lugar determinado, para tratar ou resolver un asunto. 2. Conversa en que unha persoa lle fai preguntas a outra... 3. Texto en que aparecen as respostas que unha persoa dá a un xornalista nunha conversa deste tipo” (DRAG 1997: 473).

“Vista, concurrencia y conferencia de dos o más personas en lugar determinado, para tratar o resolver un negocio” (DRAE 2001: 935).

Nós concibimos a entrevista como un medio utilizado para a obtención de datos, algo que acontece tanto en disciplinas lingüísticas coma non lingüísticas (socioloxía, periodismo, antropoloxía...), perseguindo en cada unha obxectivos distintos. A partir de aquí pódese orixinar un proceso de *interdisciplinariedade* polo que o lingüista-entrevistador colabora con outros profesionais; dito doutra maneira, o material obtido polo lingüista nalgúns traballos de campo pode ser aproveitado por estudosos doutras áreas non lingüísticas como etnógrafos, antropólogos, historiadores, sociólogos, musicólogos que recollen cantigas e folclore tradicional, etc.

A entrevista ha procurar a consecución do maior número posible de datos, isto significa que en principio canto máis *ampla* e *diversa* sexa, mostras e conclusións máis precisas se poderán extraer. A duración de cada entrevista-gravación tende a oscilar entre un mínimo de quince minutos e un máximo de cento vinte².

1.2.1. O lugar e o momento da entrevista

A entrevista pódese realizar no *domicilio* do entrevistado, nalgún espazo apartado da *rúa* e en calquera outro punto previamente acordado ou improvisado. No caso dos nenos e adolescentes tamén se pode acudir a un *centro escolar*: colexio, instituto, universidade, residencia de estudantes. E tratándose de persoas maiores é posible desprazármonos ata os *centros da terceira idade* (Monxardín 1998).

Os lugares nos que se vaian obter mostras gravadas han ser preferentemente *espazos cotiáns* que favorezan o discurso espontáneo e coloquial, tamén tranquilos (o domicilio, unha aula baleira, un despacho, un

² En máis dunha ocasión temos realizado gravacións de dúas ou tres horas a un único informante, pero son casos excepcionais.

banco afastado do barullo nun parque...), *sen ruidos* de fondo (unha televisión acesa, balbordo de voces nun café, ruído de coches na rúa, nenos xogando no patio dun colexio, etc.) que poidan máis tarde dificultar a transcripción obrigando a prescindir de palabras, sons e matices que, segundo os casos, poidan ser máis ou menos relevantes. Neste sentido son pouco recomendables, como acabamos de indicar, os cafés con xente, un restaurante, lugares públicos como parques ou patios de colexio, o interior dun vehículo circulando, as rúas urbanas, un taller, etc.

Se o lugar ten que ser o axeitado tamén o *momento* ha de ser o oportuno, isto é, que os informantes no se atopen excesivamente ocupados nin estean en horas de máximo descanso. Neste sentido hai que distinguir entre dous períodos de tempo como son a semana e a fin de semana. A semana e os días laborables ou lectivos son a ocasión propicia para realizar entrevistas nun centro escolar e nalgunhas empresas. A fin de semana constitúe o período de tempo adecuado para dirixírmonos a aqueles domicilios onde os posibles entrevistados traballen de luns a venres ata moi tarde. Dentro da fin de semana, en Galicia acostuma a ser máis eficaz para gravacións o sábado có domingo (sempre falando en xeral), dado que neste día algúns familiares e amigos adoitan reunirse, comer xuntos, ir ó café, ver fútbol, pasear... Durante o día tampouco son bos momentos en Galicia as horas de forte traballo, a hora da comida e da sesta, a cea e, segundo acabamos de sinalar, as ocasións de reunión familiar, xa que os informantes non se han prestar a falar e se acceden será con présa, desexando rematar canto antes para proseguiren co que estaban a facer ou para volveren cos seus amigos. Vexamos algún exemplo rexistrado en Aciveiro-Forcarei-Pontevedra:

ENTREV: Boas tardes, quería falar con José, o carpinteiro.

X- Estache aí cortando leña na eira, entras por esa cancela è x'ò ves.

ENTREV: Ola, viña falar un momento con vostede sobre o oficio de carpinteiro para un traballo que estou facendo.

J- Entra, entra pa dentro. Chita, pasa d' aí [diríxese a un can].

ENTREV: Fálame do traballo de carpinteiro.

J- Era un traballo de facer carros, sobre todo...

[De novo ponse a traballar picando leña]

ENTREV: ¿Que máis se facía á parte de carros?

J- Portas, cancelas, portales tamén se facían, moitas cousas...

[contesta con présa e amósase inqueda, quere seguir traballando]

ENTREV: ¿O dos carros como era?

J- Os carros teñen moitas partes, per' eso è longo de contar...

X- [avisan en voz alta] José, mira que vén aí o trator...

[van cargar leña nun tractor, a conversa iniciada detense de todo, o informante non atende]

1.2.2. Formas de realizar unha entrevista: o entrevistado e o entrevistador

O destinatario da entrevista pode ser unha *persoa individual* á que se lle pregunta sobre un determinado *tema*, favorecéndose o dinamismo discursivo que se concreta en *intervencións prolongadas* (conversar sobre os anos de xuventude, os xogos da infancia, a relixión, os costumes do lugar...), ou se lle formulan *preguntas curtas* de resposta tamén curta, coas que entrevistador e entrevistado se converten ambos os dous en *interlocutores activos alternantes*. Nas entrevistas a nenos obtéñense frecuentemente respostas breves. Velaquí algúns exemplos recollidos en Forcarei-Pontevedra, que corresponden a unha informante de sesenta e cinco anos sen estudos e a un neno de dez anos:

ENTREV.: ¿Cómo eran os casamentos antes?

A- antes eran moi bonitos, hòme. Os casamentos de antes, **buèno**, hòme, buèno. Eran na casa, faciámolos na casa. Eu fíxeno na do Xilo, porqu'a miña casa non era moi ghrande, però despois fomos aí onda un que tèn un comercio aghor'alí en Pontevedra, qu'è primo aquí de Bouzas..., fómose axudarll' a polas mesas na casa, apartaron as camas pa un lado è puxéron alí as mesas. Fíxéron'us **bizcochos** na casa, cor ghuèvos da casa. Se ti n'os tiñas xuntábans'os de todos e iban p'alí pa face-los ghuèvos, è había unha muller que chamábanlle Bolandina, qu'èr'à cociñeira, facía moi ben de comere. Levabas us xamós e chourizos è facías un bo cocido, è facías un bo arroz, unha boa paella de arroz, è non había marisco nin había esa cousa, pèr'había a fartar...

ENTREV.: ¿Como pasas unha fin de semana cos amigos?

M- Xogamos casi tod'ò tempo. Xogar a fútbol, botar carreiras, ó escondite e nada máis.

ENTREV.: ¿Non andas en bici?

M- Eu pouco.

ENTREV.: ¿Alternas os estudos con algún traballo no campo?

M- No.

ENTREV.: ¿Que falas máis galego o castelán?

M- Os dous **mezclados**.

ENTREV.: ¿Os nenos de aquí da túa idade en qué lingua falan?

M- Gallego e castellano, e a mèdias os dous.

Noutros casos as entrevistas realízanse a *parellas*, a *grupos* de varias persoas ou a *grupos de discusión*. Aquí os informantes son interpelados en común polo entrevistador ou propónselles un tema de conversa co fin de que se expresen libre e espontaneamente, desdebuxándose os papeis discursivos entrevistador-entrevistado e as quendas de pregunta-resposta, para así procurar a *fala máis nativa, coloquial e natural*. O entrevistador pode actuar como simple observador unha vez proposto o tema de charla

e mesmo pode abandonar temporalmente o espazo en que transcorre a entrevista deixando sós os informantes (Beswick 1998: 61). En calquera destas situacións existe a posibilidade de que os interlocutores se moderen a si mesmos, introduzan variacións sobre o tema, completen a mensaxe, a maticen e a alteren mutuamente, esquecéndose con máis facilidade de que están a ser gravados. Velaquí dous exemplos en transcripción parcialmente fonolóxica, o primeiro pertence a una parella de sesenta anos de Forcarei- Pontevedra e o segundo reproduce a charla entre dúas adolescentes de 17 anos naturais de Xinzo-Ourense e Baio-A Coruña.

ENTREV.: ¿Como foi o traballo de carpinteiro?
B- Pois arrèghlo de carros, carros do país.
Clav- È facelor novos.
B- Facelor novos, facer arados, facer alghún muèble...
Clav- Facer mesitas de **nóche**, camas...
B- Si, buèno, do que se facía naquèla època.
ENTREV.: Ós carros botábanlles graxa para que andasen.
B- A graxa botábaselle no eixo por causa de que andase máis liso.
Clav- A graxa è esa da casa.
B- Buèno graxa valía graxa...
Clav- Graxa sin sal.
ENTREV.: ¿Os ladrais cómo se facían?
B- Facíanse dunhas tablas de centímetro è cuarto d'espesor.
Clav- Dunha **tabla** sòla.
B- Pèro daban cincuenta centímetros de alto.
Clav- Èran boas.

- *A ver, falade entre vós do que queirades.*

R. Non fixèmos nada a verdá.
A. Xa pèr'estes días que fai tan bo tempo calquèra lle vai dicir a clase.
R. Ti viches María que guapa 'stá, un cambiazo impresionante. Estaba toda pintadiña hoxe, co seu pelo recollido.
A. Si tiña unha trenza feita que lla fixo Cristina, díxom'a min.
R. ¿Cristina d'Òrènsè?
A. Si, fíxolla èla porque pregunteill'eu ¿quen che fix'esta trenza?, è dice no Cristina è tal è díxenll'eu ah vale, è viña pintadiña cos labios pintados.
R. È quitounos unha foto.
A. A min o que máis me amolou foi cuando me dice eso de sacarl'a foto ò jicho este que lle gusta, sabes, è bue, que lle dab'ásí un pouco deso, è díxenlle bue pois si quères sácocha eu è tal, è si que lla saquèi.
R. E ti viches en que posturiña se puso.
A. I-ò da margarita, que colleu unha margarita que se puxo alí è tal, si está mui guai.
R. È a Tamara co outro, tía. Tamara estaba con este tío que è amigo do que lle gusta a María.
A. Ah si, pero eu non-o **conozco**.

R. Si que o conòces, este chico de melenas, que tiña **melenitas** è despois cortou o pelo è aínda me dixèches tí, aí pois gústame como lle qued'ò pelo còrto.

A. Pois eu no me dou conta, de verdá eh, te lo juro.

R. Pois cando o vexa dígocho, de **acuèrdo**. È fumei, aí canto fumei hoxe.

A. Eu pensei que m'iba chegar unha **cajetilla**, pero tuven que comprar outra.

R. Pois eu compreí lunes unha, martes unha e **miércoles** unha.

R. E nada, qué máis fixèmos hoxe, a ver.

A. No sei.

A. Aí a clase de prácticas de teoría aburridísimo.

R. Aburridísimo, compréndote.

A. Jolín eu cando saliu Daviz, o da noiva.

R. È tí viches a Sònia, Sònia non facía máis que rirse, tío.

A. Quèn è Sònia.

R. Sònia è a d'Òrense, a qu'anda con Daviz.

Os entrevistados adoitan pertencer á mesma xeración, porén ás veces son de idades e xeracións opostas (por exemplo unha parella de 50 anos cos seus fillos de 20), o que permite confrontar variedades de lingua e alternancias de código. Tamén poden proceder de puntos xeográficos dispares para así contrastar usos e variacións de carácter diatópico (vid. exemplo anterior de falantes de Ourense e A Coruña).

O entrevistado pode responder de diferentes maneiras a unha entrevista concertada ou non concertada: amosar nerviosismo, querer rematar axiña, aceptala de bo grao. Temos comprobado que en xeral as persoas de xeracións vellas préstanse a falar con bastante entusiasmo cando se lles pregunta polos seus anos de xuventude, a súa vida pasada, etc. Lembrar eses anos a moitos satisfailles enormemente, deste xeito a charla convértese para eles nunha terapia psicolóxica de desafogo. Algúns mesmo nos chegaron a agradecer o feito de seren entrevistados por acordarnos deles, escoitalos e xa que logo permitirlles contar anécdotas da súa vida.

Nos momentos previos a calquera gravación, enquisa e/ou toma de notas é conveniente que o entrevistador estableza co entrevistado un proceso de *interacción* que derreta a cortina de xeo e o escepticismo iniciais, xerando *confianza*, seguridade, familiaridade e tranquilidade no informante, volvendo a *relación afable* e distendida. Para isto, o investigador-entrevistador dispón de varios recursos como son o de explicar o contido da entrevista, a súa finalidade, o que estuda ou aquilo no que traballa, de onde procede, etc. Deste modo o informante confíase, ten menos présa por rematar, explícase con meirande naturalidade e en consecuencia a charla transcorrerá consonte o previsto, acadando os obxectivos preestablecidos (vid. tamén Vaux & Cooper 1999: 10ss.). Isto resulta aínda máis importante cando o investigador-entrevistador non é nativo da zona ou non coñece o entrevistado. En caso contrario, o tempo que se

ocupe nesta fase pódese reducir. Vexamos un exemplo en que se reflicte unha aproximación entre entrevistador e informante recollido en Coto-bade:

ENTREV.: Boas tardes, se tes un momento gustaríame charlar contigo sobre o traballo da construción.

P- ¿Que querías saber?

ENTREV.: Que me explicases un pouco en que consiste, cando empeza-ches... Porque ti antes de andar por aquí fuches emigrante, non.

P- bueno, si, estúven varios anos fóra.

ENTREV.: Dixéronme que estiveras en Francia traballando. Daquela ga-naríanse alí bastantes cartos porque ademais o cambio favorecía.

P- Si, traballei alí en varias compañías e facendo carreteras. Todo o ano. Viñas dous ou tres meses d'inverno e volvías marchar.

ENTREV.: Si, eu tamén tiven familiares que emigraron, aquí en Galicia nos anos sesenta emigrou moita xente a centroeuropa, a situación eco-nómica era tan mala que case non había outra saída se querías prosperar. Por sorte as cousas melloraron bastante, anque segue habendo moita emigración de xente nova e con estudos...

P- Aquí non facías nada.

ENTREV.: Ben, entón podemos logo gravar algo e falar dos traballos que fas aquí na construción.

P- A ver logo, pero séntate, séntate aí. [Acéndese a gravadora e centrá-monos no tema].

Unha vez lograda esa inmediatez co entrevistado, antes de iniciar a con-versa débese proceder a tomar nota de diversos datos de arquivo, a saber: nome e apelidos do informante, idade ou idade aproximada, profesión e nivel de estudos, lugar de nacemento e residencia, lugar da entrevista (es-pazo: domicilio, rúa...; aldea ou vila, parroquia, concello, provincia), data (día, mes, ano), nivel socioeconómico. Outras informacións poderanse completar máis tarde ó non ser necesaria a presenza do entrevistado.

O entrevistador pode realizar a entrevista dunha maneira libre e flexible, improvisando as preguntas, enlazando con calquera materia que se vai tratando ó longo da conversa, modelando esta en función do entre-vistado, insistindo naqueles asuntos que parecen ser do seu agrado ou que coñece mellor, mudando de tema se se decata de que o falante non mos-tra entusiasmo dabondo á hora de responder a certas cuestións, introdu-cindo algunha anécdota na entrevista, etc. É o que se denomina *técnica amplificatoria e relacionante*. Fronte a esta técnica de enquisa libre e de temas ligados, o investigador pódese servir de procedementos rigorosos coma o da *entrevista dirixida ou controlada*, é dicir, cun cuestionario pre-viamente elaborado que sirva de guía para que nada quede ó chou e para que os contidos que interesan non pasen desapercibidos, agromando sempre aquelas palabras ou expresións que ó investigador lle convén re-

xistrar. Neste caso, as preguntas serán as máis axeitadas para as nosas metas e estarán formuladas da mellor maneira posible. Por exemplo, os módulos conversacionais de moitos cuestionarios dialectais adoitan ser de carácter etnográfico, con diálogos cuxos contidos abordan temáticas non especializadas, máis familiares e comúns ó entrevistado: traballos do campo, oficios, relixión, crenzas, sentimento e vivencia da morte, a infancia e os xogos, o matrimonio, a maternidade, a familia, as celebracións ou festas da zona e as do almanaque, etc. Noutros casos, tratándose de informantes de cultura alta (un profesor, un avogado...), as temáticas poden ser máis especializadas, con charlas menos coloquiais e de orientación máis formal. En estudos de sociolingüística e sobre contacto de linguas o obxecto do cuestionario pode ser tamén o propio código no referente a usos, actitudes e coñecementos sobre o mesmo: en que lingua se expresa normalmente o informante, cal foi a súa lingua materna, cando e en que circunstancias emprega unha ou outra, con que interlocutores, que variante adopta segundo os fins, cal lle gusta máis, que nivel de coñecementos ten dunha ou doutra, etc.

Durante a conversa hase de procurar a *boa relación* e a *fluidez comunicativa* co entrevistado, mirarlle para a cara cun sorriso, ser amable, evitar longos silencios, participar asentindo coa cabeza, con xestos, mediante comentarios de reforzo, etc.

O entrevistador debe de manifestar *versatilidade nas formas*, o que significa que cando se presente diante do informante ten que saber coidar e adaptar as mesmas, por exemplo con falantes do medio rural convén non utilizar demasiados cultismos nin expresións rebuscadas que poidan incrementar o nivel de distanciamento entre interlocutores. Nunha ocasión un falante do rural pensou que viñamos de parte da Xunta exercendo labores de control e vixilancia ó vernos portando un maletín e tomando notas. Transmitir este tipo de impresións acaba por causar nos informantes unha necesidade ben de autodefensa e rexeitamento ben de aco modo formal ó entrevistador desposuíndo así a fala dos seus trazos máis característicos e vernáculos.

1.2.3. A mostra

A mostra ha ser o suficientemente *representativa* do obxecto de investigación que se está a levar a cabo, se ben a medida da representatividade é moi difícil de determinar (Caravedo 1999: 54ss.).

1.2.3.1. Rexistro de mostras

Para rexistrar as mostras pódese tomar notas, cubrir un cuestionario e máis comunmente *gravar* utilizando unha gravadora portátil -cuxas características técnicas poden ser diversas-, que é o procedemento ordina-

rio máis rápido e cómodo, ou os diferentes métodos a un tempo. Para traballos de campo, en xeral utilízanse pequenos aparellos de fácil traslado así coma de bo rexistro e reprodución de son (ALGa 1990: 9-24; Soto Andión 1995: 8; Vaux & Cooper 1999: 69).

A gravación pódese realizar *in situ*, presentándonos diante do informante para obter del directamente os datos, ou pódese efectuar de forma non dirixida a través dun *medio*, é dicir, gravar dun programa de radio ou televisión.

No tocante á *calidade* da gravación, canto máis alta sexa máis partido lle ha poder sacar o lingüista para as súas investigacións ou para didáctica e perfeccionamento de linguas: individuos que as escoitan con vistas a mellorar a súa pronuncia e entoación, estudantes estranxeiros, no caso de Galicia nativos que falan habitualmente español e desexan perfeccionar o galego, locutores de medios de comunicación, actores e dobradores.

A gravadora adoita estar *á vista* do entrevistado, pero tamén pode permanecer *oculta* se o que desexamos é captar a lingua da maneira máis franca e instintiva, sen ningún tipo de condicionamento, cos seus tics e expresións máis directas e irreflexivas.

1.2.3.2. Organización da mostra

Os informantes son agrupados tendo en conta variables cuxos datos obtidos no momento da gravación han ser indicados ó comezo de cada transcrición; son datos como a *idade*, o *nivel cultural*, a *procedencia xeográfica*, etc. Outros parámetros máis ou menos relevantes segundo o perfil da investigación (vid. algunhas aplicacións dos corpóra en McEnery & Wilson 2001: 103ss.) son o *rexistro* (culto e elaborado fronte a popular e espontáneo), o *sexo* (usos, actitudes e comportamentos do home e da muller en relación con distintos fenómenos considerados), o *nivel socioeconómico* e a *profesión* (en traballos de sociolingüística ou naqueles que buscan léxico específico ou subcódigos dun determinado grupo como son en Galicia o *arxinas*, o *barallete*, o *daordés*, o *latín dos chafontas*).

En canto á idade, as mostras tenden a incluír falantes preferentemente dos tres sectores básicos agrupados en xeracións, isto é, xeración *vella*, *intermedia* e *nova*. No primeiro grupo ou xeración estarían aqueles informantes que superan os 65 ou 70 anos *grasso modo*. A xeración intermedia estaría composta por individuos de entre 35-50 anos e o grupo dos novos formarían os menores de 35 anos. En función dos obxectivos da investigación estes grupos poden ser desdoblados, delimitándose unha xeración *moi nova*, configurada por nenos que comezan a falar e adolescentes de 12, 14 e 16 anos, fronte a unha xeración *moi vella* que engloba informantes de máis de 85 anos. Este proceder serve de base a estudos lingüísticos de carácter contrastivo por idades e permite dar conta da variación lingüística de acordo con este parámetro. A mostra pode tamén

ser reducida a un único grupo ou xeración, sexa a nenos ou adolescentes para estudar fenómenos de aprendizaxe e adquisición de linguas, contacto de linguas, así como de léxico e de estruturas gramaticais, sexa a persoas vellas ou moi vellas para estudos de carácter etnolingüístico, dialectal, de léxico patrimonial, interferencias lingüísticas, etc., entre outros.

En función do nivel cultural e académico, os entrevistados son clasificados en informantes de cultura alta, media ou baixa, algo que ten a súa principal aplicación en investigacións de corte sociolingüístico.

Dende o punto de vista do lugar de nacemento ou residencia, as mostras poden incluír informantes de contorno urbano, semiurbano -pequenas vilas- e rural. Cando o lugar de nacemento coincida co de residencia bastará con anotar o primeiro. Se son distintos apuntaremos ambos para dar conta das mudanzas ou das resistencias ós cambios que se operan nos falantes. Do mesmo modo, é importante ter en conta, sobre todo en Galicia, se o individuo entrevistado é unha persoa que estivo moitos anos na *emigración*, se se trata dun emigrante retornado ou dun emigrante que está de vacacións, dado que estes factores puideron ter condicionado os seus hábitos de vida e a súa lingua.

2. *Transcripción de mostras orais*

A transcripción é un labor lento e paciente, pero necesario para a posterior manipulación dos textos (Vaux & Cooper 1999: 26) por parte do investigador-entrevistador ou doutros investigadores.

2.1. *Selección do material gravado para transcribir*

A selección do material recollido obedece a dous criterios fundamentais. Un criterio extralingüístico polo que son discriminadas aquelas gravacións en *bo estado*, máis intelixibles e con mellor son, doutras deficientes ou menos intelixibles, para despois, se se quere, proceder a separalas polo *tipo de texto*: contos, narracións de feitos, vivencias e anécdotas, cantigas populares...(Vaux & Cooper 1999: 183), ou polo seu *interese* e amenidade dende o punto de vista temático ou de contido. É un criterio lingüístico polo que son seleccionadas aquelas gravacións máis *representativas* do noso obxecto de estudo por posuíren un gran número de trazos caracterizadores (fonéticos, sociolingüísticos, léxicos, morfosintácticos...).

2.2. *Transcripción*

Transcribir é para o dicionario da RAG (2000: 1154) “copiar un texto co mesmo sistema de escritura ou con outro distinto”. Para o da RAE (2001: 2210) é “representar elementos fonéticos, fonolóxicos, mor-

fológicos y léxicos de una lengua o dialecto a través de un sistema de escritura”. En Dubois (1994: 606) enténdese por transcribir o feito de “hacer corresponder término por término las unidades discretas de la lengua hablada y las unidades gráficas”.

2.2.1. A partir de aquí, pódese elaborar unha clasificación das transcricións en tres grandes clases:

a) *Fonéticas* ou *fonolóxicas*: tamén denominadas transcricións estreitas. Servede preferentemente do alfabeto fonético internacional (IPA) e poden incluír ou non marcas prosódicas: entoación, pausas, modalidade oracional (Taylor 1996: 20; Knowles 1996: 87). A diferenza das ortográficas moitos segmentos non se puntúan (Taylor 1996: 20). Están orientadas a investigacións de índole fonético e fonolóxico, ensinanza de segundas linguas ou patoloxías da linguaxe.

b) *Fonolóxicas (parcialmente)*: tamén denominadas transcricións amplas. Nelas só se marcan os trazos fonético-fonolóxicos máis relevantes (Soto Andión 1995; Fernández & Hermida 1996). Son moi adecuadas para investigacións de lingüística diatópica, variedades de lingua e interferencias lingüísticas. Ó seren de máis doada lectura tamén poden ser utilizadas para distintos traballos gramaticais ou doutra índole.

c) *Ortográficas*: son aquelas que seguen as prescricións e convencións de escritura de cada lingua, podendo respectar asemade a presenza de vulgarismos e dialectalismos (Monxardín 1998). Interesa aquí sobre todo o contido do texto e as súas características morfolóxicas, sintácticas, de léxico ou pragmáticas. Están pensadas para a súa posterior utilización tanto por lingüistas estudosos da pragmática (Kasher 1998), da orde e a combinatoria de palabras, da gramática, da sociolingüística... (De Kock 2001), coma por outros investigadores alleos a este campo, chámense historiadores, etnógrafos e antropólogos, sociólogos.

2.2.2. A transcrición ten que ser *literal* e o máis *fiel* posible á gravación orixinal, non debendo de resumirse nin de modificarse. O que ó transcritor lle pode parecer irrelevante ás veces ten un importante significado (certos tics reiterativos e certos silencios poden ser utilizados en estudos de pragmática). Outra cousa será que a transcrición persiga única e exclusivamente obxectivos moi concretos e sexa puída de todo aquilo que non resulte de interese para esa meta trazada. Calquera cambio ou interpretación é responsabilidade do transcritor e ha de ser indicado mediante algún tipo de nota, convención, etc., agás casos obvios como a puntuación do texto.

O *tempo* empregado en transcribir está en función do tipo de transcrición³. Obviamente unha transcrición fonética cuadriplica o tempo

³ Non entro a analizar aquí instrumentos tecnolóxicos en permanente cambio, como son certos programas informáticos de transcrición automática ou diferentes aparellos

dunha ortográfica. Tamén inflúe o número de convencións empregado, certas dificultades como ruídos de audición, as habilidades do transcriptor en canto a rapidez tecleando, etc. En liñas xerais unha transcripción ortográfica de sesenta minutos supón seis horas de traballo. A isto convén engadir unha revisión final, é dicir, escoitar de novo toda a gravación ó mesmo tempo que se le e corrixe a transcripción.

As *intervencións do entrevistador* poden non ser incluídas cando, amais de facer máis lenta a lectura e dificultar o seguimento dalgunhas pasaxes, non son necesarias nin para a comprensión do texto nin para as análises posteriores do mesmo. En caso contrario, as preguntas formuladas transcribíense habitualmente ben en letra cursiva ben precedidas de guión ou das abreviaturas ENTREV., E., seguidas de guión, dous puntos, espazo en branco, etc.

As *intervencións do entrevistado* poden ir en estilo de letra normal cun guión ó comezo ou coa letra inicial do seu nome ben en cor automática ben en letra grosa. Se son varios os entrevistados hanse identificar cada un por unha letra ou abreviatura. Se fala un descoñecido adóitase identificar coa letra X. (cfr. tamén Edwards & Lampert 1993: 50). Exemplos:

¿A que hora se levanta?
- Normalmente ás nove.

ENTREV.: ¿A que hora se levanta?
P- Ás dez.

E.- ¿A que hora se levanta?
P.- Ás dez.

ENTREV ¿A que hora se levanta?
P Ás dez
X- Non sempre, eh. [fala un dos presentes].

2.2.3. Nas transcripcións manéxase de ordinario unha serie de *convencións tipográficas* así coma de signos fonéticos, prosódicos e ortográficos. En transcripcións fonéticas e fonolóxicas seguimos o IPA (International Phonetic Alphabet), menos frecuentemente o RFE (alfabeto da Revista de Filología Española). Nas ortográficas as normas ortográficas de cada lingua. Ademais, é habitual valerse de certas convencións como as seguintes:

manexados co pé que permiten os movementos das cintas sen necesidade de retirar as mans do teclado da computadora.

i) Entre corchetes sitúase algo que non aparece no texto, pero que o completa ou clarifica. Por exemplo:

[risos], [move a cabeza], [xesticula], [entra unha persoa], [soa o timbre ou o teléfono]... A aclaración pode ser tan precisa como se queira: [xesticula inflando a boca], [xesticula e move as mans], [xesticula e abre moito os ollos reflectindo asombro]...

ii) En situacións de contacto de linguas ás veces indícanse as principais interferencias dunha lingua noutra marcando palabras, estruturas alleas, préstamos, etc. en letra grosa sobre todo.

iii) Os dialectalismos e vulgarismos pódense subliñar, deste modo imos obter información da mudanza de rexistro (neste caso máis descoidado e espontáneo) ademais de datos léxicos, gramaticais e/ou sociolingüísticos...

iv) O signo # pode indicar que varias persoas falan simultaneamente.

v) As parénteses con puntos (...) indican corte ou supresión de texto.

vi) Os signos <> poden indicar reconstrución de texto: imp<osi>ble.

vii) As marcas xxx sinalan unha palabra da gravación que non se entende, e xxx xxx xxx varias palabras ou frases que non se entenden.

viii) ____ Esta liña a comezo de frase indica unha pausa do falante, que pode ser curta __, media_____ ou longa_____. A pausa pode non ser un silencio (*hesitation pause*) senón un son (*filled pause*) e hase indicar ben mediante algunha palabra gramatical (isto...) ben de forma onomatopéica: uhhh, mmm, eeh... (Bibe *et alii* 1999: 1053).

ix) As frechas ascendentes ou descendentes ↑↓ adoitan indicar inflexións de entoación, que se marcarán cando consideremos que son significativas para o noso traballo.

x) Dous interrogantes despois dunha palabra sinalan que é dubidosa ??

xi) O apóstrofo serve para indicar a elisión dunha vogal, son ou sílaba.

xii) O guión utilízase, ademais do seu uso convencional, para indicar a unión de dous sons ([j] i-entonces; i-os nenos) ou para sinalar que un *n* velar se transforma en alveolar por mor do son seguinte (non-o sabe; ben-a vin).

xiii) Unha liña en diagonal / indica o final da cara dunha fita e dúas liñas en diagonal // o remate da fita.

xiv) Observacións e comentarios ó comezo da transcripción de cada informante. Por exemplo:

INFORMANTE: ESPERANZA BARREIRO

IDADE: 60

INSTRUCCIÓN: primaria

DATA DE GRAVACIÓN: 11-10-97

LUGAR: COVAS, na súa casa.

PARROQUIA: FORCAREI

CONCELLO: FORCAREI

PROVINCIA: PONTEVEDRA

REXISTRO: coloquial

TEMA: traballos do campo

OBSERVACIÓNS: Utiliza gheada e rotacismo. Sempre falou galego. Prodúcense algunhas interrupcións do fillo. Ó principio non se prestaba a falar, sentía algo de vergoña.

2.2.4. Rexistro do material transcrito

Unha vez concluído o proceso de transcripción é conveniente gardar as gravacións nalgún tipo de *arquivo*. En Galicia podemos destacar, entre outros, o arquivo do *Atlas Lingüístico Galego* do ILG ou o *Arquivo Sonoro de Galicia* do Consello da Cultura Galega, que dispoñen dunha importante cantidade de gravacións realizadas por toda a xeografía galega. Tratándose de fitas, en breve convén mudar o soporte dixitalizándoas en previsión do seu posterior deterioro. Tanto as gravacións en soporte material coma as transcripcións poden ser arquivadas seguindo o esquema dunha *ficha técnica* elaborada de acordo cos parámetros da investigación que se estea a desenvolver, por exemplo en traballos de lingüística diatópica arquivamos por bloques e áreas lingüísticas, en sociolingüística o criterio pode ser a idade, a variedade de lingua, etc., noutros traballos primará o tema tratado, determinados aspectos de contido... Normalmente adoitan ser comúns datos coma os seguintes:

- a) Gravación: data, hora, autor, lugar en que se efectúa (espacio, parroquia, municipio, provincia), tipo (entrevista in situ, telefónica, dun medio de comunicación como radio ou televisión), temática, variedade e rexistro, duración.
- b) Informante: nome do entrevistado ou entrevistados, idade, sexo, nivel de estudos, nivel socioeconómico, profesión, lugar de nacemento e residencia.
- c) En situacións de contacto de linguas dentro dunha comunidade: lingua materna do informante, lingua en que se expresa actualmente, grao de destreza, predilección...

Os datos que consideremos de meirande relevancia hainos que indicar nalgunha parte visible do soporte material para que nunha primeira ollada, sen necesidade de escoitar a gravación, saibamos de que informante se trata, de onde procede, cal é o tema de conversa, a que área pertence, etc.

3. Conclusión

Tentamos expoñer algúns dos principais problemas e posibles solucións a ter en conta á hora de afrontarmos tarefas como a de recoller e transcribir mostras orais, tarefa que se ben en calquera lingua resulta necesaria como fase previa á de constitución e posterior manipulación de córpora lingüísticos, en linguas minorizadas coma o galego, cunha base de

usuarios limitada e baixo unha forte presión do castelán, vólvese fundamental e constitúe un *obxectivo en si mesmo*. Tomamos como base as nosas experiencias e desenvolvemos unha forma de proceder que pode ser contrastada coa metodoloxía empregada por outros investigadores, coa que ha converxer nalgúns puntos e diverxer noutros.

Referencias bibliográficas

- Beswick, J. E. (1998): “Observacións sobre as actitudes e os comportamentos relativos ós cambios de código en Santiago de Compostela”, *Cadernos de Lingua* 18, 53-78.
- Biber, D. & S. Johansson & G. Leech & S. Conrad & E. Finegan (1999): *Longman grammar of spoken and written English*, London: Longman.
- Caravedo, R. (1999): *Lingüística del corpus. Cuestiones teórico-metodológicas aplicadas al español*, Salamanca: Universidad de Salamanca.
- De Kock, J. (2001): *Gramática española. Enseñanza e investigación. I. Apuntes metodológicos. 7. Lingüística con corpus*, Salamanca: Universidad de Salamanca.
- Dubois, J. et al. (1994): *Diccionario de lingüística*, Madrid: Alianza.
- Du Bois, J.W. & S. Schuetze-Coburn & S. Cumming & D. Paolino (1993): “Outline of discourse transcription”, en Edwards, J.A. & M. D. Lampert (eds.), 45-90.
- Edwards, J.A. & M. D. Lampert (eds.) (1993): *Talking data. Transcription and coding in discourse research*, Mahwah, NJ: LEA.
- Fernández Rei, F.& C. Hermida (eds.) (1996): *A nosa fala: bloques e áreas lingüísticas do galego*, Santiago de Compostela: Consello da Cultura Galega.
- ILG (Instituto da Lingua Galega) (1990): “Introducción”, en *Atlas lingüístico galego*, volume I: *Morfoloxía verbal*, A Coruña: Fundación Barrié de la Maza, 9-24.
- Jahandarie, K. (1999): *Spoken and written discourse: a multi-disciplinary perspective*, Connecticut: Ablex Publishing Corporation.
- Kasher, A. (ed.) (1998): *Pragmatics. Critical concepts*. Vol. II: *Speech act theory and particular speech acts*, London/New York: Routledge.
- Knowles, G. (1996): “The value of prosodic transcriptions”, en Knowles, G. & A. Wichman & P. Alderson (eds.), 87-105.
- Knowles, G. & A. Wichman & P. Alderson (eds.) (1996): *Working with speech*, London/New York: Longman.
- McEnery, T & A. Wilson (2001): *Corpus linguistics*, Edinburgh: Edinburgh University Press.
- Real Academia Galega (1997): *Diccionario da Real Academia Galega*, A Coruña/Vigo: Real Academia Galega/Xerais/Galaxia.

- Real Academia Española (2001): *Diccionario de la lengua española*, Madrid: RAE.
- Sánchez Rodríguez, H. (1986): “Unha suxerencia metodolóxico-didáctica para a recollida e transcrición de textos orais”, *Adaxe*, 2, 195-209.
- Soto Andión, X. (1995): *Estudio lingüístico do galego da Terra de Montes*, tese de licenciatura, Universidade de Santiago, Facultade de Filoloxía, inédita.
- Taylor, L. (1996): “The compilation of the spoken English corpus”, en Knowles, G. *et al.* (eds.), *Op. cit.*, 20-37.
- Vaux, B. & J. Cooper (1999): *Introduction to linguistic field methods*, München: Lincom Europa.
- Vázquez-Monxardín, A. (coord.) (1998): *A cultura popular de tradición oral nos centros da terceira idade. Unha experiencia de recollida en Centros de Ourense, Santiago, Tui e Viveiro*, Santiago de Compostela: Consello da Cultura Galega.