



Revista de Filología y Lingüística de la Universidad de Costa Rica

Publicación Semestral, EISSN: 2215-2628

Volumen 48 - 2

Julio 2022 - Diciembre 2022

**Analyse de l'évolution de la disponibilité
lexicale d'un échantillon d'élèves apprenant
le FLE durant le cycle secondaire
en Espagne**

Gabriel Díez Abadie

Díez Abadie, G. (2022). Analyse de l'évolution de la disponibilité lexicale d'un échantillon d'élèves apprenant le FLE durant le cycle secondaire en Espagne. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 48(2), e50959. doi: <https://doi.org/10.15517/rfl.v48i2.50959>



Doi: <https://doi.org/10.15517/rfl.v48i2.50959>

URL: <https://revistas.ucr.ac.cr/index.php/filyling/index>

Analyse de l'évolution de la disponibilité lexicale d'un échantillon d'élèves apprenant le FLE durant le cycle secondaire en Espagne

Analysis of the Evolution of the Lexical Availability of a Sample of Students Learning French as a Foreign Language Throughout the Stage of Secondary School in Spain

Gabriel Díez Abadie

Universidad de las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España

gabriel.diez@ulpgc.es

<https://orcid.org/0000-0003-0107-3917>

DOI: 10.15517/rfl.v48i2.50959

Reçu le : 07-02-22

Accepté le: 17-03-22

RÉSUMÉ

Si les avancées en termes de disponibilité lexicale ne sont que très récentes, depuis ces dernières années, les recherches dans ce domaine se sont considérablement approfondies. Bien que de telles recherches foisonnent dans le monde hispanique, ce travail s'inscrit, modestement, dans cette lignée et revient à ses origines. En effet, les recherches sur le concept de disponibilité lexicale ont vu le jour en France. En outre, ce projet expérimental propose d'établir un état des lieux de l'évolution du vocabulaire disponible en langue étrangère grâce à l'analyse d'un corpus réalisé auprès d'élèves de langue maternelle espagnole apprenant le français comme seconde langue étrangère en Espagne. Notre expérience avec un échantillon composé de 294 élèves met l'accent sur les possibilités que peuvent offrir ce type d'étude à l'enseignement du vocabulaire pour l'élaboration de supports en fonction des besoins d'apprentissage plus axés sur la qualité de l'apprentissage dans le temps. En effet, la courbe reflétant l'évolution du bagage lexical n'est pas toujours croissante mais plutôt irrégulière. Ce qui confirmerait peut-être d'autres études sur le fait que l'apprentissage n'est pas linéaire mais en spirale, fait d'allées et venues.

Mots clés: vocabulaire disponible; disponibilité lexicale; unités lexicales; vocables; FLE.

ABSTRACT

Even if advances regarding lexical availability are only very recent, during the last few years, research in this area has progressed considerably. Although such studies abound in the Hispanic world, this research intends to modestly support these lines of investigation, returning to their origins. Indeed, the concept of research as lexical availability was born in France. In addition, this experimental project proposes to establish an inventory of the evolution of the available vocabulary in a foreign language thanks to the analysis of a corpus produced with students of Spanish mother tongue learning French as a second foreign language. Our experience with a sample made up of 294 students highlights the possibilities that this type of study can offer with regard to the teaching of vocabulary for the development of resources responding to specific learning needs focusing on the quality of the language learning over time. Indeed, the curve which reflects the evolution of lexical knowledge is not always increasing but rather irregular. This would perhaps confirm other studies stating the fact that learning is not linear but spiral, made of comings and goings.

Keywords: available vocabulary; lexical availability; lexical units; lemma; FLE.

Lorsque l'on demande aux apprenants d'identifier les principaux enjeux qui empêchent la communication en langue étrangère, le manque de vocabulaire apparaît comme le principal obstacle. D'ailleurs, il est fréquent de constater que les apprenants se plaignent de ne pas avoir suffisamment de vocabulaire. Afin d'illustrer cette hypothèse nous avons mené une enquête auprès de 294 informateurs (tous les élèves d'un établissement scolaire espagnol inscrits en Français Langue Étrangère, FLE dès à présent) parmi lesquels 41 % indiquent que le manque de vocabulaire est le principal frein à la communication, 27 % pensent que c'est la prononciation tandis que 32 % évoquent la grammaire. Comme l'indique Courtillon (1989, p. 147) « le lexique est le pivot de l'acquisition à partir duquel s'organise la syntaxe et plus tard, la morphosyntaxe. Cela s'explique aisément par le fait que le lexique, haut porteur d'informations, contribue, avec l'intonation, à donner rapidement aux élèves l'accès à la communication ».

Cette prise de conscience de l'importance du lexique, très longtemps considéré comme de second plan face à la grammaire, apparaît comme une revendication héritée d'une tradition de la didactique des langues étrangères, DLE dès à présent, trop axée sur cette dernière.

Par conséquent, il faudra attendre la deuxième moitié du XX^{ème} siècle pour que l'on accepte de nouveau l'idée que de bonnes connaissances lexicales sont fondamentales pour un usage réel de la langue. Même s'il n'est pas question de tomber dans le piège de trop se focaliser sur un manque de vocabulaire qui bloquerait éventuellement l'apprentissage ou la communication, il paraît essentiel de lui redonner sa place dans nos pratiques pédagogiques au même titre que la grammaire et la prononciation.

Pour comprendre cette position secondaire, il paraît nécessaire de revenir sur l'histoire de la place du lexique dans la DLE au XX^{ème} siècle, c'est pourquoi nous reprendrons cet aspect dans la première partie. Bien que la question des modalités d'apprentissage du vocabulaire ait été beaucoup

abordée nous n’y reviendrons pas de manière exhaustive mais nous nous centrerons sur certains points que nous considérons essentiels tels que le principe de *sélection lexicale* ou encore la différence essentielle entre *vocabulaire disponible* et *vocabulaire actif*. En effet, notre intérêt se porte sur la question du *vocabulaire disponible* des apprenants et plus précisément sur son évolution au cours de la scolarité obligatoire.

Si les avancées en termes de *disponibilité lexicale* ne sont que très récentes, depuis ces dernières années, les recherches dans ce domaine se sont considérablement approfondies. Par ailleurs, l’avènement de l’ambitieux projet coordonné par Humberto Lopez Morales, le *Projet Panhispanique* (*Proyecto Panhispánico*), peut être considéré comme un tournant dans l’histoire de la *statistique lexicale*. L’objectif principal de ce projet est de recenser le *vocabulaire disponible* en langue maternelle d’échantillons d’apprenants scolarisés dans l’enseignement secondaire avant leur entrée à l’université, et cela pour toutes les variantes de l’espagnol. Bien que de telles recherches foisonnent dans le monde hispanique, ce travail s’inscrit, modestement, dans cette lignée et revient à ses origines, dans la mesure où les recherches sur le concept de *disponibilité lexicale* (Michéa, 1950; Gougenheim et al., 1964) ont vu le jour en France. En outre, ce projet expérimental propose d’établir un état des lieux de l’évolution du *vocabulaire disponible* en langue étrangère grâce à l’analyse d’un corpus réalisé auprès d’élèves de langue maternelle espagnole apprenant le français comme seconde langue étrangère.

Les applications de la *disponibilité lexicale* sont diverses, elles ne se limitent pas à la DLE mais abordent aussi la didactique de la langue maternelle, la dialectologie, la sociolinguistique, la psycholinguistique ou encore l’ethnolinguistique. En DLE il est possible de réaliser un corpus à partir de locuteurs natifs ou alors à partir de locuteurs apprenants la langue étrangère. Dans le premiers cas les publications sont nombreuses, qu’elles soient d’ailleurs intégrées au *Projet Panhispanique* ou non, alors que dans l’autre cas peu nombreux sont les chercheurs qui se sont lancés dans cette expérimentation, bien qu’ils soient de plus en plus, selon Florentino Paredes Garcia (2012), mais toujours très peu au niveau du secondaire.

Notre expérimentation consiste en l’analyse de l’évolution du *vocabulaire disponible* sur des thèmes de la vie quotidienne d’apprenants espagnols de FLE pendant toute la période de l’enseignement secondaire obligatoire (*Educación Secundaria Obligatoria, ESO*) et non obligatoire (*Bachillerato*), de la première année (*1° ESO*) jusqu’à la dernière (*2° de Bachillerato*). Les préoccupations de nos élèves par rapport à leur manque de vocabulaire correspondent-elles à la réalité ? Qu’en est-il de l’évolution du *vocabulaire disponible*, le bagage lexical des élèves stagne-t-il ou est-ce qu’il s’accroît au fil des années ? Le *vocabulaire disponible* actualisé par nos apprenants dans leurs

enquêtes est-il un vocabulaire réel ou artificiel produit de la classe ? Voici les questions autour desquelles sont articulés nos objectifs et les hypothèses de cette étude.

Après l'exposition des différents aspects théoriques afin de comprendre les enjeux de l'application de la disponibilité lexicale en DLE mais aussi ses apports et ses limites, nous présenterons la méthodologie utilisée dans la réalisation de notre étude. Enfin, la dernière partie sera dédiée à la comparaison et à l'analyse des résultats dans laquelle nous mettrons à l'épreuve nos hypothèses et les concepts repris dans la première partie. Dans la conclusion, nous tentons de donner des pistes de réflexion d'abord sur l'évolution de la *disponibilité lexicale* pendant toute la période de l'enseignement secondaire obligatoire et non obligatoire, puis sur l'exploitation de nos résultats.

1. La disponibilité lexicale et la DLE

1.1. Le lexique dans l'enseignement des LE

Pour comprendre la position secondaire du lexique il paraît nécessaire de revenir sur l'histoire de la place du lexique dans la DLE au XX^{ème} siècle. Galisson (1991) distingue deux périodes : celle des *vaches grasses* et celle des *vaches maigres*. La première correspond à la période des méthodes plus traditionnelles dans laquelle un vocabulaire riche était présenté aux apprenants par le biais de textes littéraires et à la période des méthodes directes qui étaient à l'opposé des premières mais qui maintenaient une grande pression au niveau lexical. La période des *vaches maigres* est, selon Galisson, celle qui correspond à l'avènement des méthodes audiovisuelles (années 50-60) dans lesquelles un *vocabulaire minimal* suffit à une meilleure acquisition de *mécanismes de bases*. Galisson (1991, p. 12) conclut de la manière suivante : « tout le travail effectué pour l'élaboration du f. f. [*français fondamental*] a surtout servi... à desservir le vocabulaire et abouti à sa paupérisation dans les méthodes. (...) le f. f. [a joué] le rôle inattendu de fossoyeur du vocabulaire dans la pédagogie des langues ». Paul Boggards (1994, p. 7) signale que dans les années 70 « tous les aspects lexicaux sont presque totalement absents dans les textes classiques ». Il indique également que la préoccupation première dans les livres de didactique est celle de la complexité de l'acquisition grammaticale et que la question de l'apprentissage du vocabulaire n'est pas traitée. Coady et Huckin en 1997 nous résumaient parfaitement la situation en indiquant que l'attitude et les convictions entre les enseignants et les apprenants ne convergent pas forcément. Ainsi, d'une part les auteurs soulignent toute l'importance accordée aux mots par les apprenants et d'autre part, le fait que les enseignants pensent que le problème se réduit à la grammaire. Pour eux, les mots sont faciles à apprendre et leur dédier du temps n'est pas une priorité. Même si dès la fin des années 70 de nombreuses publications voient le jour sur la question du lexique, il faut attendre le milieu des années 80 pour que l'on accepte l'idée que

le manque de vocabulaire nuit à la bonne compréhension et par conséquent à la communication. Lewis, de son côté, dénonçait en 2000 les problèmes que supposent les défaillances au niveau lexical, particulièrement chez les élèves ayant atteint un certain niveau.

La revendication du rôle fondamental que joue le vocabulaire dans l'enseignement / apprentissage des langues étrangères n'est donc que récente. On ne pourra nier l'influence et les apports des théories linguistiques telle que l'analyse du discours, la pragmatique, la psycholinguistique, ou encore la linguistique cognitive. C'est d'ailleurs de cette préoccupation que sont nées de nouvelles méthodes, comme l'approche lexicale que cite Michael Lewis (2000), *The Lexical Approach*, ou le Français sur Objectifs Spécifiques. Ainsi, apparaissent des dictionnaires spécifiques à la LE, ou encore, du matériel spécifiquement focalisé sur des besoins lexicaux donnés. Le *Cadre européen commun de référence pour les langues*, ou *CECRL* (Conseil de l'Europe, 2001, p. 115), reprend également cet aspect fondamental :

La quantité, l'étendue et la maîtrise du vocabulaire sont des paramètres essentiels de l'acquisition de la langue et, en conséquence, de l'évaluation de la compétence langagière de l'apprenant et de la planification de l'enseignement et de l'apprentissage de la langue.

Si l'on s'accorde à reconnaître que le vocabulaire est important il n'en demeure pas moins complexe. De surcroît, on insistera sur le fait que c'est la composante la plus instable, bien plus que la composante phonétique ou grammaticale. Comme le soulignent de nombreux spécialistes (Lewis, 2000 ; Coady et Huckin, 1997 ou encore Martín Martín, 2007), cette complexité qui complique ou empêche sa systématisation est due aux trois facteurs suivants : le grand nombre d'unités qui la composent, la variation lexicale et la vitalité de la langue. Dans le premier cas, le nombre d'unités qui conforment le lexique d'une langue est quasi illimité puisque selon la situation d'énonciation un même mot peut avoir des sens, des interprétations différentes. Ensuite, lorsqu'on évoque la variation lexicale, très étudiée par Humberto Lopez Morales, on l'entend sous toutes ses formes, *diatopique*, *diastratique* ou *diaphasique*. Et finalement, le facteur vitalité de la langue sous-entend que c'est au niveau lexical que l'évolution est la plus fréquente. En effet, si l'évolution est quelque chose d'inhérent à la langue, c'est au niveau lexical que se manifeste ce principe de la manière la plus évidente. En effet, les mots doivent s'adapter à de nouvelles réalités, en revanche les différentes réalisations phonologiques n'altèrent pas la compréhension tant qu'elles ne modifient pas les traits pertinents et que la grammaire suit des règles bien définies.

Faut-il procéder à un enseignement explicite du vocabulaire ? De nos jours de nombreux courants linguistiques, mentionnés précédemment, font ce postulat et organisent le vocabulaire autour de différents champs lexicaux ou notionnels. Si on accepte d'enseigner ce vocabulaire « illimité » il va donc falloir procéder à une sélection adéquate. Un choix de vocabulaire à enseigner à un moment

donné de l'apprentissage pour pouvoir organiser une gradation lors de la construction de la situation d'apprentissage, en situation de besoin. Il s'agit de trouver comment fournir aux apprenants les outils nécessaires à la communication, identifier des objectifs spécifiques pour constituer des actes de paroles correspondant à une réalité sociale, se détacher du caractère institutionnel de l'enseignement / apprentissage de la langue étrangère en milieu scolaire pour que les élèves y voient une utilité. D'autre part, le *CECRL* (Conseil de l'Europe, 2001, p. 115) affirme de façon catégorique que « les concepteurs d'examens et de matériel pédagogique sont tenus de choisir le vocabulaire qu'ils y feront entrer ». Bien qu'en ce qui concerne cette sélection, le *CECRL* ne mentionne en aucun cas la planification initiale pour l'enseignement du lexique. L'approche communicative, comme le souligne Galisson (1997, p. 64), « n'a pas suffisamment valorisé le lexique. L'intégration de documents authentiques a provoqué l'abandon de tout principe de sélection ».

Si l'on accepte le principe de *sélection lexicale*, comment faut-il alors procéder et quels sont les critères à suivre pour nous guider dans cette tâche ? Même s'il est difficile de s'accorder sur les critères, on peut considérer celui d'utilité, comme un critère clé dans notre quête de sélection. Ainsi, il s'agit de munir les apprenants d'un vocabulaire utile à chaque étape de leur apprentissage dans des situations de communication variées et des contextes de communication divers. Oui mais, comment mesurer le degré d'utilité d'un mot ? Quels sont les mots considérés comme utiles dans une langue ?

De nos jours, de nombreux auteurs s'accordent à dire que le concept d'utilité dépasse celui de fréquence. Danielle Bailly (1997) indique que la *sélection lexicale* dans le domaine scolaire doit se faire en accord avec le principe d'*utilité* et du *degré de facilité* d'apprentissage. De surcroît, il lui paraît fondamental que les dictionnaires de mots utiles en LE soient composés de *vocabulaire disponible*, de *vocabulaire fréquent*, de *vocabulaire instrumental*, et de *vocabulaire de la classe*. Dans cette lignée s'inscrivent les recherches de Izquierdo Gil (2004), qui développent le principe de *facilité* d'apprentissage. Selon lui, composé d'une part les *mots ressemblants* à ceux de la langue maternelle et d'autre part, par les mots ayant un indice élevé de *familiarité*. Comment doit-on procéder à la *sélection lexicale* ?

Le *CECRL* (Conseil de l'Europe, 2001, p. 149) propose quatre options possibles afin de procéder à cette sélection :

- Choisir des mots et des expressions clés **a.** dans les domaines thématiques exigés pour réaliser les tâches communicatives correspondant aux besoins des apprenants, **b.** qui concrétisent la différence culturelle et/ou les valeurs et croyances significatives partagées par le ou les groupes sociaux dont on étudie la langue.
- Suivre les principes de statistiques lexicales en sélectionnant les mots à grande fréquence dans un large corpus ou dans des domaines thématiques réduits.
- Sélectionner des textes (authentiques) oraux et écrits et apprendre/enseigner sans restriction les mots qu'ils contiennent.
- Ne pas planifier l'enrichissement du vocabulaire mais lui permettre de se développer organiquement en réponse à la demande de l'apprenant lorsqu'il entreprend des tâches communicatives.

Même si les deux derniers points peuvent être facilement remis en question, les deux premiers indiquent clairement la tendance suivie jusqu'alors. D'un côté, la sélection *subjective* (alinéa *a* du Conseil) basée sur l'intuition des auteurs de manuel. Malgré le fait que les résultats de son utilisation dépendent évidemment des connaissances et de cette intuition, cette méthode a été pendant longtemps la plus sollicitée. Elle peut produire des différences remarquables dans les recueils lexicaux. Et d'un autre côté, la sélection *objective*, basée sur un corpus dans lequel on applique des méthodes statistiques d'analyse. Ce genre de sélection est facilité par les progrès informatiques de ces dernières décennies qui ont vu fleurir ce genre de travaux. En effet, maintenant il existe de nombreux logiciels (ne serait-ce qu'Excel) qui permettent de faciliter le traitement des données en calculant des formules complexes, ou donnant la possibilité de trier les données, on gagne donc en temps et en précision.

1.2. Le vocabulaire disponible

On entend par *vocabulaire disponible* l'ensemble de mots qu'un locuteur, de n'importe quelle langue, est capable de mobiliser rapidement à partir d'associations d'idées grâce à un stimulus. Dans les enquêtes, ce stimulus est appelé *centre d'intérêt*, il s'agit d'un thème donné qui traite de la vie quotidienne dans lequel l'informateur s'applique à restituer spontanément le champ lexical. En DLE, les recherches sur le concept de *disponibilité lexicale* sont relativement récentes et ont vu le jour en France grâce aux apports de la statistique lexicale entre les années 1950 et 1960. Dans un premier temps, elles étaient destinées à déterminer quel vocabulaire enseigner aux apprenants de FLE en fonction des mots les plus utilisés par les locuteurs natifs. Au départ, le principe de fréquence était pris en compte mais c'est après s'être rendu compte que des mots usuels tels que « métro, autobus, lettre et timbre » n'apparaissent pas dans les listes que la méthodologie a été modifiée. C'est à Michéa en 1953 (pp. 338-344) que l'on doit la distinction très importante entre *mots fréquents* et *mots disponibles* et la définition suivante d'un *mot disponible* : « un mot qui, sans être particulièrement fréquent, est cependant toujours prêt à être employé et se présente immédiatement et naturellement à l'esprit au moment où l'on a besoin ». Selon lui, le *vocabulaire disponible* est donc celui que l'on utilise dans une situation de communication donnée, si celle-ci n'est pas abordée, certains termes n'ont pas lieu d'apparaître. D'après cette définition, on comprend qu'une langue se compose d'un ensemble de mots, quasi illimités, sémantiquement très concrets et mobilisés uniquement si le thème de la conversation le permet. C'est à dire qu'il existe des termes qui ne s'actualisent que s'il faut communiquer une information spécifique. Voilà pourquoi le facteur de fréquence ne peut être pris en compte. Pour cet auteur « la *disponibilité* s'oppose à la *fréquence*, et la complète » : la fréquence aide à « déterminer le vocabulaire de base d'une langue ».

Malgré le succès avec lequel les projets d'études de *disponibilité lexicale* se sont développés, il demeure encore beaucoup d'imprécisions dans la définition même de *lexique disponible*. Comme l'affirme Hernández Muñoz (2006, p. 28), l'action d'actualiser les mots disponibles est une « tâche cognitive complexe qui implique la mise en marche de différentes stratégies linguistiques et psychologiques », qu'il est nécessaire d'analyser pour comprendre et valoriser de manière globale ce qu'est le *vocabulaire disponible*. López Morales (1999, p. 26) constate l'importance de l'appareil psychologique dans la propre identité de « disponibilité » : « nous espérons bientôt pouvoir compter non seulement sur une théorie psychologique de la *disponibilité lexicale* qui conjugue des facteurs comme la mémoire sémantique, la familiarité, etc., mais aussi sur une théorie intégrée de la disponibilité ».

Les recherches en *disponibilité lexicale* dans l'état actuel des choses semblent s'inscrire dans cette lignée et coïncider avec l'idée avancée par Hernández Muñoz (2006). Elles partent donc d'une vision de la disponibilité comme une tâche cognitive dans laquelle les processus psychologiques sous-jacents conditionnent dans une grande mesure le vocabulaire produit. Comme Hernández Muñoz (2006, p. 29) l'indique, la plus grande partie des définitions récentes de *mots disponibles* font référence au « capital lexical » utilisable dans une situation de communication donnée. C'est ainsi que le définissent de nombreux auteurs. Traditionnellement, le lexique disponible est défini avec les concepts de « termes communs ou usuels », dans le but de refléter l'usage réel que les locuteurs font du vocabulaire. Par conséquent, selon cette auteure, le *vocabulaire disponible* désigne un « lexique potentiel qui appartient au vocabulaire actif des locuteurs et qui, comme tel, pourrait être actualisé dans une situation de communication quotidienne ». Mais elle nuance son propos en indiquant que bien que le concept « d'usage » associé implicitement aux mots disponibles est important, « dans les enquêtes qui recueillent ce *vocabulaire disponible*, rien ne nous indique que les locuteurs sont capables d'utiliser ou non un terme déterminé » (Hernández Muñoz, 2006, p. 46). Effectivement, être capable de nommer un mot ne veut pas forcément dire que l'on soit capable de l'utiliser de façon autonome dans une situation communicative donnée.

Les définitions apportées par les chercheurs français dans les années 50 font références aux mots qui viennent à l'esprit du locuteur au cours d'une association d'idées, pas dans le discours quotidien. La perspective qu'une langue est, avant tout, « langage en usage » n'est pas nouvelle en linguistique. Pour Michéa (1953) et Galisson et Coste (1976), le lexique disponible n'était pas un lexique utilisé dans un contexte communicatif réel, mais un vocabulaire produit sous la pulsion d'une situation expérimentale, induite expérimentalement, puisque très rarement dans la vie nous ne sommes amenés à réaliser ce type d'énumérations d'éléments d'une catégorie.

En réalité, souligne Hernández Muñoz (2006), la tâche qui consiste à énumérer des mots doit être considérée dans les limites de l'instrumentalité méthodologique, dans sa juste mesure, comme une « manière d'évaluer les capacités linguistiques d'un sujet de manière isolée et expérimentale ».

Pour conclure, nous pouvons dire que le vocabulaire disponible, bien qu'il se rapproche de l'usage que pourrait faire un locuteur du lexique, cet usage n'est pas forcément équivalent à celui qu'il ferait dans une situation communicative concrète. Réfléchir et apporter un mot comme réponse à un stimulus isolé dans le délai imparti ne peut être que le reflet très partiel et imparfait de la compétence communicative (qui comprend, entre autres, la compétence lexicale) d'un locuteur dans une situation de communication ordinaire.

1.3 La disponibilité lexicale et la DLE

L'histoire de la *disponibilité lexicale* compte deux grandes étapes unies par un lien étroit : d'un côté, on considère les apports de Michéa et Gougenheim dans les années 50 et 60, et d'un autre côté, les apports récents essentiellement en espagnol. Les premières recherches ayant permis d'asseoir les bases méthodologiques ont donc commencé en France il y a maintenant un demi-siècle, avec l'élaboration du *français fondamental*. L'objectif était de maintenir le lien linguistique avec les ex-colonies. Toutefois, en essayant de sélectionner le vocabulaire à enseigner le plus usuel, les auteurs, Gougenheim, Rivenc, Michéa et Sauvageot se sont rendus compte que certains mots très connus atteignaient une fréquence très basse et n'apparaîtraient pas dans le *français fondamental*. Ils ont alors décidé de compléter les données de fréquence avec leurs recherches sur la *disponibilité lexicale* : plus un mot est disponible pour un informateur, plus il y aura de chances que celui-ci l'inclue parmi ses réponses. C'est là toute la différence entre le *lexique de base* et le *lexique disponible* ; nous verrons que dans notre corpus ce sont les substantifs qui prédominent.

Au Canada, W. Mackey (1971), à l'instar de G. Dulong, reprend essentiellement la méthode utilisée par les chercheurs français pour confronter le *lexique disponible* de France à celui du Québec. Le fait d'appliquer la même méthodologie favorise l'objectif comparatif. Les études de *disponibilité lexicale* naissent directement en relation avec la linguistique appliquée. Le lexique fondamental d'une langue (formé par le vocabulaire de base et le *vocabulaire disponible*), s'éloigne beaucoup de celui qui compose les dictionnaires généraux. Ainsi, déjà en 1964, Gougenheim et al. insistent sur l'importance des applications pédagogiques du *français fondamental*. Comme nous le savons, la didactique du FLE était la motivation principale de Michéa et de son équipe avant d'entreprendre leurs recherches dans le champ de la *disponibilité lexicale*. En outre, les publications ayant suivi le *français fondamental* ont dans leur grande majorité un but éducatif. Certains chercheurs le précisent

explicitement comme objectif ultime. En effet, Mackey (1971) prétend appliquer les résultats de son travail à la didactique du français au Canada. En effet, les données extraites des recherches sur le *lexique disponible* s'avèrent utiles quand elles s'appliquent à la didactique.

Les recherches sur le lexique disponible entreprises par Gougenheim et ses collaborateurs s'inscrivent dans la lignée de l'enseignement du français. Dans cette perspective, plusieurs auteurs ont décrit les avantages que cette méthode d'étude apporte à la DLE, toutefois cet aspect n'a pas été suffisamment approfondi. En outre, si l'on s'intéresse concrètement aux travaux ayant mis en exergue l'importance des études de *disponibilité lexicale* pour l'enseignement du français à partir de l'analyse de matériaux obtenus à partir de locuteurs apprenants d'une langue étrangère reste modeste voire quasi inexistante. En revanche, en ce qui concerne l'enseignement de l'espagnol comme langue étrangère, les travaux de recherche de Alberto Carcedo (1999 et 2000), professeur à l'université de Turku, en Finlande, et Marta Samper Hernández (2002), professeur à l'université de las Palmas de Gran Canaria, en Espagne, ont été les premiers à avoir approfondi cet aspect et marque le point de départ d'une bibliographie abondante.

Alberto Carcedo est à l'origine de nombreux articles, et même si la majorité est de nature pratique certains sont à caractère éminemment théorique. Il insiste sur la pertinence des enquêtes réalisées avec les étrangers, puisqu'elles permettent, entre autres, d'examiner les différentes phases d'apprentissage du lexique de la langue étrangère, l'espagnol dans son cas, de découvrir autant les erreurs les plus fréquentes que la tendance des apprenants, au long de cet apprentissage, d'organiser la progression de l'enseignement du lexique et de réaliser des comparaisons fructueuses avec d'autres listes de *vocabulaire disponible*.

En 2002, Marta Samper Hernández a repris cette méthodologie et l'a appliquée à des étudiants universitaires étrangers de l'université de Salamanque. Elle se base sur un échantillon de 45 élèves (dont 13 de niveau supérieur, 10 de niveau avancé et 14 de niveau intermédiaire et 8 de niveau élémentaire). Dans cette étude elle reprend plusieurs variables dont la langue maternelle de ses étudiants, le niveau d'études, ainsi que le sexe des informateurs. Samper Hernández reprend la méthodologie du *Projet Panhispanique*, qui correspond également à celle de Carcedo, ce qui lui permet de comparer leurs résultats. Ces dernières années les recherches en *disponibilité lexicale* se sont développées, et les travaux appliquant cette même méthodologie se sont multipliés mais la bibliographie reste encore modeste. Toutefois, bien que nous ayons des références en langue espagnole : López González (2014), Hidalgo Gallardo (2017), Aabidi (2019) et Nalesso (2019), nous n'en avons que très peu en français : De la Maya Retamar (2015) et Santos Díaz (2015). Le premier

également réalisé auprès d'étudiants du secondaire et l'autre auprès d'étudiants universitaires mais tout deux rédigés en langue espagnole. Notre travail a donc la spécificité d'être en français.

Cette recherche poursuit différents objectifs. Tout d'abord, celui d'apporter la quantité absolue et relative de réponses et de *vocables*¹ des élèves dans chaque *centre d'intérêt*. Ensuite, celui d'analyser l'incidence de la variable choisie, le niveau scolaire de l'apprenant, autant sur la quantité de *mots* que sur le nombre de *vocables*, dans les différents *centres d'intérêt*. Dans le but de pouvoir observer les relations associatives qui existent, mettre en avant les *vocables* qui occupent les premières positions dans les listes de disponibilité. De la même façon, nous avons décidé de réaliser la même analyse en fonction de la variable choisie, afin de vérifier l'incidence de chacun de ces facteurs. Enfin, il s'agit de décrire du point de vue qualitatif les faits les plus pertinents à partir des résultats quantitatifs.

Dans notre travail nous nous proposons de répondre aux questions de recherche suivantes :

1. La *disponibilité lexicale* des étudiants de FLE est-elle inférieure (autant en quantité de *mots* qu'en quantité de *vocables*) à celle des locuteurs natifs ?
2. Trouve-t-on forcément des coïncidences entre les *centres d'intérêt* qui présentent le plus de *mots* et ceux qui présentent le plus de *vocables* ?
3. Existe-t-il une relation associative importante entre le niveau scolaire et la quantité de *mots* et *vocables* ?
4. Existe-t-il une relation associative entre l'influence de l'espagnol et la quantité de *mots* et *vocables* ?

La collecte du matériel pour ce travail a eu lieu pendant le mois d'avril 2021, à partir d'enquêtes réalisées auprès de 294 élèves de FLE étudiant dans un établissement d'études secondaires espagnol, aux Canaries.

2. Méthodologie

L'objectif principal de cette recherche est de refléter d'un côté, l'état du bagage lexical d'un échantillon d'étudiants en FLE dans l'éducation secondaire publique obligatoire en Espagne à un instant *t*, et de l'autre, l'évolution de celui-ci à travers les années scolaires et de l'évolution du *vocabulaire disponible*. En ce sens, notre enquête se concentre exclusivement sur le *vocabulaire disponible*, dans le but de pouvoir identifier de quel bagage lexical disposent nos étudiants et comment celui-ci évolue pour pouvoir trouver des pistes de planification de l'enseignement du vocabulaire. Quel vocabulaire un élève est capable de mobiliser pour s'exprimer sur des thèmes quotidiens ? Les mots appris sont-ils

¹ Dans le langage courant, *vocable* et *mot* s'utilisent comme des synonymes. En revanche, en linguistique et plus précisément dans le domaine de la statistique lexicale, les mots font référence à la quantité totale d'unités lexicales, les *vocables* aux mots différents.

utiles ou le reflet d'une langue de la classe artificielle au regard des standards de la communication ? Dans quelle mesure ce bagage lexical augmente au fil des années d'apprentissage ? Augmente-t-il d'ailleurs au fil des années ou stagne-t-il au bout d'un certain temps ?

Pour ce faire, nous avons réalisé une enquête auprès de 294 informateurs, soit l'ensemble des apprenants de français de l'établissement, allant de la première année après la primaire (*1^o de la ESO*) jusqu'à la dernière avant l'université (*2^o de Bachillerato*).

2.1. L'échantillon

Comme nous l'avons dit plus haut, notre travail constitue une approche à l'étude de la *disponibilité lexicale* en français de locuteurs non natifs. Ce caractère préliminaire, ainsi que les objectifs qu'il poursuit, justifient la sélection d'un échantillon de 294 enquêtes d'étudiants de FLE, tous les élèves inscrits en français de l'établissement. Nous n'avons retenu qu'une seule variable : celle du niveau scolaire dans la mesure où nous sommes intéressé par l'évolution du bagage lexical pendant la scolarité de nos élèves. Toutefois, dans un souci d'harmonisation des résultats, nous avons pris le soin d'interroger nos apprenants pour savoir s'ils suivaient des cours particuliers de français en dehors de leur cursus. Les enquêtes des étudiants ayant répondu favorablement n'ont pas été prises en compte. Il est également à noter qu'aucun d'entre eux ne présente de besoins éducatifs particuliers qui pourraient altérer les résultats de l'enquête.

Voici la répartition des différents élèves en fonction de leur niveau scolaire :

Niveau : 1^o (102), 2^o (93), 3^o (66), 4^o (25), Bach (8)

- a. *1^o de la ESO* (5^{ème}) : 102 élèves
- b. *2^o de la ESO* (4^{ème}) : 93 élèves
- c. *3^o de la ESO* (3^{ème}) : 66 élèves
- d. *4^o de la ESO* (2^{nde}) : 25 élèves
- e. *Bachillerato* (1^{ère} et terminale) : 8 élèves

Le choix de ces facteurs correspond aux intérêts spécifiques que comporte la recherche en *disponibilité lexicale* d'une langue qui n'est pas la langue maternelle. En outre, le niveau de connaissance est fondamental lorsque l'on analyse les résultats, puisque l'hypothèse de départ est qu'une meilleure maîtrise du français implique une disponibilité plus riche.

La sélection des informateurs s'est faite de façon plutôt arbitraire puisque nous avons interrogés tous les élèves de l'établissement. En ce qui concerne le niveau des élèves la distribution correspond à la réalité du secondaire. En Espagne la deuxième langue est obligatoire jusqu'en troisième année (*3^o ESO*), à partir de la quatrième année les élèves sont moins nombreux car ils ont plus de choix d'options

(obligatoires comme en seconde) et s'orientent en fonction de leur projet professionnel. Après la quatrième année, les élèves de *Bachillerato* ont la possibilité de poursuivre le français mais cette fois comme réelle option qui leur apportera un bonus à l'examen qui précède l'entrée à l'université, s'ils ont une note supérieure à la moyenne. Les langues sont souvent peu choisies.

2.2. L'enquête

Le corpus de notre étude se compose de 294 questionnaires, un par informateur, collectés au mois d'avril 2021. Ils sont de type *ouvert* et reprennent 8 des 16 *centres d'intérêt*, utilisés dans l'élaboration du *Français Fondamental* et repris dans l'essentiel des travaux sur le *vocabulaire disponible*. Ces *centres d'intérêt* sont d'ailleurs utilisés dans le *Projet Panhispanique*. Les enquêtes se sont déroulées en classe sur les heures de cours normales, nous avons distribué à chaque élève une feuille de questionnaires correspondant à un tableau reprenant les principales informations pour pouvoir classer les enquêtes en fonction de notre variable. Sur l'autre partie de la feuille apparaissent 8 cases à remplir représentant les 8 *centres d'intérêt* sélectionnés pour l'étude. Pour chaque case est inscrit le titre de la catégorie considéré comme le stimulus. C'est à partir de ce stimulus que les réponses spontanées des apprenants, obtenues à travers ce qu'on appelle des *associations d'idées* en psycholinguistique, vont pouvoir s'inscrire dans les différentes cases dans le temps imparti (2 minutes par catégorie). Avant le déroulement de « l'épreuve » toutes les considérations méthodologiques ont été communiquées aux apprenants. En effet, nous leur avons expliqué quel était l'objectif de ces questionnaires et qu'ils ne seraient aucunement pris en compte dans leurs notes de trimestre. Par conséquent, nous leur avons demandé de participer honnêtement en leur expliquant que le copiage fausserait les résultats et qu'ils ne se sentent pas coupable si aucun mot n'apparaissait dans une ou plusieurs catégories, qu'ils étaient justement là pour apprendre. Pour les niveaux débutants, il a été facile de rassurer les élèves en leur indiquant que toutes les enquêtes étaient les mêmes pour tous les niveaux et que certains thèmes étaient abordés plus tard. Pour les niveaux plus avancés, la différence entre *vocabulaire disponible* et *vocabulaire actif* a été clairement expliquée de façon à ne pas provoquer de frustration si une case restait vide. Comme nous l'avons expliqué dans la première partie, le *vocabulaire disponible* est parfois le reflet imprécis du vocabulaire que l'on utiliserait dans une situation réelle de communication, puisque parfois, le stimulus et le temps imposé ne permettent pas de restituer un mot que nous connaissons ou au contraire ce type d'épreuve peut donner lieu à l'apparition de mots en langue maternelle que l'informateur n'utiliserait probablement pas dans un autre contexte. Les détails sur l'anonymat ont également été évoqués, bien que les enquêtes restent anonymes nous leur avons

demandé d'indiquer leurs noms pour des raisons pratiques : pour pouvoir s'adresser à eux en cas de problème de lecture, d'interprétation d'un mot...

Nous avons organisé l'épreuve de manière à ce que chaque élève traite la même catégorie pendant deux minutes en surveillant le temps à l'aide d'un chronomètre. Le but étant d'écrire le plus de mots possibles, en relation avec le thème, dans le temps imparti, il est possible d'écrire au dos de la feuille s'il n'y a pas suffisamment d'espace dans une des cases, en revanche il n'est pas possible de revenir sur une catégorie déjà remplie. Par ailleurs, la distinction entre langue écrite et la langue orale n'étant pas prise en compte, les informateurs ont été invités à inscrire tous les mots qu'ils connaissaient même s'ils n'étaient pas sûrs de l'orthographe.

Pour des raisons pratiques l'enquête distribuée à chaque élève a dû se concentrer sur une seule page. Les 8 *centres d'intérêt* ont donc été repartis sous forme de tableau soit 8 cases dont le titre de chaque *centre d'intérêt* (C. I.) apparaissait en haut. Nous n'avons pas choisi de numéroter les lignes, ni de les définir pour ne pas rendre le questionnaire illisible, mais jamais il n'a été question de ne pas prendre en compte la position d'apparition de chaque mot car celle-ci est essentielle à la réalisation des calculs. Étant donné que certains élèves écrivent en colonnes d'autre en lignes il a fallu mettre des stratégies en place pour pouvoir définir la position d'apparition de chaque mot. En principe, l'organisation en colonnes avec des lignes numérotées facilitent la prise en compte de la position du mot dans l'ensemble de la liste de mots apportés par l'informateur, aspect qui, pondéré à la fréquence, s'avère fondamental pour la classification selon la disponibilité. Par rapport à la quantité d'unités qui peuvent être obtenues dans chaque cas, nous avons choisi le système de listes ouvertes, ainsi la limite est seulement d'ordre temporelle : 2 minutes par catégorie.

En ce qui concerne les *centres d'intérêt* choisis, nous avons sélectionnés les 8 premiers des 16 *centres d'intérêt* utilisés par Gougenheim et al. (1964) :

Les parties du corps ;

- (01) Les vêtements (d'homme ou femme) ;
- (02) La maison (mais pas les meubles) ;
- (03) Les meubles de la maison ;
- (04) Les aliments et boissons des repas ;
- (05) Les objets placés sur la table pour les repas ;
- (06) La cuisine, ses meubles et ustensiles ;
- (07) L'école, ses meubles et son matériel ;
- (08) Le chauffage et l'éclairage ;
- (09) La ville ;

- (10) Le village ou le bourg ;
- (11) Les moyens de transport ;
- (12) Les travaux des champs et le jardinage ;
- (13) Les animaux ;
- (14) Les jeux et distractions ;
- (15) Les métiers.

Comme en témoignent les programmes officiels ces 8 premiers thèmes² représentent ceux de la vie quotidienne qui sont en général ceux abordés dans les premières années d'apprentissage. Notre sélection est de surcroît motivée par notre objectif qui poursuit l'obtention du lexique utilisé normalement par les locuteurs d'une communauté linguistique donnée.

2.3. Traitement des données

Après avoir terminé nos enquêtes, il a fallu harmoniser les données pour que celles-ci puissent être numérisées. Dans les études de disponibilité, cette étape est la plus complexe, car il est nécessaire non seulement d'harmoniser les enquêtes, mais aussi de résoudre certains problèmes comme celui de la pertinence des associations ou l'unification de différentes variantes de certaines unités lexicales. Notre étude présente encore plus de difficultés, puisque, comme il s'agit d'informateurs espagnols, elle comporte des phénomènes tels que des interférences linguistiques ou la ressemblance de certaines unités avec des mots français. La prise de décisions face à ces problèmes, qu'ils soient communs à tous types de travaux de disponibilité, ou qu'ils soient spécifiques à ceux en langue non maternelle, requiert un processus considérable de réflexion de la part du spécialiste, puisque ces décisions se reflètent postérieurement dans les résultats quantitatifs ainsi que dans comparaisons des résultats avec d'autres sources. Pour toutes ces raisons, il est nécessaire, de suivre dans la mesure du possible des critères d'édition communs, prenant en compte certaines décisions approuvées majoritairement dans les recherches antérieures. Étant donné que A. Carcedo, premier chercheur à s'être penché sur la question, comme nous l'avons mentionné dans la première partie, n'expose pas ses critères d'édition, nous avons utilisé l'article de José Antonio Samper Padilla (1998) dont l'objectif principal est précisément d'offrir des principes qui facilitent la comparaison entre variétés.

Pour le calcul des données nous avons utilisé le programme informatique Lexidisp (version 2.1), crée par Moreno Fernández et Garcia de la Heras, aidés par les linguistes : Moreno Fernandez et

² Nous utiliserons dès maintenant et dans tous les tableaux suivants les abréviations pour les *centres d'intérêt*. Voici l'équivalent de chaque abréviation : 01. Les parties du corps : COR ; 02. Les vêtements : VET ; 03. Les parties de la maison : MAI ; 04. Les meubles de la maison : MEU ; 05. Nourriture et boissons : ALI ; 06. Objets placés sur la table pour le repas : OBJ ; 07. La cuisine et ses ustensiles : CUI ; 08. L'École : ECO.

Bénitez Pérez. Ce programme parrainé par la *ALFAL*³, l’institut Cervantes et l’Université d’Alcala, et applique la formule développée par López Chavez et Strassburger (1987). Grâce à ce programme, nous pouvons connaître de façon immédiate l’apparition, la fréquence en fonction du total des mots, la fréquence accumulée (somme des relatives) et, évidemment, la disponibilité de chacune des unités lexicales qui apparaissent dans nos listes. Par ailleurs, ce programme permet de réaliser une analyse comparative de l’ensemble de ces données.

3. Résultats

La quantité totale de mots obtenue s’élève à 8548 ce qui offre une moyenne de 29.07 réponses par élève et de 1068.5 par *C. I.* Dans le Tableau 1, ci-dessous, apparait le rang (ou position relative) qu’occupe les champs lexicaux en fonction du nombre total de réponses (mots) et indique également la moyenne par sujet dans chacun d’eux. Comme nous pouvons observer, la moyenne de mots par élève, résultat de la division du nombre total de réponses par le nombre total d’informateurs, est bien inférieure aux résultats obtenus normalement dans les enquêtes menées en langue maternelle. Ce fait dérive, évidemment, de la moins bonne compétence lexicale de nos élèves, qui ne connaissent le français que comme langue étrangère. Ainsi, alors que dans les autres enquêtes les résultats atteignent presque toujours 20 unités par *C. I.*, dans notre cas la moyenne est de 3,63 mots par élève.

Tableau 1.
Nombre de mots et position de chaque *centre d’intérêt*

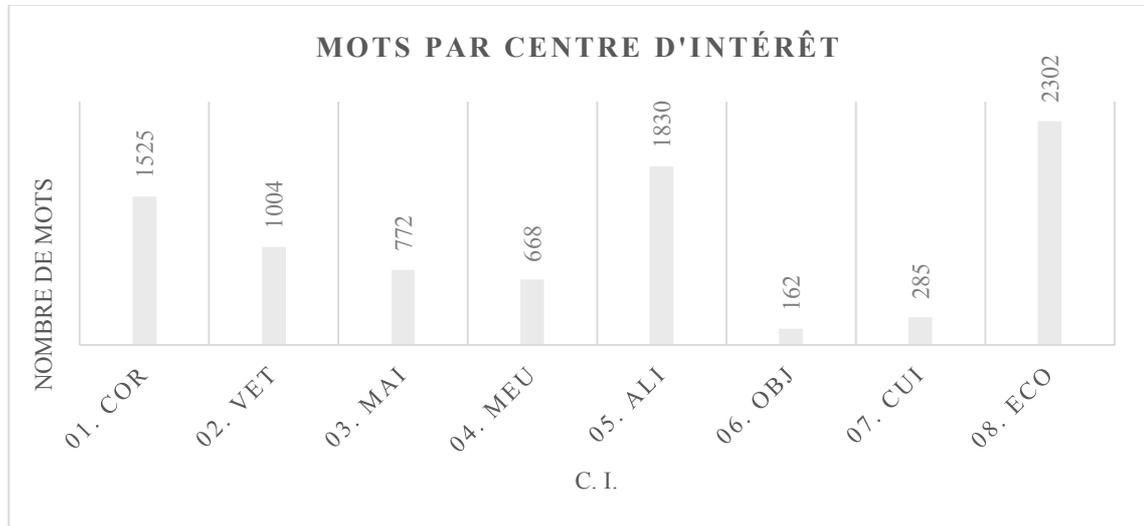
Rang	C.I	Mots	Moyenne de mot par élève
3	01. COR	1525	5,19
4	02. VET	1004	3,41
5	03. MAI	772	2,63
6	04. MEU	668	2,27
2	05. ALI	1830	6,22
8	06. OBJ	162	0,55
7	07. CUI	285	0,97
1	08. ECO	2302	7,83

Le *C. I.* le plus productif est le 08 « l’école », suivi de « les aliments et boissons des repas » (05) et de « les parties du corps » (01). Les deux derniers sont « la cuisine, ses meubles et ustensiles »

³ Asociación de Lingüística y Filología de América Latina.

(07) et « les objets placés sur la table pour le repas » (06) avec une moyenne très faible puisqu'inférieure à 1.

La différence entre les différents champs s'exprime dans le Graphique 1 :



Graphique 1.

Nombre de mots par *centre d'intérêt*

À ce stade, nous présenterons les 20 unités lexicales les plus disponibles (cf. Tableau 2) :

Tableau 2.

Liste des 20 mots les plus disponibles

Rang	Mots	Indice de disponibilité	C. I.	Rang	Mots	Indice de disponibilité	C. I.
1	Tête	0,68608	01. COR	11	Croissant	0,34207	05. OBJ
2	Pied	0,57277	01. COR	12	Livre	0,33029	08. ECO
3	Table	0,55727	04. MEU	13	Stylo	0,32286	08. ECO
4	Cuisine	0,51899	03. MAI	14	Télé(vision)	0,30699	04. MEU
5	Main	0,50788	01. COR	15	Poisson	0,29882	05. ALI
6	T-Shirt	0,50627	02. VET	16	Eau	0,28616	05. ALI
7	Pantalon	0,50340	02. VET	17	Salon	0,27980	03. MAI
8	Bras	0,45009	01. COR	18	Table	0,26357	08. ECO
9	Yeux	0,35343	01. COR	19	Coca-cola	0,25203	05. ALI
10	Toilettes	0,34597	03. MAI	20	Crayon	0,25097	08. ECO

Comme nous pouvons l'observer le mot « table » apparaît deux fois. C'est effectivement le mot le plus disponible du C. I. 04 « les meubles de la maison » et un des plus disponibles du 08 « l'école ». Nous distinguons les mots qui sont les mêmes en espagnol que ce soit au point de vue

phonétique ou au point de vue orthographique : « pied », « pantalon », « télé(vision) » et « salon ». Le mot « Coca-cola » arrive tout de même en 19^{ème} position.

Le *C. I.* le plus représenté est « les parties du corps » (01) avec 5 mots, suivi à égalité par « les aliments et boissons des repas » (05) et par « l'école, ses meubles et son matériel » (08) avec 4 mots. Ensuite, « les parties de la maison » avec 3 mots et enfin « les meubles de la maison » (04) ainsi que « les vêtements » (02) avec 2 mots chacun. Sans surprise, les deux autres *C. I.* (« objets placés sur la table pour le repas » 06 et « la cuisine et ses ustensiles » 07) n'apparaissent pas dans ce tableau.

Une fois l'ensemble des résultats de l'étude exposé, il est très intéressant d'observer la façon dont le niveau scolaire influence les résultats. Nous aborderons ainsi autant les moyennes générales de réponses (mots) et de *vocables*, que les particularités de chaque *C. I.*, selon les différents facteurs. Nous avons opté pour l'étude par moyenne parce que nous pensons que c'est la méthode la plus fiable pour la confrontation des résultats dans le type de recherche qui, comme c'est notre cas, présente des différences numériques dans la répartition de notre variable. Ces différents résultats sont illustrés plus bas dans le Tableau 3 et le Graphique 2.

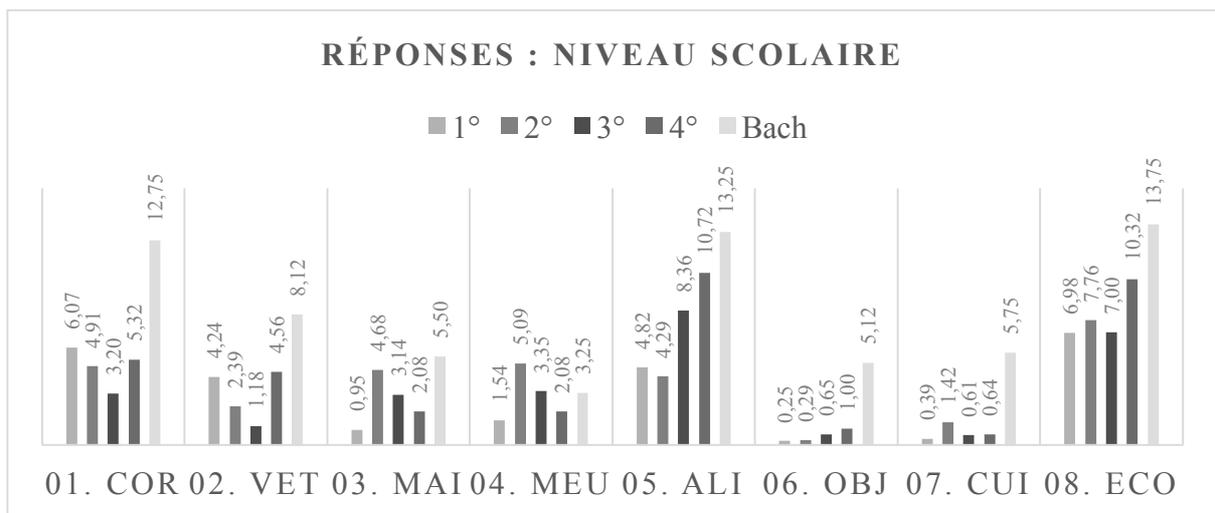
Le niveau scolaire est un facteur très intéressant car il reflète deux choses, d'un côté, l'état du bagage lexical à un moment *t* et de l'autre, l'évolution de celui-ci à travers les années scolaires. En principe plus l'élève passe de niveau, plus la moyenne de mots proposés devrait être importante. Cependant ce n'est pas le cas, les résultats de 3^o sont anormalement bas, en effet, dans les deux premiers *C. I.* les élèves de 1^o ont une moyenne de mots plus élevée que ceux de 3^o atteignant dans « les parties du corps » (01) presque le double (plus élevée également que les élèves de 4^o) et dans « les vêtements » (02) presque 4 fois plus, rejoignant presque les résultats des élèves de 4^o. Si on compare les niveaux de 1^o et 3^o dans le *C. I.* « l'école, ses meubles et son matériel » (08) on constate que les résultats sont quasi identiques. Les élèves de 2^o eux dépassent les résultats des 3^o dans 6 des *C. I.* Ensuite, l'évolution est normale à partir de 3^o à l'exception du *C. I.* 04 « les meubles de la maison » dans lequel les élèves de 2^o ont plus de vocabulaire que les élèves de *Bachillerato*. Cette évolution se reflète dans le Graphique 3 ci-dessous.

Ces résultats s'expliquent peut-être par les différents programmes scolaires. En effet, en 1^o on aborde « le matériel scolaire », « les parties du corps » et « les vêtements », en 2^o « les matières scolaires », « la maison » et « les meubles », en 3^o « l'alimentation ». Cet élément de réponse pose une question essentielle dans l'apprentissage du vocabulaire, non pas sur la quantité mais sur la qualité de l'apprentissage et sa durée dans le temps. Effectivement, cela signifie peut-être qu'on apprend pour le court terme et pas le long terme, qu'en cours d'apprentissage n'est pas proposé de traitement de ces

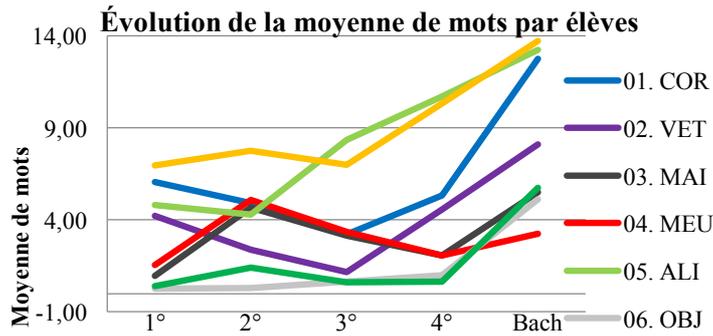
informations qui permettrait leur stockage dans la mémoire à long terme. Ce qui expliquerait la forte perte.

Tableau 3.
Moyenne de mots en fonction du niveau scolaire

C. I.	Rang	1°	Rang	2°	Rang	3°	Rang	4°	Rang	Bach
01. COR	2	6,07	3	4,91	4	3,20	3	5,32	3	12,75
02. VET	4	4,24	6	2,39	6	1,18	4	4,56	4	8,12
03. MAI	5	0,95	4	4,68	5	3,14	5	2,08	5	5,50
04. MEU	6	1,54	2	5,09	3	3,35	5	2,08	8	3,25
05. ALI	3	4,82	5	4,29	1	8,36	1	10,72	2	13,25
06. OBJ	8	0,25	8	0,29	7	0,65	6	1,00	7	5,12
07. CUI	7	0,39	7	1,42	8	0,61	7	0,64	6	5,75
08. ECO	1	6,98	1	7,76	2	7,00	2	10,32	1	13,75



Graphique 2.
Moyenne de mots en fonction du niveau scolaire



Graphique 3.

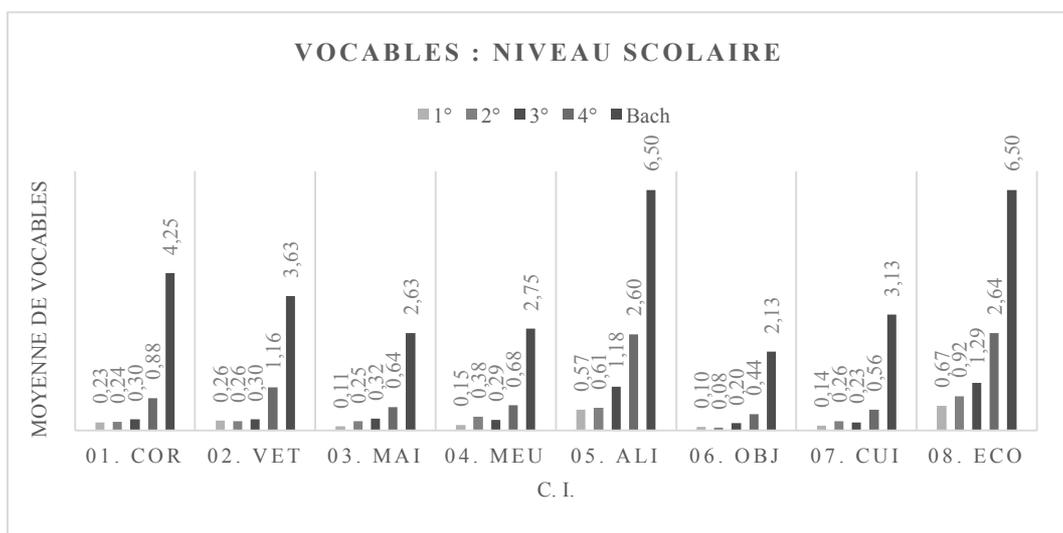
Évolution de la moyenne de mots par élèves

Comme nous l'avons vu précédemment, l'incidence de cette dernière variable confirmait ce que nous pensions, les élèves de *Bachillerato* apportent plus de réponses que ceux de 1°, mais ce qui nous surprenait le plus c'était que l'évolution du bagage lexical n'est pas toujours croissante proportionnellement à l'avancée en termes de niveau. Lorsque l'on compare la moyenne de *vocables* par niveau scolaire cette évolution est croissante, (à l'exception du C. I. 04 et du 07 dans lesquels les élèves de 2° ont apporté légèrement plus de *vocables* que ceux de 3°). Cela veut dire qu'à mesure que les élèves avancent dans leur scolarité ils diversifient leurs réponses, ils apportent parfois moins de mots que l'année précédente mais offrent plus de mots différents.

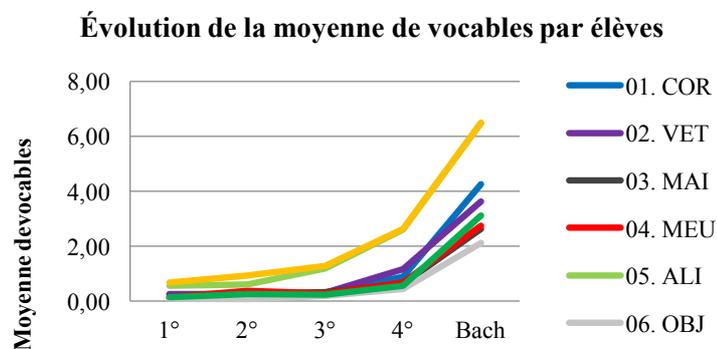
Il est également important de souligner le fait que les élèves de *Bachillerato* et de 4° ont une moyenne de *vocables* par élève beaucoup plus élevée que les autres élèves (cf. Tableau 4 et Graphique 4). Ceux de *Bachillerato* ont à leur tour une moyenne plus élevée que ceux de 4°. Ces résultats indiquent une richesse du vocabulaire croissante. Deux grands fossés caractéristiques de deux étapes importantes se dessinent. D'une part, le passage en 4° où le français devient optionnel donc les élèves qui restent sont en général les plus motivés et les meilleurs puisqu'on oriente les élèves en fonction de leurs résultats. Et d'autre part, le passage en *Bachillerato* où il y a encore moins d'élèves puisque le français passe au rang de « bonus » puisqu'il ne fait plus partie des « options obligatoires ». Lorsque l'on regarde les résultats (cf. Graphique 5) ces fossés se définissent clairement par les écarts importants dans la production lexicale.

Tableau 4.
Moyenne de *vocables* en fonction du niveau scolaire

C. I.	Rang	1°	Rang	2°	Rang	3°	Rang	4°	Rang	Bach
01. COR	4	0,23	6	0,24	4	0,30	4	0,88	2	4,25
02. VET	3	0,26	4	0,26	4	0,30	3	1,16	3	3,63
03. MAI	7	0,11	5	0,25	3	0,32	6	0,64	6	2,63
04. MEU	5	0,15	3	0,38	5	0,29	5	0,68	5	2,75
05. ALI	2	0,57	2	0,61	2	1,18	2	2,60	1	6,50
06. OBJ	8	0,10	7	0,08	7	0,20	8	0,44	7	2,13
07. CUI	6	0,14	4	0,26	6	0,23	7	0,56	4	3,13
08. ECO	1	0,67	1	0,92	1	1,29	1	2,64	1	6,50



Graphique 4.
Nombre de *vocables* en fonction du niveau scolaire



Graphique 5.
Évolution de la moyenne de vocables par élèves

Pour l'analyse qualitative de nos résultats nous allons nous intéresser à tous les *C. I.*

Comme nous pouvons observer dans le Tableau 5, dans le premier *C. I.*, « les parties du corps » (01), 4 *vocables* coïncident : « pied », « tête », « main » et « cheveux ».

Tableau 5.

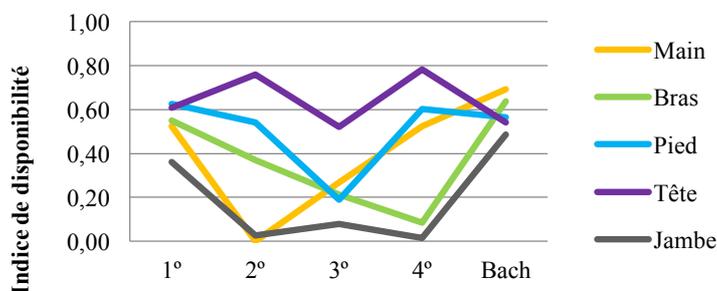
Premiers *vocables* disponibles en fonction du facteur niveau scolaire. *C. I 01*

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Pied	0,62	Tête	0,76	Tête	0,52	Tête	0,78	Main	0,69
2	Tête	0,61	Pied	0,54	Yeux	0,33	Pied	0,60	Bras	0,64
3	Bras	0,55	Main	0,42	Main	0,27	Main	0,52	Pied	0,56
4	Main	0,52	Bras	0,37	Bras	0,21	Yeux	0,21	Tête	0,54
5	Jambe	0,36	Yeux	0,24	Pied	0,19	Dos	0,18	Jambe	0,49
6	Yeux	0,31	Oreille	0,15	Cheveux	0,15	Oreille	0,18	Bouche	0,41
7	Cheveux	0,19	Cou	0,12	Nez	0,09	Cheveux	0,17	Oreille	0,39
8	Cou	0,14	Épaule	0,12	Doigt	0,09	Bouche	0,13	Cheveux	0,35
9	Oreille	0,13	Bouche	0,12	Jambe	0,08	Dent	0,13	Dos	0,35
10	Nez	0,09	Cheveux	0,08	Bouche	0,07	Nez	0,10	Doigt	0,32

Ces trois premiers mots font tous référence à des extrémités et sont tous cités avant « cheveux ». Ensuite les mots se recoupant sont essentiellement constitués de parties du visage comme « bouche », « yeux » ou encore « oreille ». Ce sont d'ailleurs ceux qui sont enseignés en premier. À l'exception de « épaule » les *vocables* se recourent amplement, mais malgré la grande homogénéité de ces réponses les résultats des 3^o surprennent par leur pauvreté.

Dans le Graphique 6, nous pouvons voir l'évolution des 5 mots les plus disponibles des élèves dans la dernière étape du cycle secondaire. Il est curieux de constater que les résultats en dernière année sont très proches de ceux de la première année et sont soumis à une évolution inattendue.

Évolution de 5 unités lexicales de 01. COR



Graphique 6.

Évolution de 5 *vocables* en fonction du facteur niveau scolaire. *C. I. 01*

Le vocabulaire associé à ce *C. I.*, au programme de *1^o*, présente une moyenne de mots par élève décroissante jusqu'en *3^o*, ensuite en *4^o* elle est croissante mais l'indice reste inférieur à celui de *1^o* puis en dernière année cette évolution continue et la moyenne finale est bien supérieure puisqu'elle est multipliée par plus de deux.

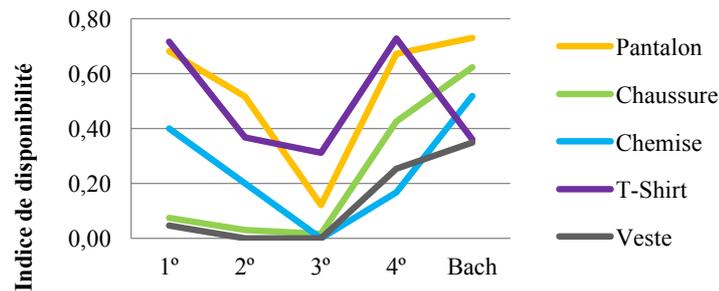
Comme dans le *C. I.* précédent, pour « les vêtements » (02) nous avons 4 *vocables* communs : « t-shirt », « pantalon », « pull », « tennis ». D'ailleurs, à l'exception de *4^o*, « t-shirt » et « pantalon » et « tennis » apparaissent avant « pull » et jusqu'à *4^o* inclus « t-shirt » et « pantalon » arrivent en premier (cf. Tableau 6).

Tableau 6.
Premiers *vocables* disponibles en fonction du facteur niveau scolaire. *C. I* 02

Rang	1 ^o	I. D.	2 ^o	I. D.	3 ^o	I. D.	4 ^o	I. D.	BACH	I. D.
1	T-Shirt	0,72	Pantalon	0,52	T-Shirt	0,31	T-Shirt	0,73	Pantalon	0,73
2	Pantalon	0,68	T-Shirt	0,37	Pantalon	0,12	Pantalon	0,67	Chaussure	0,62
3	Chemise	0,40	Casquette	0,25	Pull	0,11	Chaussure	0,42	Chemise	0,52
4	Baskets	0,36	Chemise	0,20	Jean	0,09	Veste	0,25	T-Shirt	0,36
5	Pull	0,20	Pull	0,13	Tennis	0,09	Tennis	0,25	Veste	0,35
6	Tennis	0,15	Tennis	0,09	Bouton	0,04	Jean	0,22	Pull	0,32
7	Jupe	0,14	Jean	0,08	Blouse	0,04	Pull	0,17	Chaussette	0,18
8	Jean	0,12	Baskets	0,07	Mode	0,04	Chemise	0,17	Jupe	0,16
9	Casquette	0,11	Lunettes	0,07	Botte	0,03	Lunettes	0,06	Robe	0,14
10	Robe	0,10	Jupe	0,05	Lunettes	0,03	Maillot de bain	0,06	Tennis	0,13

Nous avons essentiellement des vêtements de la garde-robe générale et mis à part « jupe », ils sont tous mixtes. On peut également constater l'évolution décroissante du bagage lexical de *1^o* à *3^o* et encore une fois des résultats faibles en *3^o* et 4 occurrences particulières, dont deux secondaires (« mode » et « bouton »). L'évolution décroissante en premier lieu (de *1^o* à *3^o*) est illustrée dans le Graphique 7, sauf pour « t-shirt » la croissance normale de l'indisponibilité s'observe à partir de *4^o*.

Évolution de 5 unités lexicales de 02. VET



Graphique 7.

Évolution de 5 *vocables* en fonction du facteur niveau scolaire. C. I. 02

Le vocabulaire des vêtements (02), également au programme de 1°, présente la même évolution que précédemment, sauf pour les élèves de 4° qui ont une moyenne de mots un peu plus élevée que ceux de première année.

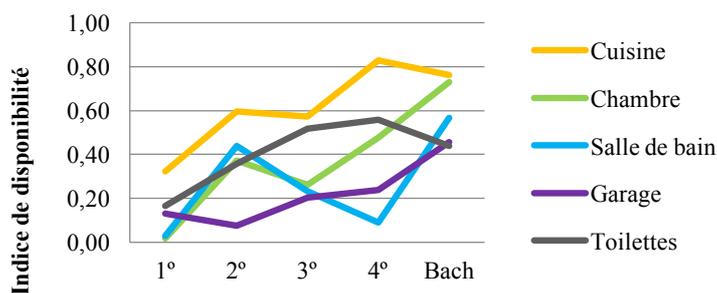
Pour le C. I. suivant, les parties de la maison (03), 5 *vocables* sont identiques : « cuisine », « toilettes », « salon », « salle de bain », et « chambre », par ailleurs « cuisine » arrive en premier place dans chacun des groupes. Ces mots représentent les principales pièces d’habitation d’une maison. Ici les indices de disponibilité évoluent d’une manière plus attendue puisqu’ils sont faibles en 1°, puis ils augmentent au fur et à mesure (cf. Tableau 7).

Tableau 7.

Premiers *vocables* disponibles en fonction du facteur niveau scolaire. C. I 03

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Cuisine	0,32	Cuisine	0,59	Cuisine	0,57	Cuisine	0,83	Cuisine	0,76
2	Toilettes	0,17	Salon	0,45	Toilettes	0,52	Toilettes	0,56	Chambre	0,73
3	Garage	0,13	Salle de bain	0,44	Salon	0,37	Chambre	0,48	Salle de bain	0,57
4	Jardin	0,08	Chambre	0,37	Chambre	0,26	Salon	0,37	Garage	0,45
5	Salon	0,06	Toilettes	0,36	Salle de bain	0,23	Garage	0,24	Toilettes	0,44
6	Hall	0,03	Terrasse	0,15	Garage	0,20	Jardin	0,16	Jardin	0,30
7	Salle de bain	0,03	Entrée	0,15	Terrasse	0,08	Salle de bain	0,09	Couloir	0,23
8	Porte	0,02	Bureau	0,13	Escalier	0,07	Hall	0,05	Salle à manger	0,22
9	Chambre	0,02	Escalier	0,11	Jardin	0,07	Piscine	0,05	Salle	0,18
10	Balcon	0,01	Salle à manger	0,08	Hall	0,06	En haut	0,03	Salon	0,17

Évolution de 5 unités lexicales de 03. MAI



Graphique 8.

Évolution de 5 vocables en fonction du facteur niveau scolaire. C. I. 03

Comme le reflète le Graphique 8, ci-dessus, en ce qui concerne le vocabulaire de ce C. I., au programme de 2°, l'évolution est croissante jusqu'en deuxième année, puis est décroissante jusqu'en 4° (mais sans toutefois descendre en dessous des résultats de 1°), puis reprend à la hausse en *Bachillerato*, mais les résultats y sont à peine plus élevés que ceux de 3°.

Dans le C. I., les meubles (04), 3 *vocables* coïncident : « table », « télé(vision) » et « chaise ».

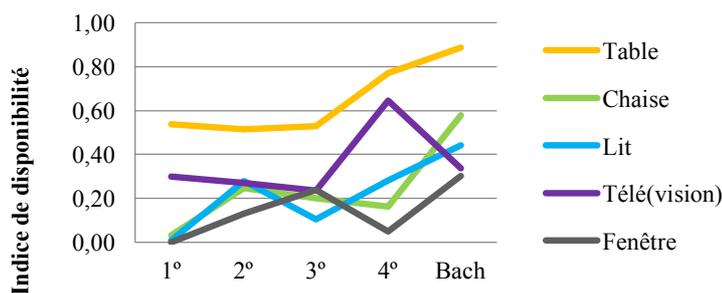
Tableau 8.

Premiers *vocables* disponibles en fonction du facteur niveau scolaire. C. I. 04

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Table	0,54	Table	0,51	Table	0,53	Table	0,77	Table	0,89
2	Télé(vision)	0,30	Lit	0,28	Bureau	0,27	Télé(vision)	0,64	Chaise	0,58
3	Sofa	0,16	Lampe	0,27	Sofa	0,27	Lit	0,28	Lit	0,44
4	Ordinateur	0,10	Télé(vision)	0,27	Fenêtre	0,24	Sofa	0,25	Télé(vision)	0,34
5	Tableau	0,05	Chaise	0,25	Télé(vision)	0,24	Chaise	0,16	Fenêtre	0,30
6	Armoire	0,04	Sofa	0,22	Chaise	0,20	Ordinateur	0,13	Porte	0,18
7	Chaise	0,03	Canapé	0,19	Armoire	0,20	Tableau	0,08	Frigo(rifique)	0,18
8	Toilettes	0,02	Bureau	0,16	Porte	0,12	Douche	0,06	Lampe	0,16
9	Lampe	0,02	Ordinateur	0,16	Lampe	0,11	Fenêtre	0,05	Ordinateur	0,13
10	Téléphone	0,01	Étagère	0,15	Lit	0,11	Porte	0,04	Radio	0,13

« Table » apparaît en tête de chaque groupe mais, à part dans le dernier groupe, « chaise » est toujours éloigné dans les réponses alors qu'il est naturel d'envisager cette relation associative. D'autre part, en observant le Tableau 8, nous pouvons remarquer les résultats assez bas de 1° et le fait que nous n'ayons pas exclusivement des meubles mais aussi de l'électroménager, des objets et des éléments comme la « porte » et la « fenêtre » etc. Une question surgit inévitablement, ces réponses correspondent-elles à une conception plus ample de la définition de « meuble » ou simplement à un besoin de performance qui consisterait à écrire le plus de mot possible indépendamment du thème traité ?

Évolution de 5 unités lexicales de 04. MEU



Graphique 9.

Évolution de 5 *vocables* en fonction du facteur niveau scolaire. C. I. 04

Pour « les meubles de la maison » (04), aussi au programme de 2°, l'évolution est similaire à celle du C. I. précédent, sauf pour les élèves de *Bachillerato* qui ont une moyenne de mots par élève inférieure à ceux de 2° (cf. Graphique 9).

Si on s'intéresse au vocabulaire de la nourriture et aux boissons (05), nous nous trouvons face à un C. I. qui, avec celui de « l'école, ses meubles et son matériel » (08), offre la moyenne de *vocables* par élève de *Bachillerato* la plus élevée. Ces résultats se matérialisent dans le Tableau 9, ci-dessous.

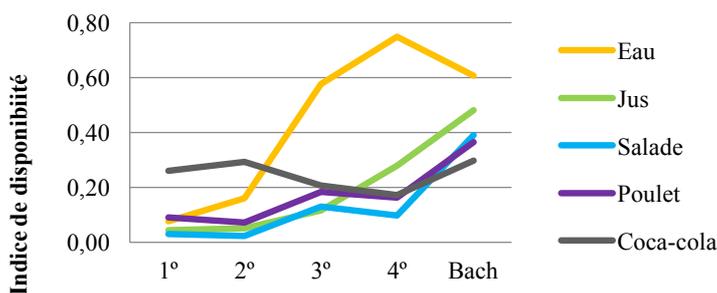
Tableau 9.

Premiers *vocables* disponibles en fonction du facteur niveau scolaire. C. I. 05

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Croissant	0,55	Poisson	0,33	Eau	0,58	Eau	0,75	Eau	0,61
2	Baguette	0,45	Spaghetti	0,33	Poisson	0,43	Poisson	0,67	Jus	0,48
3	Coca-cola	0,26	Coca-cola	0,29	Fromage	0,32	Croissant	0,39	Salade	0,39
4	Chocolat	0,23	Pizza	0,28	Spaghetti	0,25	Lait	0,32	Poulet	0,36
5	Pizza	0,22	Croissant	0,22	Croissant	0,22	Viande	0,29	Coca-cola	0,30
6	Jambon	0,15	Baguette	0,19	Pizza	0,21	Jus	0,28	Soupe	0,29
7	Tomate	0,14	Eau	0,16	Coca-cola	0,21	Fromage	0,28	Pomme de terre	0,26
8	Spaghetti	0,13	Chocolat	0,13	Pain	0,20	Crêpe	0,21	Orange	0,25
9	Sandwich	0,12	Orange	0,12	Poulet	0,18	Orange	0,20	Pain	0,23
10	Orange	0,11	Fanta	0,11	Tomate	0,17	Coca-cola	0,17	Riz	0,21

Il est intéressant de constater que le seul mot qui coïncide est « Coca-cola ». Ensuite les mots qui se recoupent le plus sont « croissant », « orange » et « eau ». Les deux mots différents pour les élèves de 1° sont « jambon » et « sandwich », quant aux élèves de 2° ils se distinguent avec une marque de soda : « Fanta ». Tous les *vocables* apportés par les élèves de 3° coïncident avec les autres. Ceux de 4° amènent trois nouveautés : « lait », « viande » et « crêpe » alors que ceux de *Bachillerato* proposent « salade », « soupe » et « pomme de terre ».

Évolution de 5 unités lexicales de 05. ALI



Graphique 10.

Évolution de 5 *vocables* en fonction du facteur niveau scolaire. C. I. 05

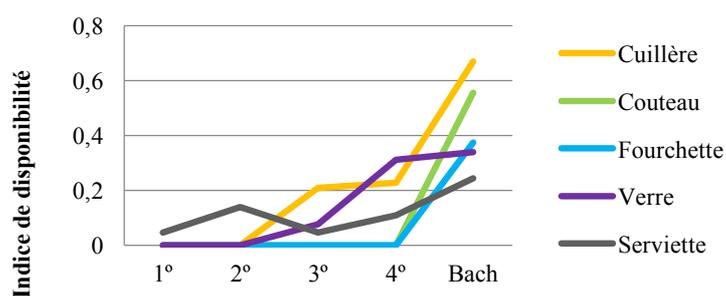
Pour ce vocabulaire, au programme de 3°, l'évolution est légèrement décroissante de 1° à 2° puis croît jusqu'à la dernière année (cf. Graphique 10).

Pour le C. I. suivant, les objets placés sur la table pour les repas (06), la moyenne de *vocables* par élève de *Bachillerato* est la plus basse. Nous sommes ici en présence d'un centre diffus et les coïncidences sont bien moindres que dans les autres C. I. Malgré tout nous avons une corrélation intéressante avec « serviette » et « eau ». Il faut attendre 2° pour qu'apparaissent les « couverts », en 3° apparaissent la « cuillère » et le « verre » et c'est réellement en *Bachillerato* que l'on trouve les mots qu'on pourrait attendre puisque les objets placés sur la table pour le repas sont avant tout les couverts (cf. Tableau 10).

Tableau 10.

Premiers *vocables* disponibles en fonction du facteur niveau scolaire. C. I. 06

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Baguette	0,07	Serviette	0,14	Cuillère	0,21	Verre	0,31	Cuillère	0,67
2	Vin	0,01	Eau	0,04	Eau	0,11	Cuillère	0,23	Couteau	0,55
3	Serviette	0,05	Pain	0,03	Verre	0,08	Serviette	0,11	Fourchette	0,38
4	Ciseaux	0,02	Baguette	0,02	Serviette	0,05	Eau	0,04	Verre	0,34
5	Ordinateur	0,02	Fruit	0,01	Bouteille	0,03	Couverts	0,04	Serviette	0,24
6	Stylo	0,01	Couverts	0,01	Ordinateur	0,03	Fruit	0,04	Petite cuillère	0,22
7	Eau	0,01	Lait	0,01	Pain	0,02	Pain	0,04	Assiette	0,19
8	Fruit	0,02			Fleur	0,02	Papier pour la cuisine	0,04	Jus	0,16
9	Bouteille	0,01			Lunettes	0,02	Tasse	0,04	Eau	0,13
10	Pain	0,03			Baguette	0,01	Dessert	0,04	Feuille	0,11

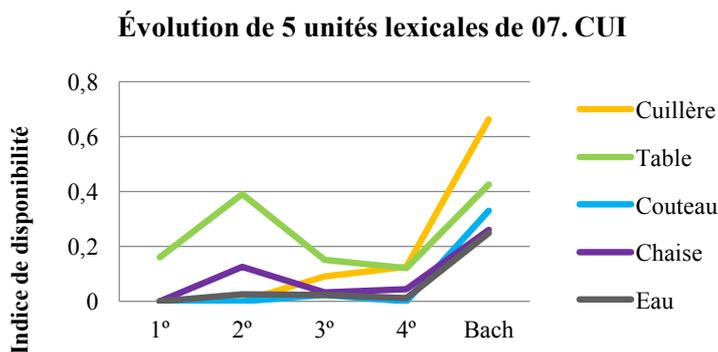
Évolution de 5 unités lexicales de 06. OBJ

Graphique 11.
Évolution de 5 *vocables* en fonction du facteur niveau scolaire. C. I. 06

Pour « les objets placés sur la table pour le repas » (06), un des deux centres qui ne sont pas au programme et un peu plus abstraits, comme le reflète le Graphique 11, l'évolution est croissante du début jusqu'à la fin.

Pour le vocabulaire de la cuisine et ses ustensiles (07), nous voilà de nouveau en présence d'un centre diffus dans lequel nous retrouvons une grande hétérogénéité dans les réponses de nos informateurs. Le seul *vocabulaire* qui revient dans chaque groupe est « table » qui n'est pas un ustensile. Nous trouvons quelques couverts, de la nourriture, de l'électroménager, des objets de cuisson, du mobilier mais aucun ustensile à proprement parlé. À l'exception du dernier, les indices de disponibilité sont très bas dans tous les groupes (cf. Tableau 11).

Tableau 11.
Premiers *vocables* disponibles en fonction du facteur niveau scolaire. C. I. 07

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Table	0,16	Table	0,39	Table	0,15	Cuillère	0,13	Cuillère	0,66
2	Télé(vision)	0,03	Chaise	0,13	Cuillère	0,09	Table	0,12	Table	0,42
3	Micro-ondes	0,02	Frigo(rifique)	0,07	Verre	0,05	Verre	0,12	Couteau	0,33
4	Ciseaux	0,02	Vitrocéramique	0,06	Bol	0,04	Serviette	0,05	Chaise	0,26
5	Serviette	0,02	Micro-ondes	0,05	Chaise	0,03	Chaise	0,05	Eau	0,25
6	Fruit	0,01	Porte	0,04	Eau	0,02	Casserole	0,04	Frigo(rifique)	0,24
7	Frigo(rifique)	0,01	Étagère	0,04	Frigo(rifique)	0,02	Couverts	0,04	Plat	0,21
8	Nespresso	0,01	Fenêtre	0,03	Armoire	0,02	Farine	0,04	Verre	0,17
9	Étagère	0,01	Plaque (électrique)	0,02	Casserole	0,02	Bol	0,04	Fourchette	0,13
10	Bouteille	0,01	Évier	0,02	Couteau	0,02	Cafetière	0,02	Chaudron	0,13



Graphique 12.

Évolution de 5 vocables en fonction du facteur niveau scolaire. C. I. 07

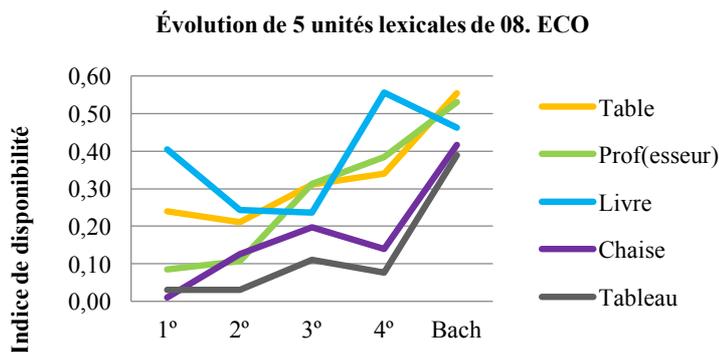
Comme l'exprime le Graphique 12, dans « la cuisine, ses meubles et ustensiles » (07), le deuxième centre n'étant pas au programme et plus abstrait, l'évolution est croissante de 1° à 2°, mais décroissante jusqu'en 4°, et redevient croissante mais de manière plus marquée jusqu'à la fin.

Pour le dernier C. I., nous remarquons les principales différences dans les réponses des élèves de 2° qui apportent plus de matières scolaires alors que les élèves de 1° citent plutôt du matériel scolaire. Ces divergences s'expliquent certainement par les thématiques des différents programmes scolaires comme nous l'expliquions précédemment (cf. Tableau 12).

Tableau 12.

Premiers vocables disponibles en fonction du facteur niveau scolaire. C. I. 08

Rang	1°	I. D.	2°	I. D.	3°	I. D.	4°	I. D.	BACH	I. D.
1	Gomme	0,48	Math(ématique)s	0,35	Stylo	0,35	Stylo	0,68	Table	0,55
2	Crayon	0,45	Classe	0,32	Prof(esseur)	0,31	Livre	0,56	Prof(esseur)	0,53
3	Livre	0,40	Livre	0,24	Table	0,31	Sac-à-dos	0,52	Livre	0,46
4	Trousse	0,35	Cantine	0,24	Classe	0,27	Prof(esseur)	0,39	Chaise	0,42
5	Stylo	0,38	Français	0,23	Sac-à-dos	0,26	Cahier	0,37	Tableau	0,39
6	Cahier	0,25	Table	0,21	Livre	0,24	Table	0,34	Trousse	0,35
7	Table	0,24	Bibliothèque	0,17	Math(ématique)s	0,22	Crayon	0,28	Élève	0,32
8	Règle	0,23	Anglais	0,16	Chaise	0,20	Classe	0,20	Cahier	0,29
9	Sac-à-dos	0,22	Histoire Géographie	0,16	Cahier	0,17	Math(ématique)s	0,16	Classe	0,24
10	Ciseaux	0,20	Stylo	0,16	Gomme	0,14	Règle	0,16	Stylo	0,24



Graphique 12.

Évolution de 5 *vocables* en fonction du facteur niveau scolaire. C. I. 08

Pour le vocabulaire de l'école (08), au programme de 1° pour le matériel scolaire et de 2° pour les matières, l'évolution est croissante de 1° à 2° puis décroît légèrement jusqu'à 3° avant de redevenir croissante (cf. Graphique 12).

Dans tous les C. I. la moyenne des élèves de *Bachillerato* est supérieure à celle de 1° et à l'exception du premier C.I, cette même moyenne est supérieure pour les élèves de 4° que ceux de 1°. En revanche dans les deux premiers C. I. la moyenne de mots par élève de 3° est inférieure à celle de 1° et plus surprenant encore : dans six C. I. (sur huit) la moyenne de mots par élève de 2° est supérieure à celle de 3°... enfin dans trois (presque la moitié) C. I. celle des élèves de 2° est inférieure à celle de 1°.

4. Conclusion

Une fois les résultats de nos enquêtes analysés nous pouvons tout d'abord affirmer que la disponibilité lexicale de nos informateurs est effectivement très inférieure à celle des locuteurs natifs puisque la moyenne par informateurs dans notre enquête est de 3,63 mots par *centre d'intérêt* alors que pour les locuteurs natifs elle atteint environ les 20 mots par *centre d'intérêt*. Bien que nous ayons observé une corrélation dans la position des deux premiers et le sixième *centre d'intérêt*, il n'existe pas vraiment de relation entre ceux qui présentent le plus de mots et ceux qui présentent le plus de *vocables*. En outre, on observe des différences dans la position des *centres d'intérêt* : les parties du corps (01) cinquième en quantité de mot passe à la troisième position en quantité de *vocables*, ou encore les parties de la maison (03) qui passe de la dernière position en quantité de mots à la cinquième en quantité de *vocables*, à l'inverse « la cuisine, ses meubles et ustensiles » (07) passe de l'avant dernière position en quantité de mots à la quatrième en quantité de *vocables*. Ceci s'explique par la plus grande convergence de réponses dans les deux premiers centres car ces thèmes font partie du programme

scolaire et font référence à des thèmes plus concrets que « la cuisine, ses meubles et ustensiles » (07) dans laquelle les élèves se sont fiés, semble-t-il, à leur intuition personnelle.

Il existe effectivement une relation associative entre le niveau scolaire et la quantité de mots et de *vocables*. Pour ce qui est de la quantité de *vocables*, l'évolution de la moyenne de *vocables* par élève est croissante au long de toute la scolarité, ce qui veut dire que plus les élèves étudient le français plus ils diversifient leur vocabulaire. En revanche, comme nous l'avons démontré l'évolution du nombre d'unités lexicales n'est pas toujours croissante en fonction des années. En effet, la courbe reflétant l'évolution du bagage lexical n'est pas toujours croissante mais plutôt irrégulière. Ce qui pourrait peut-être nous laisser penser que l'apprentissage n'est pas linéaire mais en spirale, fait d'allées et venues. En effet, la moyenne totale de mots est de 25,24 en 1^o par élève, ensuite elle augmente passant à 30,82 en 2^o, puis elle descend en 3^o arrivant à 27,48, après elle augmente de nouveau de façon plus marquée atteignant les 35,28 et dépassant ainsi les résultats de 1^o avant de presque se multiplier par 2 en *Bachillerato* avec 68,25 mots par élève tous *centres d'intérêt* confondus. Mais cette évolution varie en fonction des *C. I.*

Alors pourquoi avons-nous ces résultats ? Nous expliquons que les meilleurs résultats de 4^o face à ceux de 3^o s'expliquent probablement par le caractère optionnel du français en 4^o et par conséquent le facteur motivation. (Facteur renforcé dans les étapes suivantes). La motivation (liée à un choix d'option donc un certain intérêt pour la matière) serait donc un facteur déterminant en disponibilité lexicale.

Pourquoi ces résultats en 3^o ? Comme nous l'expliquons, en troisième le programme en français est nouveau et constitue une nouvelle étape pour nos élèves, mais pourquoi ne sont-ils performant que dans le *C. I.* qui est en rapport avec ce qu'ils ont au programme ? Pourquoi ne se rappellent-ils pas de ceux qu'ils ont déjà vu, alors que les élèves de 4^o en sont capables et ceux de *Bachillerato* aussi ? Comme nous l'évoquons, cet élément de réponse pose une question essentielle dans l'apprentissage du vocabulaire, non pas sur la quantité mais sur la qualité de l'apprentissage et sa durée dans le temps. Effectivement, cela signifie peut-être qu'on apprend pour le court terme et pas le long terme, et que le traitement de ces informations n'est pas proposé en cours d'apprentissage qui permettrait leur stockage dans la mémoire à long terme. Ce qui expliquerait la forte perte. Par conséquent, nous pouvons penser que leur préoccupation sur le manque de vocabulaire est donc plutôt due à une question méthodologique dans l'enseignement et probablement l'apprentissage du vocabulaire qu'une question de quantité à proprement parlé.

Il n'existe pas vraiment de relation associative entre l'influence de l'espagnol et la quantité de mots et *vocables*. Seulement 4 des 20 premiers disponibles sont identiques dans les deux langues que ce soit à l'écrit ou à l'oral « pantalon », « pied », « télé(vision) », et « salon ».

Une fois nos conclusions exposées, des questions surgissent sur les limites de la disponibilité lexicale. Comme nous l'avons expliqué dans la première partie, cette méthode révèle-t-elle réellement le vocabulaire qu'un élève utiliserait dans une situation quotidienne (vocabulaire actif) ou n'est-elle pas une épreuve de performance dans laquelle l'élève s'efforce à retranscrire le plus de mots possibles pour satisfaire le professeur et assurer avec succès son « métier d'élève ». Même si ces cas restent isolés, les mots en espagnol qui apparaissent tels que sans avoir fait l'effort de les traduire en sont un indice. En outre, un autre problème s'est posé devant des mots tels que *pie* (*pied* en espagnol) qui se prononce de la même façon dans les deux langues, mais est-il transcrit au même titre que *ventana* (*fenêtre*) qui lui ne ressemble à aucun mot français. Nous pouvons penser que l'élève n'utiliserait pas *ventana* et s'il le faisait il serait difficilement compris sans recours à d'autres ressources langagières. Mais qu'en est-il de *pie* ? L'élève commet-il une erreur d'orthographe, puisque celle-ci n'est pas prise en compte, ou l'inscrit-il dans un souci de performance ? Pour ces problèmes qui se sont manifestés à plusieurs reprises, nous avons dû faire des choix et procéder au cas par cas en fonction de notre connaissance de l'élève et de l'ensemble de sa production.

Pour toutes ces raisons, cette épreuve ne peut offrir qu'un reflet assez relatif d'une situation de communication puisque nous ne parlons pas en exposant des listes de mots. Il est donc important d'envisager le concept de « *vocabulaire disponible* » et de le distinguer du vocabulaire général. Pourrions-nous envisager une épreuve orale dans laquelle il faudrait réciter des listes ou parler d'un sujet, une conversation dont on étudierait les mots employés ? Plus compliquée à mettre en place, puisque plus longue mais qui résoudrait peut-être ces problèmes...

Pour terminer, nous pensons que l'étude de la disponibilité lexicale des élèves de FLE ouvre une nouvelle voie pour vérifier le degré d'assimilation du vocabulaire de nos élèves et l'évolution de celui à travers l'apprentissage. Les résultats peuvent s'avérer très utiles pour la linguistique appliquée à la DLE, domaine encore peu exploré par les chercheurs en FLE. En ce sens, notre travail avec des élèves de FLE est une innovation par rapport aux travaux existants jusqu'à aujourd'hui, il n'existe d'ailleurs, à notre connaissance, que très peu de travaux similaires en français. Notre expérience avec un échantillon composé de 294 élèves a mis l'accent sur les possibilités que peuvent offrir ce type d'étude à l'enseignement du vocabulaire pour l'élaboration de supports en fonction des besoins d'apprentissage plus axés sur la qualité de l'apprentissage dans le temps.

Bibliographie

- Aabidi, L. (2019). La disponibilidad léxica en español como lengua extranjera: dos décadas de investigación científica. *MarcoELE*, (28).
- Bailly, D. (1997). *Didactique de l'anglais I. Objectifs et contenus de l'enseignement*. Paris, France: Nathan pédagogie.
- Boggards, P. (1994). *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris : Hatier.
- Carcedo González, A. (1999). Estudio comparativo del vocabulario español (LE) disponible de estudiantes finlandeses y el de la sintopía madrileña: propuestas didácticas. *Documentos de Español Actual (DEA)*, 1, 73-87.
- Carcedo González, A. (2000). Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición). Turku, Finlande : Annales Universitatis Turkuensis.
- Coady, J. et Huckin, T. (1997). *Second Language Vocabulary Acquisition*. Cambridge : Cambridge University Press.
- Conseil de l'Europe. (2001). *Cadre européen commun de référence pour les langues. Apprendre, enseigner, évaluer*. Strasbourg/Paris : Conseil de l'Europe/Didier.
- Courtillon, J. (1989). Lexique et apprentissage de la langue. *Le Français dans le monde. Recherches et applications*, (6), 146-153.
- De la Maya Retamar, G. (2015). *El desarrollo de la competencia léxica en la adquisición del francés como lengua extranjera en la ESO*. (Thèse de doctorat). Université d'Estrémadure.
- Galisson, R. (1991). *De la langue à la culture par les mots*. Paris : Clé international.
- Galisson, R. (1997). Une dictionnaire à géométrie variable au service de la lexiculture. *Cahiers de lexicologie*, 70(1), 57-77.
- Galisson, R. et Coste D. (1976). *Dictionnaire de didactique des langues*. Paris : Hachette.
- Gougenheim, G., Rivenc, P., Michéa R. et Sauvageot, A. (1964). *L'élaboration du français fondamental*. Paris : Didier.
- Hernández Muñoz, N. (2006). *Hacia una teoría cognitiva integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellanomanchegos*. Salamanca : Université de Salamanca.
- Hidalgo Gallardo, M. (2017). *La disponibilidad léxica como método de detección del vocabulario y de su elección en manuales: aplicación en una muestra de estudiantes sinohablantes de ELE*. (Thèse de doctorat). Université de Jaén.

- Izquierdo Gil, M. C. (2004). *La selección del léxico en la enseñanza del español como lengua extranjera. Su aplicación al nivel elemental en estudiantes francófonos*. (Thèse de doctorat). Université de Valence.
- Lewis, M. (2000). *Teaching Collocation: Further Developments in the Lexical Approach*. Hove : Language Teaching Publications.
- López Chávez, J. et Strassburger Frías, C. (1987). Otro cálculo del índice de disponibilidad léxica. Dans *Presente y Perspectiva de la investigación computacional en México. Actas del IV Simposio de la Asociación de Lingüística Aplicada* (pp. 67-89). Mexique : Université nationale autonome du Mexique.
- López González, A. M. (2014). Desarrollo de los estudios de disponibilidad léxica en Español Lengua Extranjera (ELE). Dans N. M. Contreras Izquierdo (Dir.), *La enseñanza del español como LE / L2. Actas del XXIV Congreso Internacional ASELE* (pp. 411-422). Malaga : ASELE.
- López Morales, H. (1999). Disponibilidad léxica y estratificación socioeconómica. Dans *Dialectología y Sociolingüística. Temas Puertorriqueños* (pp. 173-181). Madrid : Hispanova.
- Mackey, W. (1971). *Le vocabulaire disponible du français*. Paris-Bruxelles-Montréal : Didier.
- Martín Martín, S. (2007). La enseñanza del léxico más allá del diccionario. *Mosaico*, (20), 43-46.
- Michéa, R. (1950). Vocabulaire et culture. *Les langues modernes*, 44, 188-189.
- Michéa, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les langues modernes*, 47, 338-344.
- Nalesso, G. (2019). *Disponibilidad y riqueza léxica en un grupo de estudiantes universitarios italianos de ELE*. (Thèse de doctorat). Université de Padoue.
- Paredes García, F. (2012). Desarrollos teóricos y metodológicos recientes de los estudios de disponibilidad léxica. *Revista Nebrija de Lingüística Aplicada*, 11(6), 78-100.
- Samper Hernández, M. (2002). *Disponibilidad léxica en alumnos de español como lengua extranjera*, Malaga : ASELE.
- Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible. *Lingüística*, 10, 311-333.
- Santos Díaz, I. C. (2015). *Evaluación de la competencia léxica bilingüe en estudiantes del Máster Universitario en Profesorado*. (Thèse de doctorat). Université de Malaga.