

Les dialectes basques : homogénéité ou dispersion ?

Michel Etchebarne

► **To cite this version:**

Michel Etchebarne. Les dialectes basques : homogénéité ou dispersion ?. Lapurdum VIII - Revue d'études basques, Centre de recherche sur la langue et les textes basques UMR 5478, 2003, pp.139-166. <artxibo-00000059>

HAL Id: artxibo-00000059

<https://artxiker.ccsd.cnrs.fr/artxibo-00000059>

Submitted on 31 Jan 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DIALECTES BASQUES : HOMOGÉNÉITÉ OU DISPERSION ?*

L'ATLAS LINGUISTIQUE : UN OUTIL PRIVILÉGIÉ POUR LA CONNAISSANCE DIALECTALE

La présente note propose une approche quantitative et un commentaire succinct de l'atlas ethnolinguistique du Pays Basque (*Euskalerriko atlas etnolinguistikoa, EAEL*) réalisé par la société scientifique Aranzadi (Saint-Sébastien). Cette enquête s'inscrit dans le cadre des recueils de données lexicales opérés selon la méthodologie des atlas linguistiques.

Le questionnaire initial comportait 570 questions (une majorité de « mots », quelques syntagmes et une quarantaine de phrases). Les mots testés (lexèmes-sources) étaient exprimés en langue française pour les questionnaires administrés au Pays Basque Nord et en espagnol pour les questionnaires administrés au Pays Basque Sud. Les informateurs devaient indiquer le terme basque (lexème-cible) correspondant au lexème-source testé. Le questionnaire n° 1 testait les lexèmes-sources *tête/cabeza*. Tous les informateurs associèrent à ces lexèmes-sources un même lexème-cible, à quelques variantes phonologiques près : *buru*.

Les questionnaires furent administrés entre 1977 et 1983 à des informateurs résidant en 190 points du domaine basque représentatifs de l'ensemble des dialectes. Quatre-vingts enquêtes parmi les 190 furent publiées. Les 80 informateurs d'*EAEL*, 60 % d'hommes, 40 % de femmes, étaient âgés de 69 ans en moyenne lors de l'enquête, et étaient donc natifs du début du 20^e siècle. La sélection des enquêtes publiées se fit en s'appuyant sur le découpage dialectal et sous-dialectal réalisé par le prince L. L. Bonaparte, repris en 1981 par Pedro de Yrizar dans sa *Contribución a la dialectología de la lengua vasca*. Parmi les 570 questionnaires, 27 ne furent pas publiés car ils n'obtinrent pas ou peu de réponses, étant trop éloignés de la culture locale.

EAEL fut publié en deux tomes, en 1983 et 1990.

Chaque question donne lieu à une fiche de présentation occupant deux pages. Sur la page de droite figure, en en-tête, la question posée (lexème-source) en français, anglais et espagnol (fiche 1 : *tête, head, cabeza*), puis les 80 réponses données par les informateurs (lexèmes-cibles). Les réponses sont subdivisées en 7 zones dialectales et 80 communes : Alava (3 communes), Basse-Navarre (6 communes), Biscaye (13 communes), Guipuscoa (18 communes), Labourd (7 communes), Navarre (26 communes), Soule (7 communes). Sur la page de gauche figurent une carte où sont symbolisés les 80 points où le questionnaire a été administré et les réponses faites. Un comptage des réponses et quelques commentaires accompagnent parfois la carte.

* DEA d'études basques, 2003, CIEB, Bayonne, direction : C. Videgain. Je remercie mes informateurs de Soule : Teija Eppherre et Panpeia Etxebarne.

L'exploitation des résultats d'un atlas linguistique n'est pas une entreprise aisée. La quantité d'information récoltée est d'une part particulièrement riche, ce qui pose d'emblée la question du degré et du type de synthèse à opérer ; certaines questions de méthode devant d'autre part recevoir des réponses dès le début de l'exploitation.

S'agissant des questions de méthode, je ne détaillerai pas ici l'ensemble des problèmes que soulève l'exploitation d'un tel corpus. Il est néanmoins indispensable de cerner quelques protocoles d'élaboration de l'atlas avant de proposer un commentaire. L'enquête est le paradigme fondateur de la méthode de recherche en sciences sociales. Quatre étapes rythment traditionnellement l'enquête : 1) le projet d'enquête ; 2) la mise en œuvre des méthodes et des techniques d'enquête ; 3) l'analyse des données ; 4) la présentation des résultats. Les étapes 2) et 3) sont fortement solidaires car la qualité et la rigueur mises en œuvre lors de l'étape 2) influenceront sur la pertinence des étapes 3) et 4). Il est connu en épistémologie des sciences sociales que toute intervention auprès d'informateurs crée, parallèlement au recueil d'information proprement dit, des « perturbations » qui vont influencer sur les réponses.

Ainsi, le recueil de lexèmes-cibles basques à partir de lexèmes-sources appartenant à deux langues de départ différentes, le français et l'espagnol, pose des difficultés théoriques et pratiques. On sait en effet que les significations de deux lexèmes appartenant à deux langues différentes — même s'ils sont équivalents au point de vue du sens — ne se superposent pas totalement. L'espagnol et le français ne découpent pas toujours le réel de la même manière. Le terme *bois* (matière), en français, reçoit deux traductions possibles en espagnol : *leña* « bois brut » et *madera* « bois travaillé ». Or le questionnaire n° 126 d'*EAEEL* testait les lexèmes-sources *bois*, au Pays Basque Nord, et *madera*, au Pays Basque Sud. Symétriquement, lorsque le français et l'espagnol disposent de termes synonymes, ceux-ci peuvent ne pas recevoir un équivalent basque de même périmètre : *sœur/hermana* reçoit deux équivalents en basque : *ahizpa* « sœur d'une femme » et *arreba* « sœur d'un homme ». Des termes tels que *foyer* en français et *hogar* en espagnol (*EAEEL*¹²³) présentent une polysémie de même périmètre : « lieu où l'on fait du feu » et « lieu où l'on vit » (les termes français et espagnol ont ici la même étymologie : latin populaire *focarium*, du latin *fōcus* « foyer, feu »). Ce double sens se retrouve dans les réponses faites par les informateurs lorsque l'enquêteur n'a pas levé l'ambiguïté, certains informateurs ayant répondu en se référant à la première acception (*supazter*, *sutondo*, *suete*, etc.), d'autres à la seconde (*etxe*, *bizi leku*).

Toute enquête ethnolinguistique est contrainte de médiatiser son projet par un ou plusieurs truchements. Quel que soit le soin que l'on mettra à affiner la « neutralité » des protocoles de recueil des données, la médiatisation est inévitable. On ne peut « faire dire » un lexème à un informateur sans utiliser un biais et celui-ci pourra toujours être vu comme comportant des effets parasites (une photographie n'est pas toujours plus explicite qu'un lexème-source). Si la neutralité absolue est un objectif hors d'atteinte dans une enquête en sciences sociales, il est néanmoins possible d'en atténuer les effets indésirables en utilisant des procédures d'évaluation, de recoupements et de contrôles a posteriori.

Tout comptage lexical pose d'emblée la question cruciale de savoir à partir de quel degré de variation de forme on passe d'un lexème à deux lexèmes : *esan* (aire occidentale du Pays Basque) et *erran* (aire orientale du Pays Basque), qui ont tous deux le sens de « dire » (*EAEEL*²²), sont-ils un seul et même « mot » ou deux « mots » ? ; *arrats* et *arratsalde*, « soir »

(*EDEL*⁵⁴¹), sont-ils un seul et même mot ou deux « mots » ? La solution retenue ici pour démarquer les différentes formes de lexèmes-cibles a été la suivante :

— lorsque la différence entre deux formes se situait au seul niveau du phonème, j’ai considéré que l’on était en présence d’un mot unique (*esan* et *erran* « dire » = un mot) ;

— lorsqu’elle se situait au niveau du morphème, j’ai considéré que l’on avait affaire à deux mots (*arrats* et *arratsalde*¹ « soir » = deux mots ; *orratz* et *jostorratz*² « aiguille » = deux mots).

Si l’on considérait en effet qu’*esan* et *erran* étaient deux unités lexicales distinctes, on se heurterait à une double difficulté. À supposer que l’on ait déterminé une empreinte phonologique des citations des locuteurs (ce qui n’a pas été le cas d’*EDEL*), on amplifierait de la sorte la dispersion lexicale de manière exagérée tant du point de vue de la gestion des données que de celle de la logique de la langue. Même s’il est très utile au plan de la phonologie de la langue de connaître les diverses réalisations d’une même unité lexicale, il n’est pas pertinent de considérer comme unités lexicales distinctes des termes tels qu’[eśan] et [er : an] « dire » ou [tşakur] et [sakur] « chien ». A contrario, si l’on considérait qu’*arrats* et *arratsalde* étaient un seul et même mot, il nous faudrait alors considérer, en raison de la présence d’une même base, que des lexèmes tels que *behatz* et *atzamar* (« doigt ») sont un seul et même mot, de même qu’*ilargi* et *argizagi* (« lune ») ou *emakume* et *emazteki* « femme », etc. Une telle approche serait au contraire trop englobante et incompatible avec la démarche lexico-statistique suivie ici, qui consiste à mesurer des ressemblances et des différences interdialectales³. À cet égard, il est important de souligner que la notion d’intercompréhension n’entre pas directement en jeu dans la définition qui vient d’être donnée de l’unité lexicale⁴.

Ce point de définition acquis (tous les comptages proposés sont élaborés à partir du protocole précité), on peut établir des statistiques lexicales selon deux parcours :

— en partant des lexèmes-sources (parcours onomasiologique) et en établissant des corrélations avec les lexèmes-cibles (combien de lexèmes-sources ont-ils généré un seul et même lexème-cible ?, combien ont-ils généré deux lexèmes-cibles ?, etc.) ;

— en partant des lexèmes-cibles (parcours sémasiologique) et en observant leurs formes et leurs sens (combien de monosèmes, combien de polysèmes ayant une seule et même acception principale, combien de polysèmes ayant au moins deux acceptions principales ?, etc.).

¹ *arrats* « soir » + *alde* « côté, partie ».

² *josi* « coudre » + *orratz* « aiguille ».

³ Les géosynonymes comportant une même base comme *behatz* et *atzamar* ; *ilargi* et *argizagi* ; *emakume* et *emazteki* ne sont pas interchangeables pour les locuteurs d’un même parler. Là où l’on utilise une forme, on n’utilise pas l’autre. Or le niveau visé dans la présente étude est précisément celui du lexème : dans le cadre de ce protocole, [eśan] et [er : an] sont un même « mot » mais *behatz* et *atzamar* sont deux « mots ».

⁴ Une différence se situant au seul niveau du phonème peut empêcher la compréhension ([eśan] ne sera généralement pas compris de ceux qui ne connaissent qu’[er : an], et vice versa). A contrario, des géosynonymes n’appartenant pas à leur dialecte seront parfois compris par les locuteurs (soit parce ceux-ci ont acquis une compétence interdialectale, soit parce que le contexte leur permet d’en « deviner » le sens).

PLAN DE LA NOTE

1/ Présentation des données

11/ Point de départ : lexèmes-sources

12/ Point de départ : lexèmes-cibles

121/ Distribution géographique et sémantique des lexèmes-cibles

122/ Géosynonymes ayant/n'ayant pas de morphème en commun

13/ Dispersion dialectale et catégories grammaticales

2/ Commentaire

21/ Polysèmes ayant au moins deux acceptions principales

22/ « Faux amis » : un phénomène linguistique universel

23/ Une même réalité observée sous deux angles

24/ Convergence ou divergence ?

Annexe 1 Données recueillies par *EAEI* : aperçu général

Annexe 2 Lexique commun

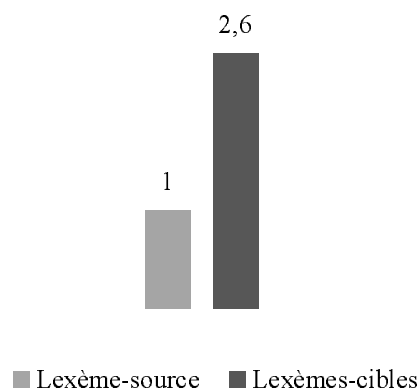
Annexe 3 Lexique quasi commun

1/ PRÉSENTATION DES DONNÉES

11/ Point de départ : lexèmes-sources⁵

La figure suivante rapporte le total des lexèmes-sources (520) au total des lexèmes-cibles (1335) recensés dans l'étude. Un lexème-source a généré 2,6 lexèmes-cibles en moyenne.

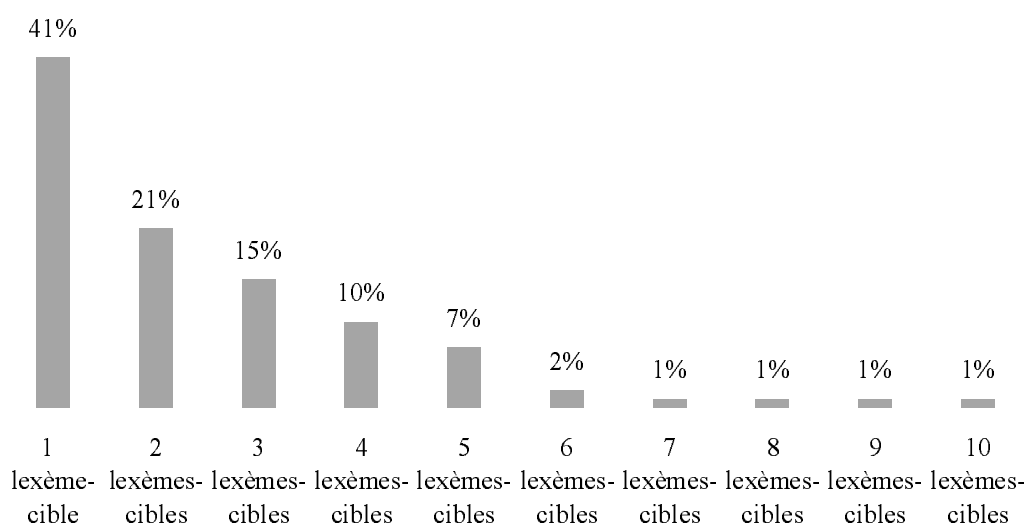
Fig. 1 — *Un lexème-source a généré 2,6 lexèmes-cibles en moyenne.*



⁵ On gardera à l'esprit qu'en réalité les informateurs n'ont pas été mis en présence d'une idée (ou d'un concept ou d'un signifié) pour laquelle il leur aurait été demandé de signaler le lexème-cible basque correspondant, mais d'un lexème-source administré en français ou en espagnol. C'est donc par facilité de langage que je parle ici de parcours onomasiologique. On sait en effet depuis Saussure que « Si les mots étaient chargés de représenter des concepts donnés d'avance, ils auraient chacun, d'une langue à l'autre, des correspondants exacts pour le sens ; or il n'en est pas ainsi. » (*CLG*)

La figure suivante représente la dispersion des réponses.

Fig. 2 — Pourcentages de lexèmes-sources ayant généré 1, 2, 3, etc. lexèmes-cibles.



Lecture du tableau :

— 41 % des lexèmes-sources ont généré une seule et même réponse de la part de tous les informateurs (« abeille » → *erle*)⁶ ;

— 21 % des lexèmes-sources ont généré deux lexèmes-cibles (« naître » → *jaio, sortu*) ;

— 15 % des lexèmes-sources ont généré trois lexèmes-cibles (« doigt » → *behatz, eri, atzamar*) ;

— 10 % des lexèmes-sources ont généré quatre lexèmes-cibles (« rivière » → *erreka, ibai, ugalde, ühaitz*) ;

— 7 % des lexèmes-sources ont généré cinq lexèmes-cibles (« arriver » → *heldu, ailegatu, etorri, jin, iritsi*) ;

— 2 % des lexèmes-sources ont généré six lexèmes-cibles (« arc-en-ciel » → *ortzadar, Jainkoan pasa, Erromako zubi, San Migelen zubi, arkoiris, ustruke*) ;

— 4 % des lexèmes-sources a généré plus de six lexèmes-cibles (« rapidité » → *azkar(tasun), arin(tasun), bizkor(tasun), zale(tasun), presa(ka), fite, agudo*).

Peut-on en déduire que tous les bascophones ont un même lexique en commun pour exprimer environ la moitié des « concepts » de base (parcours onomasiologique) ?

La réponse à cette question me semble devoir être positive, aux réserves suivantes près :

— L'homogénéité/dispersion lexicale dépend de la manière dont on définit la limite où, malgré des ressemblances, on va estimer qu'un mot n'est plus un seul mot mais deux mots. Comme précisé plus haut (voir protocole de comptage), j'ai considéré que si une différence phonématique entre deux formes n'était pas discriminante, une différence morphématique l'était.

— La dispersion lexicale dépend de phénomènes liés aux conditions d'administration des questionnaires. Comme on l'a vu, l'utilisation de deux langues-sources a vraisemblablement

⁶ Cela signifie que si l'on feuillette *EAEL*, on trouvera en moyenne près d'une page sur deux où les lexèmes-cibles signalés par les quatre-vingts informateurs sont, à des variantes phonologiques près, identiques.

joué dans le sens de l'accentuation des différences. La qualité des réponses recueillies dépend également du degré d'implication des enquêteurs et des informateurs dans la procédure de collecte des données.

— La dispersion dialectale telle qu'elle est mesurée est indépendante du nombre de locuteurs. À côté d'un mot connu par la majorité des locuteurs (*eguzki* « soleil » et ses variantes phonologiques représentent autour de 98 % des locuteurs), un mot ayant le même sens peut n'être connu que par une minorité de locuteurs (*ekhi* : autour de 2 % des locuteurs). Cela suffit pour considérer le mot connu par la majorité des locuteurs (*eguzki*) comme ne faisant pas partie du lexique commun à tous les bascophones⁷.

Si l'on s'accorde sur ces principes et en se basant sur l'enquête d'*EACL*, on peut sans doute admettre que la moitié des « concepts de base » s'exprime au moyen d'un lexique commun à tous les locuteurs (« chien » se signale partout, malgré des variantes phonologiques, par le même lexème : [tʃakur], [sakur], [tsakyr], [ʃakur]).

12/ Point de départ : lexèmes-cibles

121. Distribution géographique et sémantique des lexèmes-cibles

Le tableau suivant présente un échantillon des 1335 lexèmes-cibles recueillis par *EACL* en les organisant selon qu'ils sont connus partout ou connus en certains lieux seulement de l'aire bascophone, et selon qu'ils sont monosémiques (lexèmes n'ayant qu'une seule acception) ou polysémiques (lexèmes en ayant plusieurs).

Tab. 1 — *Distribution théorique des lexèmes selon la zone d'utilisation et le type de sens.*

	Monosèmes	Polysèmes	
		ayant une seule et même acception principale	ayant au moins deux acceptions principales
Connus partout	<i>erle</i> <i>idi</i> <i>urte</i>	<i>eman</i> <i>erdi</i> <i>joan</i>	<i>anaia</i> <i>heldu</i> <i>arotz</i> <i>hondar</i> <i>behatz</i> <i>hodei</i> <i>eltxo</i> <i>izter</i> <i>gaixo</i> <i>laino</i> <i>gaur</i> <i>lepo</i> <i>hagin</i>
Connus en certains lieux	<i>ekhi</i> <i>izeko</i> <i>neba</i>	<i>argizagi</i> <i>etorri</i>	<i>galtza</i> <i>oihan</i> <i>zuhain</i>

⁷ Mais les réponses manifestement erronées n'ont pas été incorporées dans les comptages et on peut considérer qu'une réponse représentant moins d'un pour cent de locuteurs n'a pas été retenue.

anaia : « frère d'un homme » (A, B) et « frère » (G, N, L, BN, S).
argizagi : « lune » (S) et « astre » (en certains endroits).
arotz : « charpentier » (A, B, G) et « forgeron » (N, L, BN, S).
behatz : « ongle » (L, BN), « doigt » (G, N), « pouce » (S), ailleurs : « doigt de pied ». Lexème comportant au moins quatre acceptions principales distinctes (« ongle », « doigt », « pouce » et « doigt de pied »).
ekhi : « soleil » (S).
eltxo : « moustique » (A, B, G, N, S) et « moucheron » (S) (en S, « moustique » = *mus/ztika*). *Eltxo* n'est pas signalé en L (où l'on a *mus/ztika* et *sistuli*) et BN (où seul *mus/ztika* est signalé). *Eltxo* « moustique » est présent dans les deux aires latérales (A/B et S).
eman : « donner ».
erdi : « moitié » (partout) et « accoucher » (N).
erle : « abeille ».
etorri : « venir » (A, B, G, N).
gaixo : « malade » (A, B, G, N) et « pauvre, malheureux » (N, L, BN, S) (dans cette aire, « malade » = *eri*).
galtza : « pantalon » (G, N, L, BN) et « chaussette » (S) (en S, « pantalon » = *zaragoila*).
gaur : « aujourd'hui » (A, B, G, N) et « ce soir » (N, BN, L, S) (dans cette aire, « aujourd'hui » = *egun*).
hagin : « dent » (A, B) et « molaire » (G, N, L, BN, S).
heldu : « arriver » (A/B, S) et « venir » (N, L, BN).
hodei : « nuage » (présent dans tous les dialectes) et « tonnerre » (en quelques endroits : A, B, G, N).
hondar : « sable » et « reste » (partout). En S, sable se dit *harina*, *hondar* étant monosémique avec le sens de « reste ».
idi : « bœuf ».
izeko : « tante » (A, B).
izter : « cuisse » (dans tous les dialectes sauf dans quelques endroits, notamment en S), « jambe » (A, B, N) et « quartier de fruit » (en S, « cuisse » = *azpi*).
joan : « aller ».
laino : « brume, brouillard » (S) et « nuage » (Pays Basque Sud). (Voir aussi *behe-laino* « brouillard, brume » en Biscaye.)
lepo : « dos » (A, B) et « cou » (G, N, L, BN, S) (en A, B, « cou » = *sama*).
neba : « frère d'une femme » (A, B).
oihan : « forêt, bois » (L, BN, S), « brousse, lande » et « montagne » (N). (*Oihan* est-il connu partout auquel cas ce terme devrait se trouver dans la ligne *Connus partout* ?)
urte : « année ».
zuhain : « fourrage » (L, BN) et « arbre » (S).

Le tableau suivant propose une répartition estimative des 1335 lexèmes-cibles recueillis par EAEL selon la zone d'utilisation (connus partout/connus en certains lieux) et le type de sens (monosèmes/polysèmes).

Tab. 2 — Distribution statistique des lexèmes selon la zone d'utilisation et le type de sens.

	Monosèmes	Polysèmes		Total
		ayant une seule et même acception principale	ayant au moins deux acceptions principales	
Connus partout	3 %	16 %	1 %	20 %
Connus en certains lieux	12 %	64 %	4 %	80 %
Total	15 %	80 %	5 %	100 %

Méthode de comptage :

Connus partout : total des lexèmes-cibles signalés dans tous les dialectes en réponse à un seul et même lexème-source (*erle* « abeille » est pandialectal).

Connus en certains lieux : total des lexèmes-cibles signalés en concurrence avec au moins un second géosynonyme en réponse à un même lexème-source (« chauve-souris » se dit *saguzar* en certains endroits et *gauenara* dans d'autres).

Monosèmes : je retiens un taux de monosémie de 15 %.

Polysèmes ayant une seule et même acception principale : total des lexèmes-cibles qui ne sont considérés ni comme des monosèmes ni comme des polysèmes ayant au moins deux acceptions principales.

Polysèmes ayant au moins deux acceptions principales : je considère à partir d'un sondage estimatif que 5 % des lexèmes-cibles connus partout sont des polysèmes ayant au moins deux acceptions principales. La même proportion est appliquée aux polysèmes connus en certains lieux.

122/ Géosynonymes ayant/n'ayant pas de morphème en commun

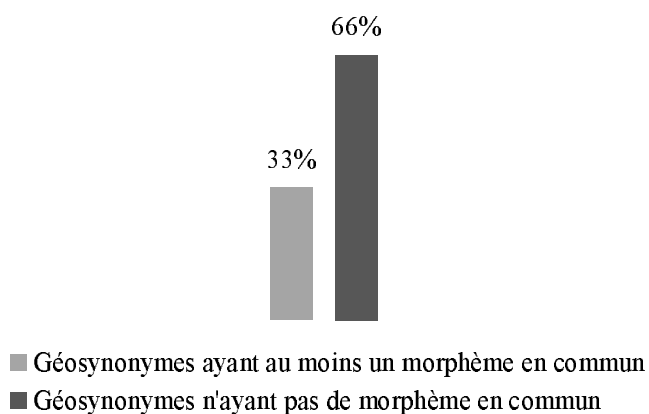
Les géosynonymes (lexèmes-cibles distincts selon les régions mais ayant le même sens) peuvent se répartir en deux catégories :

- les géosynonymes ayant au moins un morphème en commun (*behatz* et *atzamar* « doigt ») ;
- les géosynonymes n'ayant aucun morphème en commun (*eri* et *gaixo* « malade »).

Tab. 3 — Géosynonymes ayant/n'ayant pas de morphème en commun : aperçu.

Au moins un morphème en commun	Aucun morphème en commun
<i>behatz</i> / <i>atzamar</i> « doigt » <i>eguzki</i> / <i>ekhi</i> « soleil » <i>emakume</i> / <i>emazteki</i> « femme » <i>igeri egin</i> / <i>igerikan ari</i> <i>ilargi</i> / <i>argizagi</i> « lune » <i>ilun</i> / <i>ilunpe</i> « ombre » <i>izeko</i> / <i>izeba</i> « tante » <i>jaungoiko</i> / <i>jinko</i> « Dieu » <i>loba</i> / <i>alabaso</i> « petite-fille » <i>zil</i> / <i>zilko</i> « nombril »	<i>ardi</i> / <i>kukuso</i> « puce » <i>arantza</i> / <i>elorri</i> « épine » <i>bizkar</i> / <i>lepo</i> « dos » <i>dena</i> / <i>oro</i> « tout » <i>eri</i> / <i>gaixo</i> « malade » <i>etorri</i> / <i>jin</i> « venir » <i>jaio</i> / <i>sortu</i> « naître » <i>labur</i> / <i>motz</i> « court » <i>nigar</i> / <i>malko</i> « larme » <i>ortzi</i> / <i>dünda</i> « tonnerre »

Fig. 3 — *Parts respectives des deux types de géosynonymes (sondage à partir d'EAEL).*



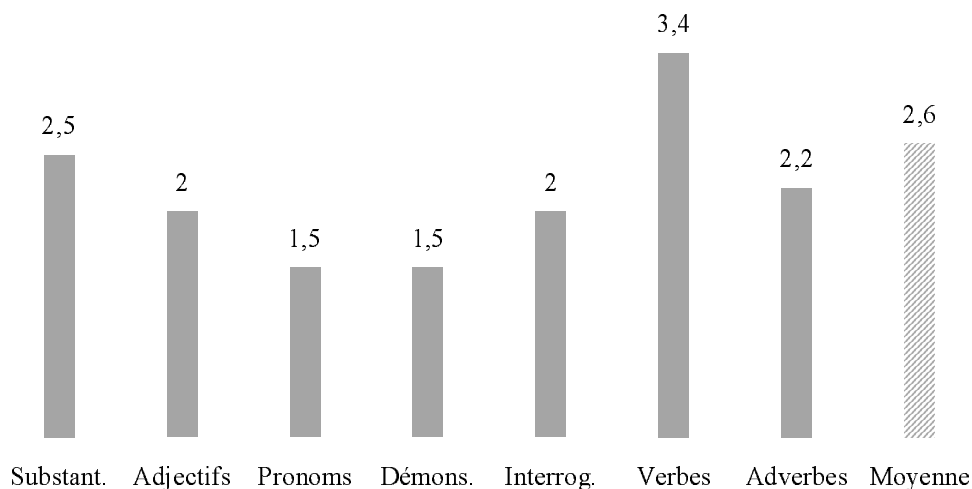
13/ Dispersion dialectale et catégories grammaticales

Il est possible de déterminer un ratio d'homogénéité/dispersion par catégorie grammaticale à partir d'un décompte des lexèmes sources et lexèmes cibles selon leur nature grammaticale.

Tab. 4 — *Ratios de dispersion dialectale selon la catégorie grammaticale (EAEL).*

	Lexèmes sources	Lexèmes cibles	Ratio
Substantifs	321	792	2,5
Adjectifs	69	139	2,0
Pronoms	6	9	1,5
Démonstratifs	2	3	1,5
Interrogatifs	1	2	2,0
Verbes	99	342	3,4
Adverbes	22	49	2,2
Total	520	1335	2,6

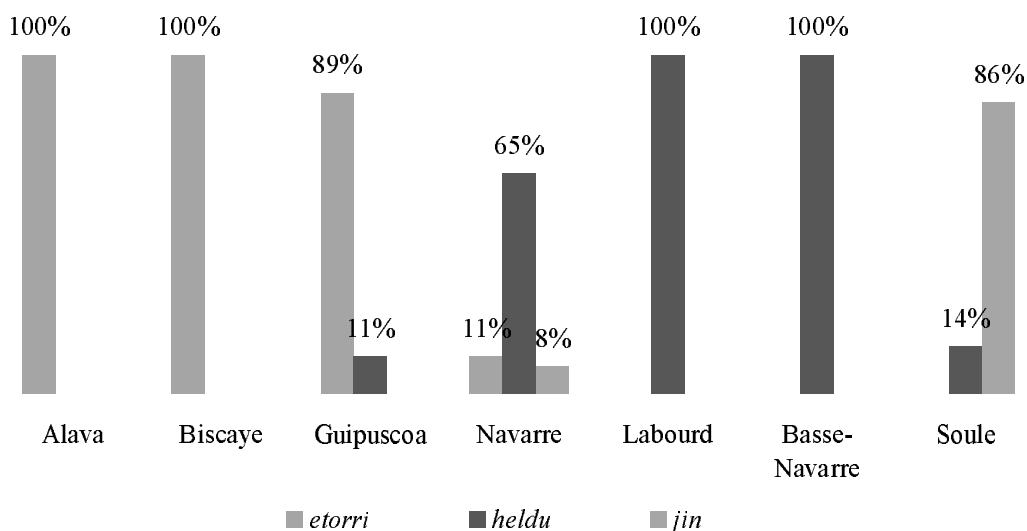
Fig. 4 — Ratios de dispersion dialectale selon la catégorie grammaticale (EAEL).



La catégorie grammaticale du verbe se singularise par le niveau élevé de son ratio de dispersion ; viennent ensuite les substantifs, les adverbes, les adjectifs et les interrogatifs qui se situent dans la moyenne ; alors que les pronoms et les démonstratifs sont plus homogènes. La catégorie du verbe est donc celle qui participe le plus à la dispersion dialectale. Ce fait observable en synchronie doit être explicable en diachronie.

Observons l'exemple suivant :

Fig. 5 — « il vient/viene » (EAEL, question n° 26).



— *etorri* est utilisé à l'ouest et disparaît progressivement en allant vers l'est (ce terme est généralement inconnu au Pays Basque Nord de ceux qui n'ont pas de pratique interdialectale) ;

— *heldu* est utilisé en Labourd et Basse-Navarre et, dans une moindre mesure, en Navarre, Soule⁸ et Guipuscoa ;

— *jin* est utilisé en Soule et quelque peu en Navarre (en Pays Basque Sud, ce terme est inconnu de ceux qui n'ont pas de pratique interdialectale, sauf dans quelques zones limitrophes du nord).

Si *etorri* est connu à l'ouest et inconnu à l'est, *jin* est connu à l'est et inconnu à l'ouest. *Heldu*, en revanche, est connu partout mais avec deux acceptions principales distinctes quoique proches (« venir » et « arriver »).

Une observation d'ordre morphosyntaxique doit ici être faite. En Alava, Biscaye et Guipuscoa, presque tous les informateurs ont répondu « *(ba)dator* », utilisant la forme synthétique du verbe *etorri*. En Navarre et Pays Basque Nord, en revanche, on trouve des formes périphrastiques telles que « *heldu da* » ou « *jiten da* ». Apparaît ici une dichotomie bien connue entre Pays Basque Sud et Pays Basque Nord quant à l'usage des formes verbales. Même si la perte de terrain des formes synthétiques au profit des formes périphrastiques est, à date historique, un phénomène général à tout l'espace bascophone, on sait que les premières ont mieux résisté dans les dialectes péninsulaires. À la différence lexicale (*etorri/heldu/jin*) se surajoute donc une différence morphosyntaxique (forme synthétique/forme périphrastique).

On peut ici faire l'hypothèse que la dispersion dialectale des formes verbales a été accentuée par la combinaison de plusieurs événements qui ont sélectivement affecté la morphologie verbale à date historique. Le principal de ces événements étant la perte de terrain des formes synthétiques au profit des formes périphrastiques, événement lié à la diffusion d'une conjugaison composée à l'instar de ce qui s'est passé dans les langues romanes.

«Así, por ejemplo, las formas fuertes, sintéticas del verbo no han cesado de perder terreno ante las formaciones perifrásticas; el pasado se ha convertido en un 'imperfecto' perdiendo el valor aorístico que en otros tiempos podía tener (vize. ant. *neuca* 'tenía lo (yo)', pero *nentorre* 'vine', *engarren* 'te traxo', etc.) De una manera general, no parece aventurado pensar que la formación de tiempos compuestos es en vasco un fenómeno paralelo al que se observa en tantas lenguas de Europa occidental, pero no solo en esa área, y por tanto relativamente reciente, aunque prehistórico, en el sentido lingüístico de la palabra⁹.»

Cela a conduit à formuler l'hypothèse que le système verbal protobasque n'aurait connu que des formes synthétiques¹⁰. Deux autres phénomènes ont sans doute joué un certain rôle dans

⁸ Dans les communes proches du domaine bas-navarrais.

⁹ Michelena, 1964 [1988 : 5].

Les remarques suivantes, qui concernent la bascule du passé simple au passé composé français, pourraient aussi s'appliquer au basque historique :

« Avec elle [la disparition du passé simple en français], commence pour les verbes une nouvelle ère, c'est un cheminement vers l'état du verbe où il n'y aura plus comme voiles que des auxiliaires faisant manœuvrer une coque qui porte l'idée. » M.J. Gilliéron, « La faillite de l'étymologie phonétique », *Neuveville*, 1919.

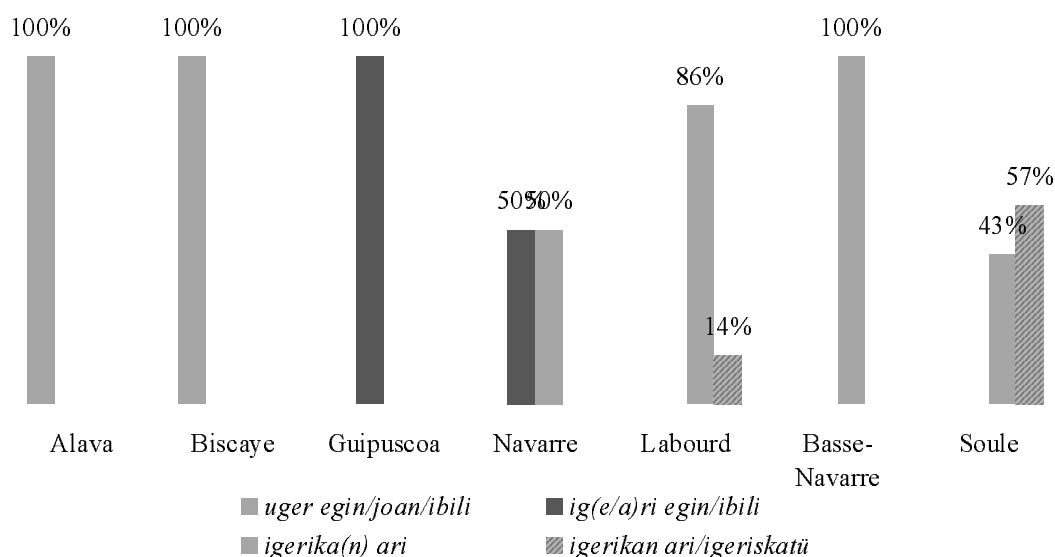
« Située dans l'ensemble dont elle fait partie, l'élimination du prétérit simple est un moment de grand développement qui entraîne les langues indo-européennes à passer du mot forme variable au mot fixé une fois pour toute. *J'ai aimé, tu as aimé*, etc. sont plus près du terme de cette évolution que ne l'étaient *j'aimai, tu aimas*, etc. » A. Meillet, 1982 : 158.

¹⁰ Cf. entres autres Azkarate et Altuna, 2001 : 152.

la dispersion des formes verbales dialectales : l'introduction des formes respectueuses de vouvoiement au Moyen Âge et la régression de l'usage des formes d'adresse de familiarité.

Prenons un exemple de verbe aujourd'hui exclusivement périphrastique et observons les modalités de la dispersion dialectale.

Fig. 6 — « *il nage/il a nagé – nada/ha nadado* » (EAEL, question n° 315).



On constate que la même racine est omniprésente mais que la morphologie du syntagme verbal varie sensiblement¹¹. Dans un cas, le verbe inergatif est exprimé par une structure de type ergatif [NOM + *egin*] : *uger/igeri egin* « nager » (littéralement « faire nage »)¹². Dans un autre, on a une structure intransitive constituée sur le type [NOM + *-ka* (affixe réitératif) + *-n* « inessif » (non systématique) + *ari* (particule signalant l'action en train de se dérouler)] : *igerika(n) ari* « nager » (littéralement « être en train [dans] la manière réitérative de la nage »). La troisième forme (Soule) comporte l'affixe réitératif *-ka* et le suffixe aspectuel *-tu*.

Un trait général à peu près constant se dégage à travers les parlers : malgré l'existence d'une forme commune, *igeri*, rendant l'idée de nage, il n'existe pratiquement pas de verbe dérivé, et on utilise soit des verbes tel qu'*egin* ou *ari izan*, soit des verbes de mouvement comme *ibili* ou *joan*. La structure non dérivée représente 98 % des citations (estimation globale pondérée), la structure dérivée 2 %¹³. Il semble que tous les verbes de déplacement obéissent à cette règle morphosyntaxique (*lasterka ibili*, *arineketan egin/joan*, *korrika egin*, etc.).

¹¹ On trouve, notamment en Navarre, une grande dispersion : *iregin*, *igerikan ibili*, *igeri*, etc.

¹² B. Oyharçabal, in Orpustan, 1996 : 119.

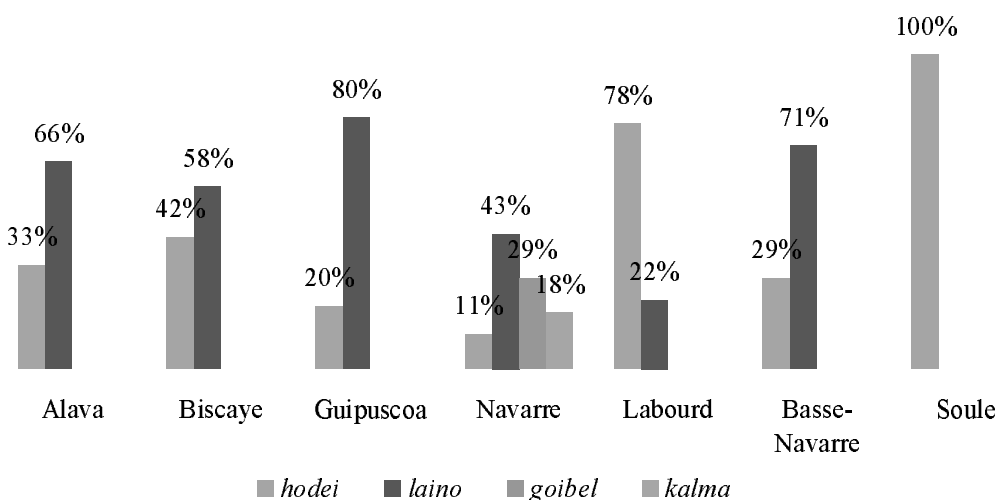
¹³ De fait, mes informateurs de Soule me précisent qu'*igeriskan ari da* (au présent) et *igeriskan erauntsi dü* (au passé) sont préférés par eux, sinon seuls possibles, dans ce contexte (on remarquera la fracture d'actance entre présent et passé, phénomène assez fréquent en basque, mais aussi dans d'autres langues, avec les verbes inergatifs).

2/ COMMENTAIRE

21/ Polysèmes ayant au moins deux acceptions principales

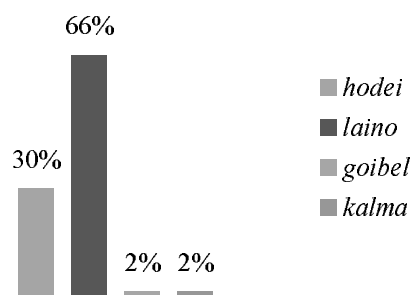
« *Nuage* ». L'équivalent de « nuage » dans *EAEL* est *laino* pour deux bascophones sur trois et *hodei* pour un bascophone sur trois. Les deux termes sont signalés dans des proportions variables dans tous les dialectes à l'exception de la Soule où seul *hodei* est cité (*laino* y signifie « brume », « brouillard »). Comme le suggère la figure suivante, le Guipuscoa et la Navarre sont le foyer où *hodei* perd le plus de terrain (*hodei* n'y est signalé que par respectivement 20 et 11 % des informateurs), l'influence s'étendant au bas-navarrais (29 % des informateurs). *Hodei* résiste mieux dans les aires latérales : 42 % des informateurs en Biscaye (*edoi* par métathèse), 78 % en Labourd, 100 % en Soule. À noter que parmi les informateurs ayant cité *laino* avec le sens de « nuage », sept d'entre eux ont cité *hodei* avec le sens de « tonnerre » : *hodei* rentre donc dans la catégorie des polysèmes ayant au moins deux acceptions principales. Sur la base de ces observations, on est tenté de proposer que le sémantisme profond d'*hodei* renvoie au nuage de pluie ou d'orage (types *nimbostratus* ou *cumulonimbus*) alors que celui de *laino* vise le nuage bas non pluvieux (type *stratus*).

Fig. 7 — Appellation du « nuage » au Pays Basque (*EAEL*, question n° 228)¹⁴.



¹⁴ La répartition des lexèmes-cibles se fait sur une base de 100 % à l'intérieur de chaque dialecte. L'échantillon d'*EAEL* n'étant pas représentatif du poids relatif des dialectes par rapport au total des locuteurs, un coefficient correcteur permet d'établir une estimation globale de la fréquence de chaque lexème-cible dans la langue.

Fig. 7 bis — Le « nuage » en basque : estimation globale pondérée.



On a la confirmation que les polysèmes ayant au moins deux acceptions principales distinctes sont révélateurs d'un mouvement de la langue. On peut à ce stade tenter de tracer une première allure de ce mouvement :

— une *extension sémantique* provoque un *abandon lexical* (*laino* prend le sens de « nuage » et fait « disparaître » *hodei*¹⁵) ;

— une *extension sémantique* provoque un *déplacement lexical* (face à *laino* qui étend son périmètre sémantique, *hodei* cède le sens de « nuage » et prend celui de « tonnerre »¹⁶).

Si l'on isole une période de temps donnée de la langue, on observera qu'une partie du lexique est toujours en mouvement. Des mots changent de sens, des mots nouveaux sont créés (par composition, dérivation, emprunt), des mots sont « poussés » à prendre d'autres acceptions, des mots sont abandonnés. La polysémie joue dans ces processus un rôle capital car hormis les changements qui s'opèrent par création ou emprunt, c'est dans le stock de mots existants que le changement va puiser ses ressources, accentuant de ce fait le phénomène polysémique (plus un mot est polysémique, plus ce mot a des chances d'être ancien). Dans le même temps, une autre partie du lexique de la langue n'est pas affectée par le phénomène du changement. La question cruciale qui se pose alors est de savoir pourquoi, au cours d'une période donnée, le changement affecte certains lexèmes et en épargne d'autres.

« L'apparition de traits ou de rapports culturels d'un type nouveau rend souvent nécessaire un enrichissement des ressources de la langue, mais cet enrichissement n'est jamais un processus d'addition arbitraire par lequel quelque chose de nouveau viendrait s'adjoindre aux éléments de contenu ou de forme déjà existants. Il s'agit seulement d'une extension dans l'application de principes déjà utilisés ou même, bien souvent, d'un simple élargissement métaphorique de significations ou de termes anciens¹⁷. »

¹⁵ En certains endroits.

¹⁶ En certains endroits.

¹⁷ Sapir, 1968 : 33.

22/ « Faux amis » : un phénomène linguistique universel

Les polysèmes que nous venons d'étudier sont communément qualifiés de *faux amis* (*false cognates*). Il s'agit de mots qui se correspondent par l'étymologie et par la forme, d'un dialecte à l'autre ou d'une langue à l'autre, mais qui ont évolué au sein d'environnements distincts et ont de ce fait pris des sens différents. Les polysèmes ayant au moins deux acceptions principales distinctes (voir tableau 1) ne sont pas une originalité de la langue basque mais un phénomène linguistique général qui s'observe, à des degrés divers, partout.

On trouve ainsi des faux amis entre le français de France et le français de Belgique : *déjeuner* veut dire « prendre le repas de midi » en France mais « prendre le petit déjeuner » en Belgique ; *dîner* veut dire « prendre le repas du soir » en France mais « prendre le repas de midi » en Belgique ; *sonner* veut dire « faire retentir une sonnerie » en France mais « téléphoner » en Belgique ; *cru* veut dire « non cuit » en France mais « froid et humide » en Belgique.

On trouve aussi de nombreux faux amis entre l'espagnol d'Europe et l'espagnol d'Amérique : *baño* veut dire « bain » en Espagne mais « toilettes » au Mexique ; *camión* veut dire « camion » en Espagne mais « autobus » au Mexique ; *cena* veut dire « dîner » en Espagne mais « repas de midi » en Colombie ; *clavo* veut dire « argent » au Texas mais « ennui » au Mexique et « farce » en Argentine ; *comida* « le repas de midi » en Espagne mais « dîner » en Colombie ; *cuadra* « écurie » en Espagne mais « pâté de maison » en Amérique latine ; *llanta* « jante » à Caracas mais « pneu » à Bogota ; *coger* « prendre » en Espagne mais a une connotation obscène en Amérique latine¹⁸, etc.

La moitié environ du lexique anglais est d'origine française : il existe de ce fait de nombreux faux amis entre ces deux langues dont je cite ici quelques exemples pour leur homologie avec le basque.

Tab. 5 — « Faux amis » : exemples.

basque dialecte 1	basque dialecte 2	anglais	français
galtza		vest	
« pantalon »	« chaussette »	« tricot de corps »	« veste »
heldu		achieve	
« arriver »	« venir »	« accomplir, réussir »	« achever, terminer »
hagin		grape	
« dent »	« molaire »	« grain de raisin »	« grappe de raisin »
gaixo		sensible	
« malade »	« infortuné, pauvre »	« raisonnable »	« sensible »

¹⁸ B. Pottier in *Encyclopædia Universalis*, article « Espagne » et relevés personnels.

Si les faux amis sont une réalité linguistique courante, ils signalent néanmoins un découplage forme/sens et témoignent donc d'une tendance à la divergence. Dans ce contexte, si la distance géographique est, entre autres, l'un des facteurs favorables à leur apparition, on remarquera que le phénomène s'observe en basque sur une bande territoriale réduite (180 kilomètres d'est en ouest, une cinquantaine de kilomètres du nord au sud).

23/ Une même réalité observée sous deux angles

Le présent travail est exploratoire et n'a pour ambition que de signaler des tendances. Il est néanmoins important de bien préciser une question de méthode. Tout comptage lexical peut se faire à partir de deux points de vue distincts mais complémentaires.

1/ Dans le commentaire de la figure 2, il est proposé qu'en se basant sur l'enquête d'*EAEL* on pouvait conclure que 41 % des « concepts de base » s'exprimaient en basque au moyen d'un vocabulaire commun à tous les locuteurs.

2/ Le tableau 2 indique quant à lui que 20 % des lexèmes de la langue de base sont connus sur tout l'espace bascophone.

Ces proportions tentent de cerner, d'une part, ce qui dans la langue est commun et échappe à la variation dialectale et, d'autre part, ce qui diverge. Mais les angles d'observation adoptés sont différents. Dans le premier cas le point de départ à partir duquel le chiffrage est effectué est les *lexèmes-sources testés par EAEL*, c'est-à-dire, si l'on veut, les « concepts de base » de la langue (onomasiologie). Dans le second cas, le point de départ est les *lexèmes-cibles recueillis par EAEL*, tous dialectes confondus, c'est-à-dire le lexique de la langue (sémasiologie).

Pour saisir cette double perspective, il importe de dissocier les notions de *vocabulaire* et de *lexique*¹⁹.

— le vocabulaire est l'ensemble des mots connus et susceptibles d'être utilisés par un locuteur donné ;

— le lexique est l'ensemble des mots que compte une langue (on parle aussi de dictionnaire de la langue).

Le total du vocabulaire est forcément inférieur à celui du lexique. En français, des études par sondages de pages de dictionnaire montrent que des personnes normalement informées disposent d'un vocabulaire de production de 30 000 à 40 000 mots et d'un vocabulaire de réception atteignant 60 000 mots. À titre indicatif, le *Grand Robert de la langue française* donne 80 000 entrées lexicales²⁰. S. Pinker donne des chiffres comparables pour l'anglo-américain : l'Américain moyen diplômé du secondaire connaîtrait 45 000 mots (1999 : 148). Il en va de même en basque : tout locuteur connaîtra forcément moins de mots que la langue n'en compte, ce d'autant plus que le fait dialectal accroît très fortement le nombre de mots.

Ceci nous amène à un double constat :

¹⁹ Picoche, 1992 [1^{re} éd. 1977] : 45-46.

²⁰ B. de Boysson-Bardies, 2003, *Le langage, qu'est-ce que c'est ?*, Paris, Odile Jacob, p. 91.

1/ Optique vocabulaire : Tout bascophone aura, en situation d'interlocution, un nombre variable de lexèmes en commun avec son interlocuteur. Dans l'hypothèse d'une conversation de base, et si l'on admet que les lexèmes-sources testés par *EDEL* représentent ce vocabulaire de base, la proportion du vocabulaire en commun sera au maximum de 100 % (locuteurs appartenant à un même dialecte ou sous-dialecte) et diminuera progressivement à mesure que croîtra l'« éloignement » dialectal pour atteindre une proportion plancher théorique d'environ 50 % de lexèmes en commun (locuteurs appartenant à des dialectes « éloignés » et n'ayant en commun aucuns géosynonymes)²¹.

2/ Optique lexicale : Tout bascophone :

- connaîtra les lexèmes-cibles connus partout (environ 20 % du lexique) ;
- connaîtra, dans la catégorie des lexèmes-cibles connus en certains lieux (géosynonymes), ceux qui appartiennent à son propre dialecte (environ 20 % du lexique) ;
- ne connaîtra pas (sauf s'il a une pratique ou des connaissances interdialectales) les géosynonymes n'appartenant pas à son dialecte²² (environ 60 % du lexique).

En effet, si les locuteurs d'un même dialecte ou sous-dialecte disposent d'une même forme pour désigner l'*arc-en-ciel*, aucun bascophone ne connaîtra spontanément toutes les autres formes par lesquelles celui-ci se signale dans les autres dialectes : *ortzadar*, *Jainkoan paxa*, *Erromako zubi*, *San Migelen zubi*, *arkoiris*, *ustruke* (*EDEL*⁵⁴⁸).

L'optique vocabulaire est certainement plus pertinente que l'optique lexicale quant à l'appréciation du degré d'intercompréhension. Lorsque deux locuteurs se rencontrent, ce sont en effet deux vocabulaires, non deux lexiques, qui se rencontrent.

Ces données fournissent en première approche une base statistique à la question du degré variable d'intercompréhension entre bascophones²³. Tous les degrés intermédiaires peuvent en effet se rencontrer entre les deux bornes que sont l'intercompréhension totale et l'absence d'intercompréhension. Même s'il ne faut pas réduire cette question à une problématique exclusivement lexicale (la phonologie et la syntaxe jouent bien sûr un rôle tout aussi capital), J. Picoche pose clairement les enjeux lexicaux de l'intercompréhension :

« La compréhension entre un locuteur X et un locuteur Y, qu'il s'agisse d'un véritable dialogue ou du décodage par X d'un monologue oral ou écrit de Y, exige donc que les mots — forcément actifs — utilisés par Y fassent partie du vocabulaire au moins passif de X, et, en cas de dialogue, vice versa, l'intercompréhension sera d'autant meilleure que les mots utilisés par X et Y seront des mots actifs pour chacun d'eux ; d'autant plus médiocre que les mots actifs de l'un seront des mots passifs pour l'autre ; d'autant plus mauvaise qu'un plus grand nombre des mots actifs de l'un seront des mots inconnus de l'autre²⁴. »

²¹ Sur la base du protocole de comptage précité. Ce vocabulaire commun ne sera pas forcément toujours intercompréhensible (cf. note 4). Si l'on voulait répondre à la question : « Combien de concepts fondamentaux s'expriment-ils en basque par un lexème comportant au moins une même base morphématique ? », on pourrait répondre sur la base d'un sondage : « environ 66 % ».

²² J'ai validé la synthèse de l'enquête d'*EDEL* avec une informatrice de Soule appartenant à la même génération que les informateurs d'*EDEL*. Elle connaissait bien tous les termes pandialectaux et ignorait et/ou a déclaré ne pas utiliser la plupart des géosynonymes autres que ceux appartenant à son dialecte.

²³ Voir introduction à *EHHA*.

²⁴ Picoche, 1992 [1^{re} éd. 1977] : 47.

24/ Convergence ou divergence ?

Tout travail de reconstruction repose sur le postulat que l'hétérogénéité actuelle fait suite à une homogénéité passée. Comparer des unités lexicales distinctes en cherchant à leur assigner une forme ancienne uniforme ne peut se faire que si la langue a évolué dans le sens de la divergence, non de la convergence. Le constat de la diversité dialectale contemporaine (qui ne saurait occulter le caractère unique de la langue) conduit donc à postuler l'existence d'une forme de langue plus homogène dans le passé :

« Nous pourrions nous aventurer à dire qu'à une époque pas très lointaine de la préhistoire de la langue, a existé parmi nous une koinè assez unitaire, dont les dialectes basques historiques sont sortis. Dire qu'il y a eu koinè est une autre manière de dire qu'il a dû y avoir une espèce d'unité supra-régionale qui contribua à étendre un type de langue en effaçant de nombreux particularismes antérieurs²⁵. »

Le pluralisme dialectal actuel résulterait donc d'évolutions divergentes à partir d'une forme passée assez unitaire, celle-ci étant elle-même la forme réunifiée de formes antérieures diverses. On aurait donc eu selon ce scénario trois séquences : hétérogénéité contemporaine (étape 3) < forme réunifiée « assez unitaire » (étape 2) < hétérogénéité initiale (étape 1). Si un tel scénario est plausible, il ne peut s'expliquer que par des causes socio-historiques ; chacune des trois étapes trouvant son origine dans des conditions socio-historiques centrifuges (étapes 3 et 1) ou centripètes (étape 2). Un tel phénomène s'est par exemple produit en grec où on sait que plusieurs séquences socio-historiques se sont succédé, faisant alterner des périodes d'uniformisation et des périodes de diversification des usages linguistiques²⁶.

N'acceptant pas sans réserves que le biscayen soit si différent des autres dialectes, L. Michelena explique l'originalité actuelle de ce dialecte par une causalité de type socio-historique : « un mouvement d'expansion, dans une direction ou dans une autre, a mis en contact des variétés préalablement distantes l'une de l'autre, effaçant ou recouvrant des parlers de transition²⁷ ».

²⁵ Michelena, in Haritschelhar, 1983 : 260.

²⁶ *1^{re} période (diversité initiale des dialectes)*. – 2000 à – 350 ans av. J.-C. La Grèce connaît une multiplicité de dialectes (dont le dorien, l'arcado-chypriote, l'éolien, l'ionien-attique).

2^e période (convergence). – 350 ans av. J.-C. jusqu'au 6^e siècle ap. J.-C. Cette seconde période commence avec la domination d'Alexandre le Grand. Une langue commune (*koinè*) fondée sur le dialecte d'Athènes s'élabore lentement.

3^e période (divergence). La domination de Byzance met Athènes sous l'éteignoir ; le centre du pouvoir se déplace vers Byzance, les dialectes se fragmentent à nouveau. Se succèdent alors plusieurs périodes d'invasions : slave (6^e siècle), arabe (10^e), normande (12^e), vénitienne (15^e au 18^e) puis turque (de quatre à six siècles selon les régions).

4^e période (convergence). Débute avec l'indépendance de la Grèce, en 1830.

N. G. Contossopoulos, 1978, « Los dialectos en la Grecia contemporánea », in *Bulletin lexicographique*, Athènes.

²⁷ Michelena, in Haritschelhar, 1983 : 260.

En montrant qu'elles se complètent, il me semble intéressant de mettre en parallèle des observations de L. Michelena (qui se situent entre autres au plan phonologique), celles faites par A. Meillet (qui se situent au plan syntaxique) dans un article relatif aux dialectes indo-européens et intitulé « Convergence des développements linguistiques²⁸ ». Dans cet article, A. Meillet montre qu'entre l'indo-européen et des langues qui en découlent comme le français, l'anglais, le persan, il y a des différences de détail infinies. On ne reconnaît pas au premier abord l'unité d'origine dans ces formes indo-européennes contemporaines. Mais ces langues manifestent cependant des concordances frappantes dans leur structure générale. Le latin ancien conservait exactement un état de choses indo-européen quand il disait *domus patris* (ou *patris domus*). Ce syntagme se traduit par *la maison du père* (français), *the house of the father* (anglais), *mân i pidar* (persan). Ces trois syntagmes se distinguent du syntagme indo-européen par deux traits essentiels : d'une part, l'ordre des mots y est fixe (alors qu'il était libre en latin ancien), indiquant ainsi que le second nom est le complément du premier ; de l'autre, la forme du complément est la même que celle que l'on utiliserait en position de sujet (alors qu'on a deux formes distinctes en latin : *pater* sujet et *patris* complément). Le français, l'anglais et le persan ont donc adopté un même changement de structure par rapport à l'indo-européen dont ils découlent.

« Si les mêmes conditions se rencontrent, les langues se développent non seulement en un même sens, mais aussi de la même manière dans le détail matériel des formes. Les mêmes innovations se produisent donc indépendamment chez des individus différents pourvu qu'ils soient placés dans les mêmes conditions. Ceci ne prouve naturellement pas que les changements naissent spontanément chez chacun des sujets, et qu'il n'y ait pas, en une plus ou moins large mesure, imitation d'un sujet par un autre ; mais il en résulte au moins qu'ils sont susceptibles de naître indépendamment chez plusieurs sujets, et souvent même chez beaucoup de sujets²⁹. »

Toute langue est donc simultanément et constamment traversée par des mouvements contraires de divergence et de convergence. Lorsque à certaines époques de la langue le balancier penche davantage dans un sens plutôt que dans l'autre, c'est pour des raisons socio-historiques doublées de phénomènes d'évolution interne de la langue. La langue basque ne doit pas échapper à cette règle. Ainsi, à l'époque où le suffixe aspectuel *-tu* s'est répandu³⁰, contribuant ainsi à la diffusion à tous les dialectes des formes verbales périphrastiques, on a assisté à un mouvement général de la langue (les verbes périphrastiques sont désormais largement majoritaires dans tous les dialectes). Cependant, ce changement convergent s'est réalisé à un rythme variable selon les parlers. On sait en effet que la perte de terrain des

²⁸ Meillet, 1982 : 61-75.

²⁹ Meillet, *ibid* : 73.

³⁰ Les spécialistes sont d'accord pour dire que ce suffixe a été emprunté au latin. Tous ont souligné que ce suffixe aspectuel était désormais le plus productif dans la formation des verbes dans la langue. Pour certains auteurs (Azkue, Lafon), des verbes autochtones anciens tels que *hartu* « prendre », *galdu* « (se) perdre », *kendu* « enlever », *sartu* « entrer » auraient même abandonné la forme ancienne du participe en *-i* pour adopter la nouvelle forme en *-tu* (*sartu* < *sarri*). Ceci paraît corroboré par l'existence dans la langue contemporaine de doublets tels qu'*itzali/itzaldu* « couvrir de son ombre, éteindre », *ahantzi/ahaztu* « oublier » et même *izan/izatu/izandu* « être » (cette dernière forme étant attestée par *EDEL*) (Azkarate & Altuna, 2001 : 190, 195).

formes synthétiques au profit des formes périphrastiques a été moins marquée au Pays Basque Sud qu'au Pays Basque Nord. Le résultat est donc aujourd'hui un état contrasté de la langue de ce point de vue-là (du type de celui que l'on observe entre espagnol d'Europe et espagnol d'Amérique latine, celui-ci ayant conservé de manière beaucoup plus prononcée l'usage des formes verbales synthétiques, dans l'expression du passé, que celui-là).

Il semble qu'il y ait quelque chose à gagner pour mieux comprendre l'essor des formes verbales périphrastiques en basque, à replacer leur évolution dans un contexte géolinguistique plus large. Certaines évolutions parallèles de la morphosyntaxe des langues d'Europe occidentale doivent en effet être soulignées.

La forme périphrastique est documentée en basque dès le 10^e siècle (dans les fameuses Gloses de San Millán) et n'a cessé de se développer depuis. Les verbes intransitifs se conjuguent avec les auxiliaires intransitifs *izan*/**edin* et les verbes transitifs avec les auxiliaires transitifs **edun*/**ezan* (*egin* en Biscaye, au moins en fin de période).

L'apparition de la forme périphrastique dans les langues romanes est contemporaine de leur naissance, à la suite de leur séparation d'avec le latin vulgaire, durant la deuxième moitié du premier millénaire de l'ère moderne (il n'existait pas en latin de forme périphrastique, hormis dans les flexions passives ainsi que dans celles des verbes déponents au passé). En français, la plupart des verbes se conjuguent aujourd'hui avec l'auxiliaire transitif *avoir* (*j'ai marché, j'ai parlé, j'ai vu*) alors que l'auxiliaire intransitif *être* continue d'être utilisé, notamment avec des verbes intransitifs de mouvement (*je suis allé, je suis venu, je suis monté*). Un fait important doit ici être rappelé : l'auxiliaire *être* a eu, en français, à date historique, une place plus importante que celle qu'il a aujourd'hui. En effet, des verbes intransitifs tels que *broncher, couler, courir, croître, marcher, périr, voler*, qui se conjuguent actuellement avec l'auxiliaire *avoir*, se faisaient encore accompagner de l'auxiliaire *être* au 16^e siècle³¹.

Contrairement au français, l'espagnol a aujourd'hui généralisé l'utilisation de l'auxiliaire transitif *haber* « avoir » pour la construction des temps à la forme active (*he ido, he venido, he subido*). Mais, fait remarquable, l'auxiliaire *ser* « être » s'utilisait encore au Moyen Âge, à côté de *haber* (*soy entrado*, pour « je suis entré », contre *he entrado* aujourd'hui)³².

En vieil-anglais, les temps périphrastiques apparaissent à la même époque et suivent le même processus de diffusion de plus en plus grande³³. Si aujourd'hui l'auxiliaire de base servant à former les temps composés est *to have* « avoir » (*I have gone, I have come, I have gone up*), deux auxiliaires furent utilisés au départ : *bēon* (*be* « être ») avec les verbes intransitifs et *habban* (*have* « avoir ») avec les verbes transitifs. Mais « Dès le vieil-anglais *habban* tend à être employé aussi avec les verbes intransitifs » (Crépin, *ibid.*).

On observe donc des processus parallèles dans plusieurs langues d'Europe de l'Ouest :

— chronologie de la mise en place de la conjugaison composée (2^e moitié du premier millénaire de l'ère moderne) ;

— logique du choix initial des auxiliaires (*être* et *avoir*, *ser* et *estar*, *bēon* et *habban*, *izan*/**edin* et **edun*/**ezan*) ;

³¹ Tritten J.-L., 1999, *Histoire de la langue française*, Ellipses, Paris, p. 60.

³² B. Darbord, B. Pottier, 1988, *La langue espagnole. Éléments de grammaire historique*, Paris, Nathan, pp. 164-170.

³³ Crépin A., 1978, *Problèmes de grammaire historique*, Paris, PUF, pp. 128-130.

— évolution de l'usage des auxiliaires (décroissance, voire disparition de l'usage des auxiliaires intransitifs au profit des auxiliaires transitifs).

Même si ces évolutions ne sont pas identiques ni même rigoureusement parallèles (le basque a conservé l'auxiliaire intransitif, alors que le français ne le maintient que pour certains verbes et que l'espagnol et l'anglais l'ont abandonné³⁴), la question se pose de savoir quelles sont dans ces changements les parts respectives des influences réciproques (linguistique aréale), des évolutions internes (linguistique structurale), des invariants morphosyntaxiques (linguistique typologique). A. Martinet expliquait ces phénomènes de covariation par la notion d'« affinité » (le fait que les langues n'appartiennent pas à une même famille ne saurait, selon Martinet, faire obstacle à l'affinité) :

« Les explications linguistiques de l'affinité se rangent en deux catégories :

1/ l'influence réciproque de deux langues en contact, avec ou sans prédominance de l'une des deux ;
2/ l'influence exercée par une troisième langue, substrat, superstrat ou adstrat.

Dès que des causes, quelle que soit leur nature, sont envisagées, tout le problème de l'affinité linguistique vient prendre sa place dans le vaste domaine, si souvent négligé, de la convergence linguistique, qui représente, ou du moins devrait représenter, une moitié de la linguistique dynamique, l'autre moitié étant, bien entendu, la divergence, celle qui, pendant près d'un siècle a paru se confondre avec le domaine de la linguistique tout entière³⁵. »

Les contours complexes qu'offre la situation linguistique actuelle du Pays Basque, observables en synchronie, sont interprétables en diachronie. Comme maints autres espaces linguistiques, l'aire bascophone a connu à travers son histoire des vicissitudes qui ont vu alterner des mouvements de différenciation et d'unification, ces mouvements pouvant, selon les périodes et les contextes socio-historiques, s'équilibrer ou bien céder alternativement du terrain. Focaliser l'attention sur les seuls phénomènes d'hétérogénéité ou, à l'inverse, sur les seuls phénomènes d'homogénéité, méconnaîtrait une partie de ce qui participe d'une même problématique de type dialectique (près de la moitié des lexèmes-sources d'*EACL* ont généré une seule et même réponse de la part de tous les informateurs alors que l'autre moitié a généré entre deux et une dizaine de réponses distinctes).

Martinet nous permet de conclure en rappelant les mécanismes sociolinguistiques qui déterminent les processus de dialectalisation et leurs conséquences :

« Les considérations qui précèdent permettent de jalonner un processus qui a dû se répéter à des milliers d'exemplaires depuis qu'il est des hommes, et qui parlent : un groupe humain, agressif ou prolifique, étend son domaine au point que les contacts entre ses différentes tribus perdent de leur

³⁴ Le caractère fondamental de la bipolarité absolutif/ergatif dans la morphosyntaxe de la langue basque n'est certainement pas étranger au maintien des deux binômes *izan*/**edin* et **edun*/**ezan* (**eradun/egin*). On remarquera néanmoins que si une tendance se dégage en basque quant à l'évolution des auxiliaires, c'est également l'auxiliaire transitif qui tend à déborder sur les tournures intransitives plutôt que l'inverse. Comme on l'a vu plus haut, des verbes inergatifs sont assez souvent exprimés par une structure de type ergatif [NOM + *egin*] comme dans *igeri egin* « nager » (littéralement « faire nage »). Et des verbes intransitifs empruntés plus récemment tels qu'*eskiatu* « skier » ou *funtzionatu* « fonctionner » se conjuguent dans beaucoup de parlers avec l'auxiliaire transitif (*makina horrek ez du fontzionatzen* « cette machine ne fonctionne pas »).

³⁵ Martinet, 1975 : 26-27.

fréquence et de leur intimité. Ceci entraîne un processus de différenciation linguistique qui ira s'amplifiant si les contacts se relâchent encore entre les différentes tribus, et si de nouveaux contacts s'établissent avec les tribus d'autres groupes. Il y aura dialectisation de la langue initiale, et cette dialectisation pourra entraîner, d'un canton à un autre, une incompréhension totale. Mais une tribu plus agressive, plus prolifique, plus inventive ou plus cultivée que ses voisines pourra un jour leur imposer son hégémonie politique ou culturelle. Son dialecte deviendra la langue officielle ou littéraire aussi loin que s'étendra son hégémonie, et, à ce titre, commencera à déloger les dialectes locaux, soit, s'ils sont encore peu différenciés, par un procès de convergences poursuivi jusqu'à confusion complète, soit par pur et simple remplacement. Il n'est pas dit que les limites de cette hégémonie coïncideront avec celles de l'expansion initiale du groupe : sur certains points, elles les outrepasseront, et la nouvelle langue couvrira des régions où le parler local est de tout autre origine ; sur d'autres points, elles seront en retrait, et certains dialectes de notre groupe pourront se voir intégrés à une autre nation ou à une autre zone d'expansion culturelle dans laquelle ils finiront par disparaître³⁶. »

³⁶ Martinet, 1970 : 156-157.

ANNEXE 1

Données recueillies par *EAEL* : aperçu général*

Lexèmes-sources		Lexèmes-cibles connus partout							%
		A	B	G	N	L	BN	S	
Environ 50 % des lexèmes-sources	abeille	<i>erle</i>							Environ 20 % des lexèmes- cibles
	aller	<i>joan</i>							
	âne	<i>asto</i>							
	année	<i>urte</i>							
	apprendre	<i>ikasi</i>							
	aveugle	<i>itsu</i>							
	belle-fille	<i>errain</i>							
	bœuf	<i>idi</i>							
	blanc	<i>zuri</i>							
	boire	<i>edan</i>							
	bon	<i>on</i>							
	bras	<i>beso</i>							
	brebis	<i>ardi</i>							

Lexèmes-sources		Lexèmes-cibles connus en certains-lieux (géosynonymes)							%
		A	B	G	N	L	BN	S	
Environ 50 % des lexèmes-sources	arc-en-ciel	<i>ustruke, arkoiris, San Migelen zubi</i>		<i>Erromako zubi, arkoiris</i>	<i>Jainkoan paxa, ortzadar</i>	<i>ortzadar</i>			Environ 80 % des lexèmes- cibles
	beaucoup	<i>asko</i>		<i>asko, anitz</i>		<i>anitz</i>			
	chauve-souris	<i>saguzar</i>				<i>gauenara</i>			
	chemise	<i>alkandora</i>			<i>atorra, alkandora</i>	<i>atorra</i>			
	domestique	<i>morroi, krijedu, otzein</i>		<i>morroi</i>	<i>morroi, mutil, sehi</i>	<i>mutil, sehi</i>		<i>mutil, miskandi</i>	
	maire	<i>alkate</i>				<i>auzapez, mera</i>		<i>mera</i>	
	naître	<i>jaio</i>			<i>jaio, sortu</i>	<i>sortu</i>			
	printemps	<i>udaberri</i>				<i>primadera</i>		<i>bedatse</i>	
	soixante	<i>hirurogei</i>			<i>hirurogei, hiruretan hogeï</i>	<i>hiruretan hogeï</i>			
	soleil	<i>eguzki</i>						<i>ekhi</i>	
	tante	<i>izeko</i>		<i>izeba, izoa, izeko, tía</i>		<i>(ma)tanta, izeba</i>			

* Pour des raisons de place seuls quelques lexèmes sont indiqués à titre d'exemple mais les pourcentages proposés portent sur la totalité des 520 lexèmes-sources et 1335 lexèmes-cibles de la synthèse de l'enquête. Les lexèmes sont transcrits en orthographe standard.

Les subdivisions dialectales correspondent grosso modo aux sept provinces historiques : A : Alava ; B : Biscaye ; G : Guipuscoa ; N : Navarre ; L : Labourd ; BN : Basse-Navarre ; S : Soule.

La totalité des « lexiques-cibles connus partout » au sens d'*EAEL* (partie haute du tableau) est présentée dans les annexes 2 et 3.

ANNEXE 2

Lexique commun (corpus EAEL)

- ahizpa* « sœur / hermana (de ♀) »⁴⁷
aho « bouche / boca »⁵
ahuntz « chèvre / cabra »⁸³
aingira « anguille / anguila »⁴⁴⁷
aita « père / padre »⁴¹
aitagarreba « père / suegro »³⁴⁹
aitzur « houe / azada »¹³⁸
aizkora « hache / hacha »¹³⁶
alaba « fille / hija »⁴⁵
algodo(i)n, kotoi « coton / algodón »⁸⁰
ama « mère / madre »⁴²
amaginarreba « belle-mère / suegra »³⁵⁰
anaia « frère / hermano »⁴⁶
ardi « brebis / oveja »⁴²⁵
ardo « vin / vino »⁴⁷²
argi « lumière / luz »⁵⁴⁹
arrain « poisson / pez »⁹⁷
arrano « aigle / águila »⁴⁴²
arratoi « rat / rata »⁴²⁸
arrautza « oeuf / huevo »⁴⁵⁸
arreba « sœur / hermana (de ♂) »⁴⁷
arroz, irris « riz / arroz »⁷⁶
arto « maïs / maíz »³⁹²
aste « semaine / semana »⁵⁴²
asto « âne / asno »⁴²⁰
atzo « hier / ayer »¹⁹³
azal « écorce / corteza »⁴¹¹
baba « fève / haba »³⁹⁴
bai « oui / sí »²⁰⁰
barau « jeûne / ayuno »⁵⁶⁸
bat « 1 »¹⁵⁵
bataio (bautismo) « baptême / bautismo »⁵⁶⁵
bederatz « 9 »¹⁶³
begi « œil / ojo »²
behi « vache / vaca »⁴²³
belar « herbe / hierba »⁷⁵
belarri « oreille / oreja »³
belaun « genou / rodilla »²⁸⁵
bele « corbeau / cuervo »⁴⁴⁴
beltz « noir / negro »²¹⁹
berde « vert / verde »²²¹
berdin « également / igualmente »⁵¹⁶
bero « chaud / caliente »²²²
berrehun « 200 »¹⁸⁰
berrogei « 40 »¹⁷³
berrogeita hamar « 50 »¹⁷⁴
beso « bras / brazo »¹²
bi, biga « 2 »¹⁵⁶
biak : « deux (les) / ambos »¹⁹¹
bigarren « second / segundo »¹⁸⁵
bihar « demain / mañana »¹⁹⁴
bihotz « cœur / corazón »⁹
bilo (ouest) / *ile* (est) « cheveu / pelo »⁶
bost (Sud), *bortz* (Nord) 5¹⁵⁹
bultzatu (du), *pusatu* (du) « pousser / empujar »³¹⁹
burdin « fer / hierro »¹⁰⁷
buru « tête / cabeza »¹
doble « double / doble »¹⁹⁰
edan (du) « boire / beber »²¹
egun on « bonjour / buenos días »²³²
egun « jour / día »¹⁹⁵
eho (du) « moudre / moler »⁴⁸¹
ehun « 100 »¹⁷⁹
ekarri (du) « apporter / traer »³¹⁰
elur « neige / nieve »⁵⁴⁵
eman (du) « donner / dar »⁶³
erbi « lièvre / conejo »⁴²⁷
erdi « moitié / mitad »¹⁸⁷
erein (du) « semer / sembrar »⁴⁸⁰
erle « abeille / abeja »¹⁰¹
erre (du) « brûler / quemar »⁴⁷⁵
errege « roi / rey »³⁹
esan (du), *erran* (du) « dire / decir »²²
esku « main / mano »¹³
esne « lait / leche »³⁰⁴
etxe « maison / casa »⁶⁵
euli « mouche / mosca »⁹⁹
eur « pluie / lluvia »²²⁴
ez « non / no »²⁰¹
ezker « gauche / izquierda »⁵²⁶
ezpain « lèvres / labio »²⁷⁸
ezi « miel / miel »⁴⁶⁶
fruitu « fruit / fruto »⁴¹⁵
garagar « orge / cebada »³⁹⁰
gatu / katu « chat / gato »⁸⁷
gatz « sel / sal »¹¹²
gau « nuit / noche »¹⁹⁶
gazta, gasna « fromage / queso »⁴⁶⁹
gazte « jeune / joven »²¹⁴
gezurti « menteur / mentiroso »⁵³⁷
gibel « foie / hígado »²⁹⁷
gizon (arra) « homme mâle / hombre macho »³⁴
gizon « homme / hombre »³³
gogor « dur / duro »⁵²⁷
goiz « matin / mañana »⁵⁴⁰
gorri « rouge / rojo »²²⁰
gu, guk « nous / nosotros »¹⁴⁸
guti « peu / poco »¹⁹⁸
gutiago « moins / menos »⁵¹⁷
haize vent / viento »²²⁹
hamabi « 12 »⁶⁶
hamahiru « 13 »¹⁶⁷
hamaika « 11 »¹⁶⁵
hamar mila « 10 000 »¹⁸³
hamar « 10 »¹⁶⁴
han « là-bas / allí »⁵²⁰
handi « grand / alto »²⁰⁸
harri « pierre / piedra »¹⁰⁵
hartu (du) « prendre / tomar »²³
hartz « ours brun / oso pardo »⁴³¹
hau « celui-ci, celle-ci, ceci / este, esta, esto »¹⁵¹
hauts « cendre / ceniza »¹²⁵
hego, hegal « aile / ala »⁴⁵³
hemen (heben) « ici / aquí »⁵¹⁹
herri « village / pueblo »⁷¹
heste « intestins / intestinos »²⁹⁸
hezur « os / hueso »²⁹³
hil (du) « tuer (transitif) / matar (transitivo) »⁵⁶
hiru 3¹⁵⁷
hirugarren « troisième / tercero »¹⁸⁶
hirurehun « 300 »¹⁸¹
hoge « 20 »¹⁶⁸
hogeita bat « 21 »¹⁶⁹
hogeita bi(ga) « 22 »¹⁷⁰
hogeita hamar « 30 »¹⁷²
hogeita hiru « 23 »¹⁷¹
hotz « froid / frío »²²³
idi « bœuf / buey »⁸²
ikasi (du) « apprendre / aprender »³³⁹
ikatz « charbon / carbón »⁴⁷¹
ikusi (du) « voir / ver »¹⁷
ile, bilo « poils / pelos »⁴⁴⁹
irin « farine / harina »⁴⁷⁰
itsaso « mer / mar »¹¹⁹
itsu « aveugle / ciego »⁵³⁸
izar « étoile / estrella »¹¹⁸
izerdi « sueur / sudor »³⁰³
jakin (du) « savoir / saber »³³⁸
jan (du) « manger / comer »²⁰
jo (du) « frapper / pegar »⁵⁵
joan (da) « aller / ir »²⁵
josi (du) « coudre / coser »⁴⁸⁶
kafe « café / café »⁴⁰⁷
kaka egin (du) « déféquer / defecar »³³¹
ke « fumée / humo »¹²⁴
kobre « cuivre / cobre »¹¹⁰
koinata « belle-sœur / cuñada »³⁵²
koinatu « beau-frère / cuñado »³⁵¹
kontatu (du) « raconter / contar »³⁴⁰
lau (Sud), *laur* (Nord) 4¹⁵⁸
lehoi « lion / león »⁹⁰
liho « lin / lino »⁸⁰
lo egin (du) « dormir / dormir »³¹

(i)loba « neveu / sobrino »³⁵⁹
luma « plume / pluma »⁴⁵⁴
lur « terre / tierra »¹⁰⁶
luze « long / largo »²¹⁰
mahai « table / mesa »³⁸⁸
mando « mulet / mulo »⁸¹
mila « 1000 »¹⁸²
mingain (Sud) / *mihi* (Nord)
 « langue / lengua »⁸
mo(s)ko, piko « bec / pico »⁴⁵⁶
mozkortu (egin) (da) « s'enivrer /
 emborracharse »³³³
negar egin (du) « pleurer /
 llorar »³²²
negu « hiver / invierno »²²⁶
ni, nik « moi, je / yo »¹⁴⁵
odol « sang / sangre »³⁰¹
ogi « pain / pan »¹³³
oilar « coq / gallo »⁴⁴⁵
oilo « poule / gallina »⁸⁸
olio « huile / aceite »¹³⁵
olo « avoine / avena »³⁹⁰
on « bon / bueno »²⁰⁶

orain « maintenant / ahora »⁵¹⁴
otso « loup / lobo »⁹¹
pazientzia « patience /
 paciencia »⁵⁶⁰
piper « piment / pimiento »⁴⁰⁸
pixa egin (du) / txixa egin (du)
 « uriner / orinar »³³⁰
putzu « puits / pozo »⁶⁹
sagardo « cidre / sidra »⁴⁷²
sagu « souris / ratón »⁴²⁸
saltsa « sauce / salsa »¹³⁴
sare « filet / red »⁵⁰⁷
sartu (da) « entrer / entrar »²⁷
sei 6¹⁶⁰
seme « fils / hijo »⁴⁴
su « feu / fuego »¹²²
sudur « nez / nariz »⁴
suge « serpent / serpiente »⁹⁵
suhi « gendre / yerno »³⁵³
tigre « tigre / tigre »⁹⁰
txakur « chien / perro »⁸⁶
txindurri, inurria « fourmi /
 hormiga »⁹⁸

txori « oiseau / pájaro »⁹²
uda « été / verano »²²⁷
ur « eau / agua »¹²¹
urki « bouleau / abedul »⁴⁰⁴
urte « année / año »²³¹
urtu (da) « fondre / derretir »⁴⁷⁶
zahar « vieux / viejo »²¹⁵
zahartu (da) « vieillir /
 envejecer »⁵⁴
zazpi « 7 »¹⁶¹
zerra, sega « scie / sierra »⁴⁸⁸
zeru « ciel / cielo »¹¹³
zezen « taureau / toro »⁴²²
zorri « pou / piojo »¹⁰⁴
zortzi « 8 »¹⁶²
zubi « pont / puente »³⁸⁵
zuek « vous / vosotros /
 vosotras »¹⁴⁹
zuri « blanc / blanco »²¹⁸
zutik (dago / da) « debout / estar en
 pie »²⁴

Exposant : n° de la question *EAEL*

Souligné : origine latino-romane.

Total lexèmes-cibles communs (TLC) = 211.

Total lexèmes-sources (TLS) = 520.

TLS / TLC = 41 %.

ANNEXE 3

Lexique quasi commun (corpus *EAEL*)*

<i>adar</i> « branche / rama » ⁴⁰⁹	<i>gehiago</i> « plus / más » ¹⁹⁹	<i>moskito, mustika</i> « moustique / mosquito » ¹⁰⁰
<i>ahal dut (dut)</i> « je peux / puedo » ²⁰²	<i>hanka</i> (Pays basque d'Espagne) / <i>zango</i> (Pays basque de France)	<i>nor</i> « qui / quién » ¹⁵³
<i>aita-amak</i> « parents / padres » ⁴³	« jambe / pierna » ¹⁵	<i>ohe</i> « lit / cama » ³⁸⁷
<i>aitzur</i> « bêche / azadón » ⁴⁷⁷	<i>har</i> « ver / gusano » ⁹⁶	<i>ontzi</i> « pot / cacharro » ¹³⁰
<i>amets egin (du)</i> « rêver / soñar » ³³⁴	<i>hari</i> « fil / hilo » ⁴⁸³	<i>orratz</i> « aiguille / aguja » ⁴⁸⁵
<i>ardi</i> « mouton / oveja » ⁸⁴	<i>hazi</i> « graine / grano » ⁴¹⁶	<i>osaba</i> « oncle / tío » ³⁴³
<i>arin</i> « léger / ligero » ⁵³¹	<i>hi, hik</i> « toi, tu » ¹⁴⁶	<i>pinu</i> « sapin / pino » ⁴⁰³
<i>bete</i> « plein / lleno » ²⁰⁴	<i>hodei</i> « nuage / nube » ²²⁸	<i>piztu (du)</i> « allumer / encender » ⁴⁷³
<i>bildots</i> « agneau / cordero » ⁴²⁶	<i>hori</i> « jaune / amarillo » ²²¹	<i>soka</i> « corde / cuerda » ¹²⁷
<i>biri(ka)(k)</i> « poumons / pulmones » ²⁹⁵	<i>huts</i> « vide / vacío » ²⁰⁵	<i>titi</i> « mamelle / mama » ²⁸¹
<i>bular</i> « poitrine / pecho » ²⁸⁰	<i>igitai</i> « faucille / hoz » ¹³⁷	<i>txar, tzar</i> « mauvais / malo » ²⁰⁷
<i>buztan</i> « queue / cola » ⁴⁵¹	<i>ikus(i) arte</i> « au revoir / hasta la vista » ²³⁴	<i>tximino</i> « singe / mono » ⁴³⁷
<i>buztin</i> « argile / arcilla » ⁴⁶⁰	<i>iturri</i> « source / manantial » ³⁸⁴	<i>tximista</i> « éclair / relámpago » ⁵⁴⁷
<i>eder</i> « beau / bello » ²¹⁶	<i>izan (da)</i> « être / ser » ³⁶⁷	<i>urdin</i> « bleu / azul » ²²¹
<i>egosi (du)</i> « cuire (transitif) / guisar (transitivo) » ¹³¹	<i>janari, jateko</i> « nourriture / comida » ¹²⁹	<i>uso</i> « pigeon / paloma » ⁹³
<i>egur, zur</i> « bois / madera » ¹²⁶	<i>jarri (da)</i> « asseoir / sentar » ³¹⁶	<i>zahar</i> « vieillard / anciano » ⁴⁸
<i>errain</i> « belle-fille / nuera » ³⁵⁴	<i>kaka</i> « excréments / excrementos » ³⁰⁶	<i>zaldi</i> « cheval / caballo » ⁴¹⁸
<i>erreka</i> « rivière / río » ¹²⁰	<i>larru</i> « peau / piel » ²⁹²	<i>zikin</i> « sale / sucio » ⁵³³
<i>(etxe) egin (du)</i> « bâtir / construire » ⁶⁰	<i>laurden</i> « quart / cuarto » ¹⁸⁹	
<i>garbi</i> « propre / limpio » ⁵³²	<i>makila</i> « bâton / palo » ¹⁴¹	
<i>garratz</i> « amer / amargo » ⁵³⁴	<i>min</i> « douleur / dolor » ⁵⁵⁷	
<i>gau on</i> « bonsoir / buenas noches » ²³³		

Total lexèmes-cibles quasi communs (TLCQC) = 60 lexèmes-cibles.

TLC + TLCQC / TLS = 261 / 520 = 50 %.

* Lexique quasi commun : lexèmes-cibles mis en correspondance avec un même lexème-source par une majorité mais pas par la totalité des informateurs. Exemple : une majorité d'informateur a répondu *osaba* au lexème-source « oncle / tío »³⁴³ mais quelques informateurs ont répondu *tio* en Pays Basque Sud et *otto* en Pays Basque Nord. On suppose ici que les informateurs ayant répondu *tio* ou *otto* connaissent également le terme *osaba*. Ce correctif porte le lexique commun de 41 % à 50 %.

Les annexes 2 et 3 recensent la totalité du lexique commun dans le périmètre du questionnaire *EAEL*. Cette recension ne représente par définition qu'une partie du lexique commun de la langue.

[Annexes 1 et 2 : origine autochtone : 213 / 261 = 82 % ; origine latino-romane : 48 / 261 = 18 %.]

INDICATIONS BIBLIOGRAPHIQUES

- AZKARATE Miren et ALTUNA Patxi, 2001, *Euskal morfologiaren historia*, Saint-sébastien, Elkarlanean.
- AZKUE (de) don Resurrección María, 1905 [1984, réédition de l'Euskaltzaindi], *Diccionario vasco-español-francés*, Bilbao, Euskaltzaindia.
- DGV : *Diccionario general vasco*, 1967-2005, Bilbao, Euskaltzaindia.
- EDEL : *Euskalerriko atlas etnolinguistikoa*, LEIZAOLA Fermín (dir.), 1983 (1^{re} partie) ; 1990 (2^e partie), Saint-Sébastien.
- EHHA : *Euskal Herriko hizkuntz atlasa*, 1999, *Ohiko euskal mintzamoldeen antologia*, Bilbao, Euskaltzaindia.
- HARITSCHELHAR Jean (dir.), *Être basque*, 1983, Toulouse, Éditions Privat.
- Hiztegi batua*, 2000, Travaux et actes de l'Académie de la Langue basque, Bilbao, Euskaltzaindia.
- LAFON René, 1944, *Le système du verbe basque au XVI^e siècle*, Elkar, Saint-Sébastien.
- LAKARRA Joseba, 1996, *Refranes y sentencias (1596), ikerketak eta edizioa*, Bilbao, Euskaltzaindia.
- LHANDE Pierre, 1926, *Dictionnaire basque-français*, Paris, Maisonneuve.
- MARTINET André, 1970, *Éléments de linguistique générale*, Paris, Armand Colin.
—, 1975, *Évolution des langues et reconstruction*, Paris, PUF.
- MEILLET Antoine, 1905-1906, « Comment les mots changent de sens », dans *l'Année sociologique*.
—, 1928, *Les langues dans l'Europe nouvelle*, Paris, Paillot.
—, 1982 (1^{re} éd. : 1921-1936), *Linguistique historique et linguistique générale*, Genève, Slatkine ; Paris, Champion, 1948.
- MICHELENA Luis, 1964, *Textos arcaicos vascos*, Madrid, Minotauro.
—, 1990 (1^{re} éd. : 1961), *Fonética histórica vasca*, Saint-Sébastien, Anejos del Anuario de Filología Vasca « Julio de Urquijo », IV, Serie « Obras completas de Luis Michelena », I.
- ORPUSTAN Jean-Baptiste, 1996 (dir.), *La langue basque et les autres : influences, comparaisons*, Saint-Etienne-de-Baïgorry, Izpegi.
- PICOCHÉ Jacqueline, 1992 (1^{re} éd. 1977), *Précis de lexicologie française*, Paris, Nathan.
- PINKER Steven, 1999, *L'instinct du langage*, Paris, Odile Jacob.
- SAPIR Eduard, 1967, *Anthropologie*, Paris, Les Éditions de Minuit (Folio-Essais).
—, 1968, *Linguistique*, Paris, Les Éditions de Minuit (Folio-Essais).
- YRIZAR Pedro de, 1981, *Contribución a la dialectología de la lengua vasca*, t. 1 et 2, Saint-Sébastien, Caja de Ahorros Provincial de Guipúzcoa.
- ZUAZO Koldo, 2003, *Euskalkiak : Herriaren lekukoak*, Donostia, Elkar.