Jose Luis Alvarez Enparantza "Txillardegi"

POUR UNE ANALYSE MATHÉMATIQUE DES SITUATIONS BILINGUES

1. Dans une société bilingue on trouve des locuteurs *a*, monolingue en langue **A**; des locuteurs *b*, monolingues en langue **B**; et les locuteurs bilingues *x*, qui connaissent A et B, les deux langues en présence.

Si dans cette société (ou communauté bilingue, village, région, nation), on a un total de **N** locuteurs, on y trouvera $\mathbf{N_A}$ monolingues en langues **A**, $\mathbf{N_B}$ monolingues en langue **B**; et $\mathbf{N_X}$ bilingues. Et on pourra toujours poser:

$$N_A + N_B + N_X = N$$

et si nous divisons par N:

$$e_A + e_B + e_X = 1$$

$\boldsymbol{e_A}$, $\boldsymbol{e_B}$ et $\boldsymbol{e_X}$ sont les **proportions** respectives de gens capables de parler et comprendre sans aucun problème, et respectivement, la langue **A**, la langue **B**, ou **les deux** langues.

Nous savons qu'il n'y a de "bilingues parfaits" totalement équilibrés nulle part: les bilingues ont toujours une langue qu'ils maîtrisent mieux. Et il faut en tenir compte au moment de réaliser les recensements; ou d'en interpréter les résultats.

Cependant, pour tout ce qui suit (qui est une première approche à la réalité socio-linguistique), un locuteur *x* est un locuteur parfaitement bilingue. À l'opposé, un locuteur *a* est un locuteur monolingue absolu, qui ne parle et ne comprend que **A**; et, analoguement, un locuteur *b*, est un locuteur monolingue absolu, qui ne comprend et ne parle que **B**.

1.1. Si dans un village (ou ensemble fini de locuteurs) on a, par exemple, *3700* habitants, nous écrirons:

$$N = 3700$$

Et, si parmi eux, 1614 sont monolingues en **A**, 582 sont monolingues en **B**, et 1504 sont bilingues, nous pourrons poser:

$$e_A = \frac{1614}{3700} = 0,43621$$

$$e_B = \frac{582}{3700} = 0,1572973$$

$$e_x = \frac{1504}{3700} = 0,40648$$

Évidemment,

$$e_A + e_B + e_X = 0,9999873 \quad 1$$

Et on aura aussi

$$N_A + N_B + N_X = 3700$$

1.2. Il faut bien distinguer entre **NA**, nombre des habitants qui ***ne connaissent que*** la langue **A**; et **N'A**, nombre des habitants, bilingues ou pas, ***qui connaissent la langue*** **A**.

On écrira donc:

$$N'_A = N_A + N_X = 1614 + 1504 = 3118$$
$$N'_B = N_B + N_X = 582 + 1504 = 2086$$

Et en divisant par N:

$$e'_A = e_A + e_X = 0,84269$$
$$e'_B = e_B = e_X = 0,5637773$$

Ceux qui **ne parlent que *A*** sont:

$$e_A = 0,43621 \text{ (monolingues } \mathbf{a}\text{)}$$

et ceux **qui ne parlent que *B***:

$$e_B = 0,1572973 \text{ (monolingues } \mathbf{b}\text{)}$$

Mais ***tous ceux*** qui parlent **A** (monolingues **a** + bilingues) sont:

$$e'_A = 0,84269$$

et ***tous ceux*** qui parlent **B** (monolingues **b** + bilingues):

$$e'_B = 0,5637773$$

On aura donc toujours:

$$e_A + e_B = 1 - e_X \leq 1$$

Et symétriquement:

$$e'_A + e'_B = 1 + e_X \geq 1$$

Il suffit de poser:

$$e'_A + e'_B = e_A + e_B + 2e_X = 1 + e_X$$
$$e_A + e_B = 1 - e_X$$

1.3. Dans les pays à langue B propre, minorisée, plus ou moins tolérée; et langue A officielle, et seule réellement nécessaire, par une mécanisme plus ou moins inconscient et innocent, on a tendance à confondre les choses.

On dira que **N** est la même chose que **N'_A**; et on comparera **N'_B** avec **N'_A** (mais jamais **N_B** avec **N'_A**); ce qui fausse foncièrement la situation réelle des deux langues, déjà même au niveau de leur connaissance respective.

1.31. Si on se réfère, par exemple, à notre village ci-dessus, on dira (en on croira même, surtout aux niveaux "officiels") qu'il y a **N_A qui** parlent la langue **A**; face à **N_B** qui parlent la langue **B** (c'est-à-dire, 1614 face à 2086).

On voudra même faire croire, en conséquence, que la langue **B** n'est nullement minorisée, puisqu'elle est "même majoritaire et dominante" (!): 2086 personnes parlent la langue **B**, face à 1614 qui connaissent la langue **A** (2086 > 1614).

Il s'agit d'un sophisme; puisqu'on compare **N_A** (nombre de ceux qui **ne parlent que A**) avec **N'_B** (tous ceux qui parlent **B**). On oublie que les bilingues, qui se cachent dans les décomptes, **connaissent aussi A**.

Si on compare, **comme il faut, N'_A** avec **N'_B** (tous ceux **qui connaissent A** avec tous ceux **qui connaissent B**), les constats sont bien différents:

$$N'_A = 3118 > N'_B = 2086 \text{ (langue A dominante)}$$

Ou si, inversement, on compare **N_A** (nombre de ceux qui **ne parlent que A**) avec **N_B** (nombre de ceux **qui ne parlent que B**), on constate la même dominance de A:

$$N_A = 1614 > N_B = 582$$

Je suppose qu'en Bretagne on fait pareil; et que l'on compare volontiers le nombre **N'_B** de ceux *qui parlent breton*; au nombre **N_A** de ceux qui *ne parlent pas le breton* (qui ne parlent que français).

On voit que ces procédés dissimulent la gravité objective de la situation de la langue minorisée. Parce que, à la limite, et sans trop exagérer, "tout le monde" parle aujourd'hui la langue officielle et dominante (aussi bien au Pays Basque qu'en Bretagne); et parce que l'ensemble de ceux qui connaissent la langue minorisée (Basque, Breton) n'est en fait qu'un sous-ensemble presque identique à l'ensemble (proche de la totalité) de la population bilingue.

Nous y reviendrons.

2. On va réfléchir maintenant à **l'utilisation** des langues en milieu bilingue.

Commençons par les groupes de **2 personnes**.

Il existe 9 types de couples **possibles**, et seulement 9: *aa, ab, ax; ba, bb, bx; xa, xb, xx.*

Or quelles sont les possibilités de parution *p* de chacun de ces types? Nous sommes face à un problème élémentaire de calcul de probabilités.

Quelle est la **probabilité** de trouver, au hasard, un groupe de type *aa*? (deux monolingues de langue **A**, face-à-face?)

Évidemment:

$$p_{aa} = e_A \cdot e_A = e_A{}^2$$

Analoguement, quelle est la possibilité de trouver, au hasard, un groupe *ax*? (un couple de locuteurs, constitué par un monolingue **a**, et un bilingue **x**; soit *ax*, soit *xa*):

$$p_{a \cdot x} = p_{x \cdot a} = e_A \cdot e_X$$

$$\mathrm{p_{xa} = 2e_Ae_X} \text{ (si l'ordres des éléments n'a pas d'importance)}$$

2.1. Si nous revenons à notre village:

$p_{aa} = e_A{}^2 = 0{,}43621^2 = 0{,}19027$

$p_{ab} = 2e_Ae_B = 2 \cdot 0{,}43621 \cdot 0{,}1572971 = 0{,}13722$

$p_{ax} = 2e_A \cdot e_X = 2 \cdot 0{,}43621 \cdot 0{,}40648 = 0{,}35462$

$p_{bb} = e_B{}^B = 0{,}1572973^2 = 0{,}024742$

$p_{bx} = 2 \cdot 0{,}1572973 \cdot 0{,}40648 = 0{,}12787$

$p_{xx} = e_x{}^2 = 0{,}40648^2 = 0{,}16522$

3. Voyons maintenant quelle est ***la langue choisie pour la communication*** dans les différents groupes:

aa: on y choisit la langue **A** nécessairement, unique langue commune aux interlocuteurs.

ab: pas de communciation possible, Ø; l'un étant monolingue en **A** et l'autre en **B**.

ax: le bilingue **s'adapte** à l'interlocuteur qu'il a en face. C'est bien le monolingue qui **impose la langue de communication**, pas le bilingue.

ba: pas de communication possible, Ø.

bb: langue choisie, **B**.

bx: …**B**. Le bilingue s'adapte au monolingue **b**.

xa: **…**A, le monolingue *a* imposant sa langue A au bilingue.

xb: …B, le monolingue impose sa langue B.

xx: les deux étant bilingues, on **ne sait pas** quelle sera la langue choisie.

4. Probabilité totale de l'emploi des deux langues.

Langue A: sera employée dans les couples de types 3.1, 3.3 et 3.7: ***aa, ax, xa.***

La probabilité pour ces couples étant:

$$p_{2A} = 0{,}19027 + 0{,}35462 = 0{,}54489$$

Langue B:

$$p_{2B} = 0{,}024742 + 0{,}12787 = 0{,}152612$$

Langue X:

$p_{2X} = 0{,}16522$ (en fonction de la **loyauté** des bilingues, A ou B)

Pas de communication possible:

$$p_{2\emptyset} = 0{,}13722$$

Nous ignorons la langue choisie dans les couples *xx* (qui constituent 16,522% des couples possibles).

$$0{,}54489 \leq p'_{2A} \leq 0{,}710168$$
$$0{,}152612 \leq p'_{2B} \leq 0{,}317832$$

C'est-à-dire que le total $\boldsymbol{p'_{2B}}$ des groupes qui utilisent **B** se compose en deux parties:

— il y a d'un côté $\boldsymbol{p_{2B}}$, la proportion des couples qui parlent B **obligatoirement** (0,024742);

— et il y a la partie provenant des groupes ***b.x*** et ***x.b***, 0,1572973 · 0,40648 = 0,063938 (x 2 = 0,12787). Donc:

0,024742 + 0,12787 = 0,152623 (**B** obligatoirement aussi)

Mais il y a une partie de la communication ***entre bilingues***, e_{2X}, qui se fait en B aussi: en fonction de la **loyauté** des bilingues par rapport à la langue B, dans les cas (*bb*) où de choix est possible, et dépend exclusivement du choix libre des bilingues impliqués.

Si on définit cette loyauté entre bilingues par rapport à la langue B, comme la proportion des cas dans lesquels les bilingues choisissent B:

$$0 \leq m_B \leq 1$$

Si les bilingues, entre eux, choisissent **toujours** la langue B,

$$m_B = 1$$

et s'il choisissent **toujours** la langue A,

$$m_B = 0$$

Et en général:

$$\boldsymbol{p'_{2B} = p_{2B} = m_B \cdot p_{2x}}$$

5. Si nous nous référons à notre village et supposons que les bilingues ont une loyauté $m_B = 0{,}8$ (ils emploient B entre eux dans 80% des cas):

$$p'_{2B} = 0{,}152612 + 0{,}8 \cdot 0{,}16522 = 0{,}284788$$

et symétriquement:

$$p'_{2A} = 0{,}54489 + 0{,}2 \cdot 0{,}16522 = 0{,}577934$$

la proportion des couples qui ne peuvent pas communiquer étant:

$$p_{2\emptyset} = 0{,}13722$$

5.1. Et en général les totaux pour les emplois de **A** et **B** sont respectivement:

$$p'_{nA} = (1 - e_B)^n - e_X^n$$

$$p'_{nB} = (1 - e_A)^n - e_X^n$$

$$p_{nX} = e_X^n$$

et par conséquent:

$$e_A + e_X = 1$$

À ce stade les expressions se simplifient:

$$p'_{nA} = 1 - e_X^n$$

$$p'_{nB} = 0$$

$$p_{nX} = e_X^n$$

$$p_{n\emptyset} = 0$$

p_{nX} se partage en proportion de la loyauté des bilingues.

Donc on a finalement:

$$p'_{nB} = m_B \cdot e_X^n$$

Si les bilingues n'ont aucune loyauté envers B, cette langue disparaît de la circulation. La langue B n'est jamais plus nécessaire pour l'intercompréhension: on peut toujours s'entendre en langue A.

6.1. Si l'on veut que la langue B soit utilisée, par exemple, par 50% des couples, on écrira:

$$0,5 = m_B \cdot e_X^2$$

Ce qui nous donne la proportion de bilingues nécessaire pour qu'il y ait équilibre dans l'utilisation au sein des couples:

$$e_X = \sqrt{\frac{1}{2m_B}}$$

6.11. Si on pose, par exemple, m_B = 0,8, on obtient:

$$e_X = \sqrt{\frac{1}{2 \cdot 0,8}} = 0,79056$$

Si le nombre de bilingues est inférieur à 79,056%, il est impossible d'atteindre un emploi équilibré des deux langues dans les couples.

On comprend facilement que dans les communautés dites "bilingues", la présence réelle de la langue B — au niveau de la communication sociale — devienne souvent imperceptible; même si la loyauté des bilingues est forte.

7. Isotropie et anisotropie socio-linguistique

En Pays Basque nous mesurons, depuis plusieurs années, l'emploi de deux langues "dans la rue".

Et nous avons constaté, systématiquement, que le niveau d'utilisation de la langue minorisée B est nettement supérieur à celui qu'on pourrait espérer en fonction des niveaux e de connaissance des deux langues. Le phénomène est extrêmement net dans les villages où la proportion des bascophones est très bas.

Nous avons dû nous rendre à une évidence empirique: les rapports linguistiques ne sont pas isotropiques. Une partie $g \cdot N_A$ *des monolingues a est en rapport, linguistique, ave*c les bilingues; mais le reste des monolingues *a,* $(1 - g) \cdot N_A$ fonctionne en vase clos, sans rapport linguistique avec la population bilingue "autochtone" — si l'on peut dire —. Ceci améliore un peu le niveau d'utilisation de la langue B.

Le calcul montre facilement qu'on a:

$$p'_{nB} = m_B \cdot \frac{e_X^n}{(ge_A + e_X)^{n-1}}$$

p'_{nB}, utilisation totale de la langue B dans les groupes à *n* personnes

m_B, loyauté des bilingues envers la langue B

e_X, niveau de connaissance des deux langues dans la communauté

g, proportion de la population majoritaire qui est intégrée et a des rapports linguistiques avec les autochtones

e_A, proportion des habitants monolingues *a*

Cette formule est valable dans tous les cas. Si on admet g = 1 (la totalité des monolingues *a* en langue majoritaire vivant en rapport linguistique avec les bilingues autochtones), on obtient:

$$p'_{nB} = m_B \cdot e_X^n$$

qui est l'expression utilisée dans l'hypothèse isotropique initiale: tous les habitants en rapport linguistique avec tous les habitants.

l'***isotropie*** n'est qu'un cas extrême des anisotropies possibles (g = 1).

La ***ségrégation totale*** ("deux communautés", étanches entre elles), correspont à g = 0, anisotropie maximale.

Du point de vue de l'utilisation de la langue B, le niveau optimum d'utilisation de B correspond à g = 0 (anisotropie totale); l'isotropie totale, g = 1, donnant l'utilisation minimum de B.

7.1. Si l'on revient à notre village: qu'on y pose $m_B = 0{,}8$ (bilingues très loyaux à B: 80% d'utilisation entre eux), et qu'on y suppose, par exemple, g = 0,3 (c'est-à-dire, 30% de monolingues *a* sont linguistiquement "intégrés" à la communauté originaire; tandis que les 70% restant vivent sans rapport avec les autochtones), nous trouvons (pour la situation dans les couples, n = 2):

$$p'_{2B} = 0{,}8 \cdot \frac{0{,}40648^2}{0{,}3 \cdot 0{,}43621 + 0{,}40648} = 0{,}24584$$

(presque un quart des couples utilisent la langue B).

Ce chiffre est nettement supérieur à celui qu'on trouverait dans l'hypothèse naïve de rapports isotropiques:

$$p'_{2B} = 0{,}8 \cdot 0{,}40648^2 = 0{,}13218$$

L'anisotropie des rapports linguistiques améliore (en la doublant presque) l'utilisation de la langue minoritaire B.

7.2. D'un autre côté, si on connaît le niveau d'**utilisation** des deux langues (empirique) et le niveau respectif de **connaissance** des langues dans la communauté, on peut en déduire la **structure socio-linguistique** (niveau d'anisotropie) et la **loyauté linguistique** des bilingues.

En effet:

$$p'_{2B} = m_B \cdot \frac{e_X^2}{ge_A + e_X} \text{ (utilisation de B dans les couples)}$$

$$p'_{3B} = m_B \cdot \frac{e_X^3}{(ge_A + e_X)^2} \text{ (utilisation dans les groupes de 3 personnes)}$$

Si nous mesurons **empiriquement** l'utilisation dans les groupes de 3 pesonnes de la langue B; et que nous connaissons, empiriquement aussi, les niveaux de connaissance des deux langues, nous pourrons en déduire **l'anisotropie socio-linguistique et la loyauté des bilingues**.

En effet:

$$A = ge_A + e_X$$

$$ge = \frac{A - e_X}{e_A}$$

$$m_B = \frac{A \cdot p'_{2B}}{e_X^2}$$

7.3. Faisons ces calculs pour notre village.

On y avait:

$N_A = 1187$ (monolingues en **A**)

$N_X = 2425$ (bilingues) ("majorité" apparente de B)

$N = 3612$ habitants

Ceci donne:

$e_A = 0{,}3286268$ (33% environ de monolingues *a*)

$e_X = 0{,}6713732$ (67% de bilingues; B apparemment majoritaire)

Supposons maintenant qu'on a **mesuré** dans la rue l'utilisation de B au sein des couples et des groupes de 3 personnes:

$$p'_{2B} = 42\%$$

$$p'_{3B} = 34\%$$

Que peut-on en déduire?

$$A = e_X \cdot \frac{p'_{2B}}{p'_{3B}} = 0,82934$$

$$g = \frac{A - e_X}{e_A} = 0,48068$$

$$m_B = \frac{A \cdot p'_{2B}}{e_X^2} = 0,77277$$

Seulement une moitié (g = 48,068%) des monolingues *a* est en rapport linguistique avec la population bilingue autochtone.

D'un autre côté, malgré les lectures habituelles (totalement erronées) sur le "manque" d'intérêt des bilingues pour leur langue, la population bilingue nous apparaît très loyale par rapport à sa langue autochtone (m_B = 77,277%°).

Par ailleurs, la langue officielle A est largement dominante au niveau de l'utilisation **réelle**:

p'_{2A} = 58% (des couples utilisent A)

p'_{3B} = 66% (des groupes de 3 personnes)

Si l'on tient compte de la relativité bien connue en degré de bilinguisme, et de la fiabilité très relative des recensements linguistiques, on en déduit une précarité bien réelle dans la situation des langues minorisées. Même dans des cadres globaux apparemment favorables.

JOSE LUIS ALVAREZ ENPARANTZA
Prof., Université du Pays Basque