

Nun volume anterior, o primeiro da serie de tres que constituirán o compendio do MSG, centrámonos na descrición e análise da competencia lingüística e da lingua inicial en Galicia. Unha das conclusións máis salientables naquel volume foi a constatación de que a práctica totalidade da poboación pode comprende-lo galego, e que algo máis do 85% pode falalo. Pero outra das conclusións, non menos salientable, foi a constatación dunha ruptura no proceso de transmisión interxeracional do galego, posta de manifesto na asociación que se percibe entre os datos correspondentes á lingua inicial e á idade dos individuos: a medida que descendemos na escala de idade, van sendo menores as porcentaxes dos que teñen o galego como lingua inicial.

Os datos sobre lingua inicial e competencia lingüística son, sen dúbida, de suma importancia; pero non serven por si sos para facer un diagnóstico da situación do galego. Unha descrición macrosociolingüística de Galicia necesita achegar datos precisos sobre o

uso que a poboación fai do galego e do castelán, e mais das actitudes, crenzas e prexuizos que ten ante as dúas linguas.

Este volume, dedicado ó uso das linguas, está estruturado en catro capítulos. No capítulo 3 presentamos unha análise da percepción que teñen os entrevistados acerca de cal é a súa lingua habitual (p. 14 do cuestionario incluído no apéndice 3), sen referirse a ningún ámbito específico nin a ningún interlocutor en concreto. Esta percepción compárase coa que teñen da lingua que empregaban habitualmente seus pais e os seus avós (p. 21, 23 a 26). Remata o capítulo cunha visión xeral da lingua habitual de lectura e de escritura (p. 73 e 76).

Os capítulos restantes son calas en certos ámbitos, que pretenden abarcala maior parte das situacións comunicativas nas que participan os individuos nas súas interaccións cotiás. O capítulo 4 céntrase no uso das dúas linguas nas interaccións no seo da familia, ámbito ó que se lle dedican no cuestionario dezasete preguntas (p. 27, 38 a 43 e 45 a 54), co obxecto de poder estudialas diferencias debidas ó interlocutor: a lingua falada habitualmente cos avós, cos pais, coa parella, cos irmáns e cos fillos.

O capítulo 5 ten como tema central a conducta lingüística no ámbito escolar, diferenciando entre a lingua usada para escribir (p. 85), para falar cos compañeiros na clase (p. 86) e no recreo (p. 87) e para falar cos profesores (p. 88). Tamén se inclúen neste capítulo datos sobre o estudio do galego, tanto na escola (p. 82 e 83) como fóra dela (p. 118 e 119), así como a percepción dos entrevistados acerca da posible influencia exercida pola lingua das clases sobre a súa lingua habitual (p. 92 e 93).

No capítulo 6 abordámo-la descrición e análise do uso que os falantes fan das linguas noutras situacións comunicativas: a lingua falada cos veciños galegos e cos de fóra de Galicia (p. 57 e 58), cos amigos (p. 59), cun descoñecido (p. 62), no traballo (p. 31 a 37), coas institucións (p. 65 e 66), na compra diaria e na compra non habitual (p. 60 e 61), co médico de cabeceira e co médico especialista (p. 67 e 68). Tamén se analizan neste capítulo dúas preguntas dirixidas a exploralos procesos de acomodación lingüística, tanto de converxencia como de diverxencia, tal como se manifestan no cambio ou mantemento da lingua habitual cando o interlocutor se atopa nun grupo que está a falar outra lingua (p. 114 e 115).

O traballo finaliza cun breve capítulo de conclusións, no que se recollen aqueles trazos máis relevantes postos de manifesto nos outros capítulos. Inclúense tamén tres apéndices: no primeiro deles aparecen os mapas elaborados para dar conta da distribución xeográfica dos usos; no segundo presentámo-las frecuencias das preguntas sobre usos lingüísticos incluídas no cuestionario; no terceiro e último recollémo-lo cuestionario utilizado. Finalmente, incluímos tamén unha *addenda* en forma de caderno coa actualización das táboas correspondentes ó conxunto de Galicia que figuraban no volume *Lingua inicial e competencia lingüística en Galicia*. Esta actualización fíxose a partir dos datos dispoñibles do Censo de Poboación e Vivendas de 1991.

Agás algunhas preguntas que só permitían unha resposta categórica do tipo 'si' ou 'non' (p. 37, 55, etc.) e outras con opcións de resposta exclusivas para elas (p. 93, 118, 119, etc.), a maior parte das cuestións sobre uso lingüístico incluídas no noso cuestionario teñen catro opcións de resposta, ademais do correspondente N.P. (Non Procede):

4. Só galego
3. Máis galego ca castelán
2. Máis castelán ca galego
1. Só castelán

Pódense considerar, xa que logo, como escalas de medición ordinais, por permitiren ordenalos entrevistados segundo o maior ou menor uso que fan do galego (ou do castelán). Sen embargo, en ningún dos capítulos que seguen se analizan cos procedementos que poderían ser axeitados para este tipo de medición, polo que cómpre ofrecer unha xustificación da nosa decisión.

A lingua habitual, analizada no capítulo 3, tratámola como unha variable nominal, é dicir, como unha variable cualitativa na que os números das opcións de resposta serven unicamente para identificalas, polo que os datos que se subministran son sempre porcentaxes; isto non supón ningún problema metodolóxico. Nos outros capítulos, en cambio, o número de variables analizadas, especialmente no capítulo 4, e maila necesidade de mante-la extensión deste volume dentro duns límites razoables, leváronnos a buscar algún

xeito de simplificación dos datos. No cuestionario distinguíase entre a lingua falada cos irmáns, coas irmás, cos fillos e coas fillas e, nos catro casos, cos máis novos e cos máis vellos. Estas oito preguntas reducíronse ó final a dúas variables: a *lingua falada cos irmáns* e a *lingua falada cos fillos*; as correlacións de rangos (de Spearman) entre as catro variables iniciais correspondentes a cada grupo (irmáns e fillos) son moi elevadas, superando o 0.90 en calquera dos casos, é dicir, están moi próximas á correlación perfecta. O mesmo sucede coas preguntas 51 e 52 (sobre a lingua usada polos entrevistados cos avós paternos e maternos) ou coas preguntas 23 a 26 (sobre a lingua habitual dos avós maternos, das avoas maternas, dos avós paternos e das avoas paternas), reducidas a outras dúas variables.

Cabía a posibilidade de escoller simplemente unha pregunta de cada grupo e considerala como representativa de todo o subconxunto, e quizais teríamos procedido así de non ser necesario aínda outro tipo de redución. Agás no capítulo 3, onde o foco de atención é a lingua habitual, no resto da análise o foco está na comparación entre unhas condutas lingüísticas e outras. No capítulo 4, por exemplo, do que se trata é, como xa dixemos, de compara-lo uso do galego por parte do entrevistado con distintos interlocutores no seo da familia: cos avós, cos pais, coa parella, cos irmáns e cos fillos. Se presentasémo-los datos mediante táboas de continxencia con porcentaxes para cada opción de resposta, as cinco columnas de que consta cada táboa converteríanse en 20, ou, alternativamente, habería que presentar cinco táboas semellantes ás do capítulo 3 para cada unha das táboas dos outros capítulos. Aínda que as conclusións ás que se chegaría serían as mesmas, a comunicabilidade das análises diminuíría en grao moi importante, e o lector veríase sumido nunha confusa masa de datos moi difícil de asimilar. Por outra banda, resultou imposible a utilización de programas de cuantificación de variables categóricas (PRIMALS, HOMALS e outros semellantes) por falta de capacidade das versións existentes para tratar unha cantidade de datos tan inxente coma nosa: a maioría destes procedementos comezan por construír unha matriz na que as columnas son as variables e as filas son os casos, o que supón que a nosa matriz tería 38.897 filas, cantidade moi superior á máxima que permite o *software* dispoñible. Ante estas dificultades, optamos por reducir cada pregunta a un único dato, considerando que os

números 1, 2, 3 e 4 que sinalan cada opción de resposta (de 'só castelán' a 'só galego') poden ser interpretados como puntuacións que nós lles asignamos ós individuos, de xeito que un entrevistado que fala só galego coa parella ten un 4 nesa pregunta, e outro que fala só castelán ten un 1. Xa que logo, para reducir varias preguntas a unha única variable calcúlase a media a través desas preguntas, de forma que alguén que nas catro preguntas referidas ós fillos respondeu 'só galego' en tres e 'máis galego ca castelán' nunha, terá unha puntuación de $15/4$, é dicir, 3.75. Obtemos así unha escala de medición que vai de 1 ('só castelán') a 4 ('só galego'), e na que o valor de 2.5 indica que non hai predominio de ningunha das linguas.

O procedemento non está exento de dificultades teóricas e, posiblemente, non será aceptado sen cuestionamento por tódolos lectores, xa que supón tratar unha escala ordinal coma se fose de intervalos iguais. Certamente, non se pode afirmar que un individuo que fala só galego nunha determinada situación (ó que lle asignamos un 4) fale o dobre de galego ca outro que respondeu que fala máis castelán ca galego (ó que lle asignamos un 2); tampouco podemos asumir que a distancia existente entre 'máis galego ca castelán' e 'só galego' sexa a mesma en tódolos casos, aínda que nós así o consideremos ó asignar a esas opcións sistematicamente as puntuacións de 3 e 4.

Malia as dificultades mencionadas, pódese dicir que, de feito, non se introduce ningún tipo de distorsión nas análises, sempre que o lector entenda o que realmente significan esas medias e como foron calculadas. Sería máis ortodoxo traballar con medianas ou con percentís, ou simplemente con porcentaxes; pero, tras explorar todas esas posibilidades, chegamos á conclusión de que os resultados eran os mesmos, só que menos comunicables.

Por outra banda, o tema de se é posible tratar variables ordinais como variables de intervalos iguais é obxecto dunha polémica xa antiga, e renovada con frecuencia, no ámbito da estatística inferencial. Os investigadores en ciencias sociais recorren habitualmente a ese tratamento na análise das escalas de actitudes e noutros moitos casos. Aínda que nós non facemos neste volume estatística inferencial, senón soamente unha estatística

descriptiva elemental, os razoamentos aducidos na práctica diaria da estatística inferencial en favor de tratar as escalas ordinais como escalas de intervalos son pertinentes tamén no noso caso, xa que o que pretendemos xustificar é o uso de medias a partir de números de orde, e as medias están implicadas en moitas das probas utilizadas en estatística inferencial, como a proba *t* ou a análise da varianza¹³. Trátase simplemente, pois, dunha estratexia metodolóxica que obedece a unha necesidade práctica: a de resumir e sintetizar unha serie de valores, organizándoos de xeito que parezan iguais nos seus intervalos, cando en realidade non é así, pero sen que a representación escollida resulte moi afastada dos feitos empíricos.

Como xa se explicaba no volume anterior, *Lingua inicial e competencia lingüística en Galicia*, as limitacións de espazo fan inviable a análise detida de cada un dos 34 sectores en que dividimos Galicia polo que, coma naquel caso, analizamos soamente seis sectores con características ben diferenciadas, conscientes de que con eles non se esgota a tipoloxía sectorial de Galicia. Na procura da coherencia e unidade co volume anterior, os sectores escollidos foron os mesmos:

- Provincia da Coruña: A Coruña cidade (sector 1) e Santiago (sector 2).
- Provincia de Lugo: Lugo-2 (sector 13) e Lugo-5 (sector 16).
- Provincia de Ourense: Ourense-4 (sector 21).
- Provincia de Pontevedra: Pontevedra cidade (sector 26).

Con esta selección pretendemos dar conta do que ocorre en:

a) Cidades grandes coa maioría da poboación concentrada no núcleo urbano, como A Coruña, fronte a outras máis pequenas e nas que o hábitat periurbano ten un papel destacable, como Santiago e Pontevedra (estas últimas diferenciadas á súa vez pola súa estrutura económica).

b) Sectores formados por vilas grandes, como Lugo-2 (Viveiro, Monforte) fronte ós que agrupan vilas pequenas, como Ourense-4 (Amoeiro, Carballeda, A Rúa, Rubiá).

¹³ Para unha xustificación máis detallada véxanse: Binder, A. (1984), "Restrictions on statistics imposed by method of measurement: some reality, some myth", *Journal of Criminal Justice* 12, pp. 467-481; Gaito, J. (1980), "Measurement scales and statistics: resurgence of an old misconception", *Psychological Bulletin* 87, pp. 564-567; Gaito, J. (1984), "Measurement scales and statistics: a confusion which refuses to die", *Canadian Psychology* 25, pp. 249-250; Gaito, J. (1986), "Some issues in the measurement-statistics controversy", *Canadian Psychology* 27, pp. 63-68; Traylor, M. (1983), "Ordinal and interval scaling", *Journal of the Market Research Society* 25; Zumbo, B. D. e Zimmerman, D. W. (1993), "Is the selection of statistical methods governed by level of measurement?", *Canadian Psychology* 34, pp. 390-399.

c) Sectores tipicamente rurais, como Lugo-5 (Abadín, Baleira, Ribeira de Piquín, Pol, Triacastela, ...).

A pesar de que en cada sector predomina un tipo de hábitat xeral (A Coruña: urbano; Lugo-2: vilego; Lugo-5: rural, etc.), non existe neles unha homoxeneidade completa. No deseño da mostra tentouse dar conta desta heteroxeneidade, polo que as entrevistas se aplicaron tanto na cabeceira dos concellos coma no resto dos lugares (parroquias, periferia das cidades, etc.). Xa que logo, na análise dos sectores, aínda que estes correspondan a un tipo de hábitat predominante, haberá que atender tamén á diferenciación interna de cada un. Por exemplo, en Ourense-4 sepáranse, cando é pertinente facelo, as vilas (núcleos dos concellos) das zonas etiquetadas como 'rural-1' ou 'rural-2', de acordo cos criterios metodolóxicos mencionados no primeiro capítulo. Nos sectores urbanos sepárase entre o urbano e o periurbano, distinción que cómpre ter moi en conta nos tres incluídos nesta selección, por mostraren unha grande disparidade na distribución do periurbano: mentres no concello da Coruña a poboación entrevistada fóra do núcleo urbano é o 2.3% da mostra correspondente, en Santiago acada o 29% e en Pontevedra constitúe o 51.6%. Cumprirá, polo tanto, prestar atención á distinción entre os datos relativos ó total da mostra deses tres sectores urbanos, por un lado, e os datos correspondentes ó que é territorio estritamente urbano (cidade) por outro.

Coma no volume anterior, preséntanse tamén datos relativos ó conxunto de Galicia. Posto que a unidade básica sobre a que se está a elabora-lo MSG é o sector, para construír unha mostra válida para o conxunto de Galicia non se poden agregar sen máis as mostras dos distintos sectores, senón que é necesario ponderalas dando conta do peso demográfico real de cada sector no conxunto de Galicia. Téñase en conta que sectores con escasa poboación necesitan, cando se toman como unidades de mostraxe independentes, un elevado número de entrevistas para acadar a fiabilidade mostral proposta (95%) manténdose na marxe de erro máximo establecida ($\pm 3\%$). Así, Coruña-11 (formado polos dous concellos da provincia que teñen menos de 5.000 habitantes, máis dun 33% da poboación ocupada na agricultura e máis doutro 33% na industria e na construción) ten, segundo o último Censo, 5.399 habitantes maiores de 15 anos, e necesita unha mostra de 795 individuos; en cambio, a cidade da Coruña, con 199.008

habitantes maiores de 15 anos¹⁴, só necesita 1.170 entrevistas para acadar a mesma fiabilidade e co mesmo erro mostral máximo. É evidente que se se integrasen sen correccións nunha mostra única para o conxunto de Galicia as diferentes mostras sectoriais, estarían sobrerrepresentados os sectores con menos poboación, e infrarrepresentados os sectores con máis poboación.

Entre as varias posibilidades de ponderación, escollemos agregarlos datos dos sectores segundo a fórmula seguinte:

$$P = \sum_{i=1}^n W_i \cdot P_i$$

onde

P = porcentaxe de respostas a unha opción dunha pregunta para o total de Galicia.

W_i = Ponderación que se aplica para cada sector.

P_i = porcentaxe de respostas a unha opción dunha pregunta nun sector determinado.

n = número de sectores.

W_i é simplemente a proporción que representa un sector determinado en relación co conxunto de Galicia; é dicir:

$$W_i = N_i / N$$

onde N_i é o número de individuos maiores de 15 anos nun sector determinado e N o número total de individuos maiores de 15 anos no conxunto de Galicia¹⁵.

Se as distintas mostras sectoriais tiñan xa unha fiabilidade e un erro máximo máis ca aceptables na investigación sociolóxica, a mostra do conxunto de Galicia sitúase en niveis óptimos, cun nivel de confianza do 95% e un erro máximo probable de $\pm 0,62\%$.

¹⁴ As discrepancias entre o número de habitantes que se menciona nestes exemplos e o que se mencionaba no volume primeiro débese a que neste caso dispoñemos dos datos do Censo de 1991.

¹⁵ Un infortunado erro fixo que o valor de N aparece mal explicado no volume anterior, onde se definía como o número de entrevistas feitas no sector. Na práctica, o procedemento de agregación descrito na fórmula desenvolveuse calculando para cada sector un factor multiplicativo, que consiste efectivamente en dividir a poboación do sector polo número de entrevistas feitas nel.

Finalmente, non podemos deixar de comentar un problema que aparece constantemente en investigacións deste tipo. Unha descrición macrosociolingüística do uso das linguas só pode dar conta da percepción que os falantes teñen dos seus propios usos, e non dos usos reais en si mesmos; a descrición destes últimos esixiría unha metodoloxía de investigación distinta da aquí empregada, concentrándose na análise dunhas poucas interaccións reais (e, xa que logo, con escasas posibilidades de facer xeneralizacións válidas). Con todo, cómpre dicir que nas situacións de bilingüismo se ten atopado un alto grao de correlación entre a conduta lingüística declarada polos individuos e a conduta lingüística observada polo investigador, polo que non parecen moi xustificadas as críticas de que é obxecto a investigación dos usos lingüísticos mediante cuestionarios; estes permiten, máis aló de calquera dúbida razoable, saber qué lingua se fala máis e cal se fala menos, e como se distribúen eses usos a través de distintos subgrupos da poboación. Outro tema, naturalmente, é o estudio das funcións comunicativas que pode representar para un galegofalante habitual o uso do castelán ou para un falante de castelán o uso do galego, ou o estudio das alternancias de lingua como estratexia discursiva, para construír ou redefinir o contexto ou a realidade social en que a interacción se desenvolve. Nada disto se pode estudar mediante un cuestionario, polo que é imprescindible, para unha comprensión axeitada da realidade sociolingüística galega, complementalos moi importantes datos que agora ofrecemos cos procedentes de investigacións realizadas desde outra perspectiva metodolóxica, investigacións que, felizmente, xa comezaron a desenvolverse na sociolingüística galega.