

**MESURER L'IMPACT D'UN DISPOSITIF DE REMÉDIATION
ANGLAIS/ESPAGNOL SUR LE DÉVELOPPEMENT DE COMPÉTENCES
LANGAGIÈRES ET D'AUTONOMIE D'APPRENTISSAGE EN LANGUE**

Justine PARIS

Université Paris Cité, EA3967 CLILLAC-ARP, UMR7118 ATILF (membre associée)

Ismael RAMOS RUIZ

Université Paris Cité, FENIX SEJ-596 (Université de Séville)

Gabriel THIBERGE

Université Paris Cité, UMR7110 LLF

Mots-clés

Didactique des langues - évaluation de dispositifs - statistiques en didactique

Keywords

Course Evaluation - Second Language Acquisition (SLA) - Statistics applied to SLA

Résumé

Dans cette contribution, nous cherchons à évaluer l'efficacité d'un dispositif de remédiation anglais/espagnol sur le développement de compétences langagières et de compétences d'autonomie d'apprentissage chez des étudiant·e·s en première année de licence LEA. Pour ce faire, nous avons mis au point un test de niveau pour chaque langue et un questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage que nous avons soumis avant et après le déroulé du cours dans une démarche méthodologique de « pré-tests / post-tests ». Les résultats d'une étude pilote sur la base d'une première utilisation de ces outils ont montré que le dispositif avait un effet plutôt positif sur les compétences d'autonomie d'apprentissage mais pas sur les compétences langagières que nous avons étudiées. Nous revenons ici sur la méthodologie mise en place pour mesurer ces effets et sur la nature des résultats observés.

Abstract

In this contribution, we study the impact of a remediation course in English and Spanish on the development of both linguistic skills and language learning autonomy skills in first year Applied Modern Languages students. To do so, we developed language proficiency tests in English and Spanish as well as a self-evaluation questionnaire on the degree of learning autonomy for each language. These were submitted to pilot participants before and after the completion of the course using a "pre-test/post-test" methodological procedure. Preliminary results showed that the course had a positive effect on the participants' language learning autonomy but not on their proficiency level. In this paper, we describe the procedure that we developed to measure these effects and the preliminary results that we observed.

Introduction

En 2018, entre en vigueur la loi ORE (Orientation et à la Réussite des Étudiants¹) qui a pour but d'accompagner chaque étudiant·e sur la voie de la réussite. Au niveau Licence, cet accompagnement se matérialise, entre autres, par des parcours aménagés en première année pour les étudiant·e·s ne présentant pas tous les prérequis incombant à la filière d'orientation qu'ils ont choisie (parcours « oui, si » et autres dispositifs de soutien et de suivi personnalisés).

C'est dans ce contexte que l'UFR EILA (Études Interculturelles de Langues Appliquées) d'Université Paris Cité a mis en place un dispositif hybride qui a pour but d'accompagner les étudiant·e·s les plus faibles en anglais et en espagnol dans le cadre de la première année de licence LEA (Langues Étrangères Appliquées). Ce dispositif, qui prend le nom de *remédiation*, est composé de deux UE libres (une pour chaque langue) proposées en priorité aux étudiant·e·s ayant besoin d'aide dans l'une de ces deux langues afin de garantir au mieux leur entrée en L1 LEA anglais-espagnol. L'objectif général de ce dispositif est donc de donner « un coup de pouce » aux étudiants dont le niveau d'anglais ou d'espagnol serait insuffisant pour s'accomplir en licence LEA. Pour ce faire, nous avons décidé de nous concentrer sur les compétences langagières générales (compréhension de l'oral et de l'écrit, production orale et écrite) et de miser sur un maximum d'input et de pratique de la langue.

D'un point de vue pédagogique, ce dispositif consiste en un enseignement hybride qui allie des séances de cours en présentiel (pour un apport lexical, grammatical, pragmatique et socioculturel général en langue) et un temps de travail personnel à distance en autodirection (pour que chaque étudiant·e puisse se concentrer sur ses difficultés et besoins personnels dans la langue choisie afin de s'améliorer). Le temps de travail autodirigé fait l'objet d'une préparation en classe où les étudiant·e·s sont tout d'abord amené·e·s à réfléchir sur leur apprentissage des langues jusqu'alors, puis à s'auto-évaluer dans la langue choisie et, enfin, à faire le point sur leur besoins / envies / difficultés dans cette langue dans le but de se fixer des objectifs et d'élaborer un programme de travail personnel à conduire en parallèle du cours. Les étudiant·e·s sont ensuite suivi·e·s individuellement par les enseignant·e·s à travers un carnet de bord que ces dernier·ère·s lisent et commentent chaque semaine afin d'apporter conseils et soutien personnalisés à chacun·e. Le dispositif comporte également deux séances d'entretien individuel où les enseignant·e·s reçoivent chaque étudiant·e pour faire le point sur ses avancées, difficultés et incompréhensions au sein du dispositif (temps de cours en présentiel et temps de travail autodirigé), mais également au sein des autres unités d'enseignement de la L1 LEA anglais-espagnol. Enfin, chacune des évaluations

¹ LOI n° 2018-166 du 8 mars 2018 relative à l'orientation et à la réussite des étudiants. Accessible sur : <<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000036683777/>> (dernière consultation 18-01-2023).

sommatives du dispositif (présentation orale en groupe sur documents libres, essai individuel sur thématiques abordées en cours et rapport écrit en groupe sur documents vidéo/textuels imposés) fait l'objet de séances de préparation et de séances de perfectionnement afin que les étudiant·e·s puissent travailler leurs productions en profondeur avant de rendre leurs travaux définitifs.

D'un point de vue théorique, la conception de ce dispositif repose sur le paradigme de l'autonomie d'apprentissage en langue. Henri Holec, un des pionniers du domaine, conçoit l'autonomie d'apprentissage d'un·e apprenant·e comme sa « capacité de mener, activement et de manière indépendante, un apprentissage » (Holec, 1994 : 11). Cette capacité se décline en 5 dimensions (ou 5 sous-domaines de compétence) : 1) définir ses propres objectifs, contenus et modalités de réalisation, 2) organiser son apprentissage, 3) choisir des ressources et méthodes en conséquence, 4) assurer le suivi de ses avancées et 5) évaluer ses avancées (Holec, 1990). Du point de vue des langues spécifiquement, il est donc tout aussi important que les apprenant·e·s développent une certaine autonomie langagière dans leur langue d'apprentissage (pouvoir communiquer dans cette langue cible) qu'une certaine autonomie d'apprentissage de la langue (savoir comment s'y prendre pour apprendre cette langue cible). Le dispositif de remédiation anglais/espagnol, qui fait l'objet de la présente étude, s'efforce donc de répondre à ces deux objectifs : le temps de cours visant l'autonomie langagière générale ; le temps de travail autodirigé visant l'autonomie langagière sur objectifs personnels et l'autonomie d'apprentissage en langue.

Afin de voir si ce dispositif remplit effectivement ces deux objectifs, nous avons entamé un projet de recherche IdEx de type *Émergence en Recherche*, baptisé RemedLang², où nous faisons l'hypothèse que le dispositif de remédiation anglais/espagnol contribue au développement des aptitudes langagières et d'autonomie d'apprentissage des étudiant·e·s inscrit·e·s en première année de Licence LEA pour ces deux langues. L'objet de cette publication est ainsi de présenter la manière dont nous avons choisi d'évaluer l'efficacité de ce dispositif. Dans les trois parties suivantes, nous revenons sur la démarche méthodologique que nous avons mise en place, les premiers résultats que nous avons obtenus et les conclusions préliminaires que nous tirons pour la suite de la conduite du projet.

² Projet RemedLang, « Évaluation d'un dispositif de remédiation anglais et espagnol dans le but de favoriser l'autonomie langagière et l'autonomie d'apprentissage au niveau L1 » (Idex/EK/RR 2020-291, Université Paris Cité).

1. Démarche méthodologique

Afin de mesurer l'impact du dispositif de remédiation en L1 LEA sur le développement de compétences langagières et de compétences d'autonomie d'apprentissage en anglais et en espagnol, nous avons opté pour une démarche de « pré-tests / post-tests », soit une prise de mesure de chacun de ces deux indicateurs avant et après le déroulé des cours de remédiation afin de pouvoir comparer les scores obtenus suite aux deux passations. Pour ce faire, nous avons mis au point deux outils, un test de niveau d'anglais / espagnol et un questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage en anglais / espagnol, puis nous avons eu recours aux statistiques inférentielles pour exploiter les scores obtenus par ces deux outils. Nous revenons sur la construction de ces derniers en 1.1 et sur la façon dont nous avons souhaité exploiter leurs résultats en 1.2.

1.1. Les outils

1.1.1. *Le test de niveau de langue en anglais et en espagnol*

Sachant que les élèves de lycée obtiennent un niveau se situant entre A2 et B1 en fin de terminale si l'on se réfère aux notes obtenues au baccalauréat en 2017 pour les épreuves d'anglais LV1 et d'espagnol LV2 (Manoïlov, 2019 : 48), nous avons décidé de nous inspirer de sujets de CLES³ établis pour évaluer le niveau B1 afin de réaliser nos tests de niveau dans les deux langues. Dans une démarche d'économie de temps de passation pour nos étudiant·e·s, nous nous sommes limités aux premières parties des compétences de compréhension de l'oral et de l'écrit (une vidéo, un texte et trois questions de compréhension sur chaque support). Et dans un souci d'harmonisation, nous avons sélectionné des sujets de CLES B1 traitant de la même thématique dans les deux langues : l'organisation d'une mobilité entrante dans une université australienne et dans une université espagnole. Nous avons ensuite ajouté un test de grammaire pour chaque langue : un test créé par nos soins en espagnol, que nous avons adapté pour l'anglais à l'aide d'un test de grammaire réalisé par l'un des trois auteurs dans le cadre de sa thèse (Paris, 2015 : chap. 7). Nous n'avons cependant pas pu nous permettre d'ajouter d'épreuves en production : la batterie de tests finale dans les deux langues (en plus du questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage) aurait été alors chronophage pour les étudiant·e·s et trop coûteuse en temps de passation (qui était un critère déterminant dans la conception de nos outils). Cela aurait été également très chronophage d'un point de vue logistique pour réussir à réunir les étudiant·e·s à deux reprises lors de moments-clés de l'année académique (rentrée universitaire et examens de toute fin d'année) – les cohortes de L1 LEA anglais-

³ Certificat de Compétences en langues de l'Enseignement Supérieur

espagnol avoisinant les 130 étudiant-e-s chaque année à l'UFR EILA. Notons que les compétences pragmatiques et sociolinguistiques en langue ne sont pas non plus évaluées dans le cadre de ces deux tests de niveau. Nous sommes ainsi pleinement conscients que nos tests permettent d'obtenir une photographie partielle du niveau de langue de nos participant-e-s et non pas une appréciation complète de ce dernier (au terme de leur première année de LEA, nos participant-e-s auront ainsi probablement progressé sur divers autres points non testés ici ou que nous ne pourrions identifier).

1.1.2. Le questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage en anglais et en espagnol

Si la recherche sur les compétences d'autonomie d'apprentissage en langue remonte aux années 1970, il n'existe cependant pas de moyen de mesurer rapidement leur développement ni d'évaluer rapidement leur niveau à un instant-t chez des apprenant-e-s de langues. Certains tests ont été mis au point mais sont soit très généraux (mesurant les compétences d'autonomie d'apprentissage générales, donc sans viser aucune discipline en particulier – Williamson, 2007 ; Abd-el-fattah, 2010 ; Cheng et al., 2010 ; Ceylan, 2015), soit destinés à caractériser des publics bien spécifiques (comme les apprentis infirmiers – Fisher et al., 2010). À ce jour, seuls des questionnaires de très grande envergure avec une granularité très fine dans les questions posées ont été proposés dans le domaine de l'apprentissage des langues étrangères (Murase, 2010, 87 questions ; et Tassinari, 2010, 2015 et 2016, 153 questions). Un test de plus petite envergure rapidement administrable manquait donc particulièrement dans la littérature au moment où nous avons commencé à entreprendre notre collecte de données. Nous avons donc dû concevoir notre propre outil.

Pour ce faire, nous nous sommes référés à treize questionnaires d'auto-évaluation des compétences d'autonomie d'apprentissage générale et en langue existants dans la littérature, de même qu'à diverses études sur les représentations des apprenant-e-s et des enseignant-e-s sur le rôle de chacun-e dans une situation d'apprentissage générale ou en langue spécifiquement. Nous avons alors procédé à un relevé de toutes les questions et assertions nous paraissant pertinentes dans le cadre de la construction d'un questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage pour les langues (par exemple : *I am able to monitor my learning progress* ; Williamson, 2007). Nous avons ensuite regroupé ces éléments selon les 5 dimensions d'Holec vues précédemment, puis nous avons éliminé les doublons, adapté chacun d'entre eux à l'apprentissage de l'anglais et de l'espagnol et nous les avons traduits en français (par exemple : *I am able to monitor my learning progress* => *Je suis capable de voir par moi-même si je fais des progrès en anglais/espagnol*, à partir de l'item de Williamson cité précédemment). Nous avons alors soumis notre sélection

d'items (N=42) à deux chercheuses spécialistes du paradigme de l'autonomie d'apprentissage en langue. Sur une échelle de 1 à 5 (de *pas d'accord* à *tout à fait d'accord*), celles-ci ont dû se prononcer sur la pertinence de chacun d'entre eux, sur l'importance de leur conservation dans notre questionnaire final et sur la façon dont nous les avons formulés. Au terme de deux sessions d'évaluation de leur part⁴, nous avons réussi à établir un questionnaire court, comportant 28 assertions, soit 5 à 6 items par dimension. En voici quelques exemples : *Je sais sur quoi j'ai besoin de travailler pour m'améliorer en anglais* pour cibler la dimension 1 (se fixer des objectifs d'apprentissage) ; *J'arrive à établir des priorités pour m'améliorer en anglais/espagnol* pour cibler la dimension 2 (organiser son apprentissage) ; ou encore *Je vérifie régulièrement si j'ai atteint les objectifs d'apprentissage que je m'étais fixés en anglais/espagnol* pour cibler la dimension 4 (suivre ses avancées)⁵.

1.2. Exploitation des mesures collectées : démarche « pré-tests / post-tests » et recours aux statistiques inférentielles

Afin de mesurer l'impact de notre dispositif de remédiation sur les compétences langagières et d'autonomie d'apprentissage de nos étudiant·e·s en anglais et en espagnol, la comparaison de mesures avant et après le suivi du dispositif nous est apparue comme la manière la plus objective d'en observer l'efficacité. Nous aurions pu opter pour un questionnaire d'Évaluation de l'Enseignement par les Étudiant·e·s (« EEE », ce que nous faisons déjà passer depuis quatre ans) mais les données collectées se seraient limitées exclusivement aux représentations de nos étudiant·e·s. Il aurait été alors difficile de voir comment évaluer l'impact de notre dispositif autrement qu'à travers la comparaison de mesures avant / après le déroulé du cours et entre des sujets l'ayant suivi et d'autres pas afin d'obtenir des résultats les plus représentatifs et les plus précis possibles.

Nous avons ainsi soumis les versions finales de nos deux outils deux fois à tous les étudiant·e·s inscrit·e·s en L1 LEA anglais-espagnol de l'année universitaire 2021-2022 dans le cadre d'une étude pilote. Une première fois en septembre 2021 (N=117) en guise de pré-test, puis en juin 2022 (N=62⁶) en guise de post-test. Afin de voir si les étudiant·e·s du cours de remédiation (N=16) avaient obtenu de meilleurs scores lors de la session de juin comparé

⁴ Il était possible de formuler des commentaires personnels dans un champ additionnel pour chaque item, nous avons donc jugé intéressant de leur soumettre notre sélection d'items une deuxième fois après la prise en compte de leurs évaluations chiffrées et de leurs remarques plus qualitatives.

⁵ Notre questionnaire d'auto-évaluation du degré d'autonomie d'apprentissage dans son intégralité fera l'objet d'une publication ultérieure afin d'être mis à la disposition de la communauté scientifique.

⁶ Cette chute en termes d'effectifs s'explique par le fait que certain·e·s étudiant·e·s aient abandonné leur parcours de L1 LEA en cours d'année, que d'autres aient abandonné le cours de remédiation au moment où les UE libres ont été exceptionnellement rendues optionnelles en 2021-2022 en raison de la Covid-19 et que d'autres aient passé une seule batterie de tests (voire une seule langue sur les deux) lors de la première ou de la deuxième session de passation.

aux étudiant-e-s n'ayant pas suivi le cours (N=46), un recours aux études quantitatives et aux statistiques inférentielles s'est imposé. Une comparaison de leurs scores « à vue d'œil » n'aurait pas été envisageable scientifiquement car elle n'aurait pas permis de savoir s'ils étaient représentatifs (donc potentiellement généralisables) ou s'ils étaient dus au hasard (par exemple, à des traits individuels de certain-e-s de nos participant-e-s).

Nous avons alors pensé à comparer les moyennes obtenues par les deux groupes à chaque test (degré d'autonomie et niveau de langue) et à chaque sous-partie des tests (compréhension orale/écrite, grammaire et aux 5 dimensions d'autonomie d'apprentissage). Nous avons ainsi envisagé de nous tourner vers le « test *t* de Student » pour groupes appariés car il s'agit d'un test statistique qui permet de comparer les moyennes obtenues par un même groupe à deux passations espacées dans le temps d'un même outil (afin de voir si la différence entre les résultats aux deux passations est nulle ou non). Nous avons cependant ressenti le besoin d'avoir l'avis d'un expert en études quantitatives afin d'être sûrs de ne pas écarter d'autres méthodes ou d'autres modèles statistiques possibles, voire plus adaptés à notre étude, ainsi qu'à la nature de nos données. Nous nous sommes alors tournés vers un personnel d'appui à la recherche de notre laboratoire dont les compétences dans ce domaine complétaient les nôtres. Celui-ci a pensé à un type de modèle plus fin permettant de prendre en compte les scores individuels de chacun-e de nos participant-e-s plutôt que de les exploiter à l'échelle des deux groupes (remédiation ou non), de même que d'intégrer d'autres variables par la suite en cas de complément de collecte de données et d'étude de nouvelles variables ou de nouveaux paramètres. Ce qui est le cas : nous répliquons actuellement notre procédure de collecte de données auprès de la cohorte d'étudiant-e-s inscrit-e-s en L1 LEA anglais-espagnol en 2022-2023 et nous comptons le faire également auprès de la cohorte de 2023-2024. Nous détaillons ces modèles et les résultats qu'ils ont permis d'obtenir pour la cohorte pilote de 2021-2022 dans la partie suivante. Cette collaboration et la mise en commun de nos compétences ont donc été très précieuses pour notre étude car elle nous a permis d'aller plus loin dans l'exploitation de nos données.

2. Analyses et résultats généraux

Les 62 participant-e-s intégré-e-s à cette étude constituent un premier échantillonnage depuis lequel il est possible de faire une analyse préliminaire du dispositif de remédiation et de sa contribution à l'amélioration des compétences langagières et d'autonomie d'apprentissage des apprenant-e-s. Au-delà de simples résultats descriptifs, ces données permettent de formuler, par l'inférence statistique, des constats déjà relativement solides quant aux apports de ce dispositif pédagogique.

2.1. Dépasser les résultats descriptifs

Afin d'évaluer le dispositif mis en place, comme indiqué précédemment, il aurait été possible de se contenter de calculer des moyennes (et de répondre à des interrogations du type « de combien de points en moyenne un·e étudiant·e en licence a-t-il-elle progressé entre la première et la seconde passation sans suivre le cours de remédiation ? », ou encore « de combien de points en moyenne un·e étudiant·e du cours de remédiation a-t-il-elle progressé entre la première et la seconde passation ? »). On aurait pu comparer ces progrès moyens selon le groupe (avec ou sans remédiation) et selon la langue (en remédiation anglais ou en remédiation espagnol). Ces descriptions sont, certes, intéressantes, et indicatives de résultats globaux. Elles sont cependant réductrices, en ce qu'un calcul de moyenne fait par définition abstraction de la distribution des résultats au sein d'un groupe. Deux valeurs extrêmes (-5 et +5 par exemple) peuvent donner la même moyenne que deux valeurs moyennes (0 et 0) ou proches de la moyenne (-1 et +1). Par ailleurs, la tentation est grande de comparer la moyenne de distributions d'échantillons de tailles très différentes et de considérer que si la moyenne est identique, alors il n'y a pas de différence entre les deux groupes.

Pour dépasser cette analyse de surface et éviter ces écueils, les statistiques inférentielles sont importantes. Par la comparaison des valeurs de dispersion de distributions (variance, écart-type par exemple) en plus de leurs moyennes, ce type d'analyse permet de généraliser les résultats obtenus à des échantillons plus grands. La question qui les sous-tend est alors la suivante : « à quel point l'écart de moyennes observé entre deux groupes resterait semblable si l'échantillonnage avait été réalisé sur une population différente ou plus large ? ».

Parmi les outils possibles pour analyser les données collectées dans le cadre de la présente étude, nous avons retenu les modèles de régression linéaire, qui sont adaptés aux scores recueillis ici (présence de variables quasi-continues – avec des intervalles de 0.25 – comprises entre 0 et 30 pour les tests de niveau de langue et entre 0 et 140 pour les questionnaires d'autonomie d'apprentissage)⁷. Dans ces modèles, nous avons spécifié une variable dépendante (ici, le progrès numérique entre les évaluations menées en juin et les évaluations menées en septembre à l'entrée en licence) et une variable indépendante (variable contrôlée : le groupe d'appartenance de la personne, soit « groupe sans

⁷ La linéarité des variables a été vérifiée à l'aide d'un test statistique de normalité, le test de Shapiro-Wilk. La normalité des résidus a également été contrôlée (écarts entre les valeurs observées et les valeurs calculées par le modèle).

remédiation » (N=46), « groupe avec remédiation en anglais » (N=10), « groupe avec remédiation en espagnol » (N=6)). C'est principalement cet écart de tailles entre les groupes qui impose un recours aux statistiques inférentielles, pour évaluer la robustesse des observations menées (et répondre à des interrogations du type « s'il y a une différence de progrès moyen entre les groupes, à quel point peut-on généraliser cette observation au vu de la différence de taille qui les caractérise ? »).

De manière simplifiée, les modèles calculés restituent plusieurs valeurs, dont une principale : la « *p-value* » (valeur p, sans unité et comprise entre 0 et 1), qui indique à quel point le progrès moyen observé pour un groupe (les étudiant·e·s ayant suivi la remédiation) diffère significativement du progrès moyen observé pour le groupe de référence (les étudiant·e·s n'ayant pas suivi la remédiation). Par convention, la significativité de cette valeur est définie arbitrairement autour du seuil de 0.05. Si la valeur p est inférieure à ce seuil, le résultat est statistiquement significatif, si elle lui est supérieure, le résultat ne l'est pas. Un résultat significatif indique que la différence observée entre les groupes est extrême par rapport à ce qu'on pourrait attendre si on répétait la mesure un grand nombre de fois sur des groupes semblables, et qu'on peut donc en conclure que la variable indépendante (= le groupe) a eu un effet sur la variable dépendante (= le score).

Les statistiques inférentielles présentées ci-dessous ont été menées dans la suite logicielle R, version 4.2.2 (R Core Team, 2022), avec l'interface RStudio (Posit Team, 2022). Les paquets *readxl* (Wickham & Bryan, 2022), *dplyr* (Wickham et al., 2022) et *reshape2* (Wickham, 2007) ont été utilisés pour importer et manipuler les données, le paquet *ggplot2* (Wickham, 2016) pour générer les visualisations proposées et le paquet *lmerTest* (Kuznetsova, Brockhoff et Christensen, 2017) pour calculer les modèles de régression linéaire. Le script et les données associées à ces statistiques inférentielles, ainsi que les résultats complets des différents modèles présentés, sont disponibles à l'adresse <https://osf.io/3whtr/>.

2.2. Résultats généraux

Pour chaque participant·e (N=62), une valeur « progrès » a été créée en soustrayant le score obtenu en septembre 2021 à celui obtenu en juin 2022 pour chacun des quatre tests (Anglais – Langue ; Anglais – Autonomie ; Espagnol – Langue ; Espagnol – Autonomie). C'est l'ensemble de ces valeurs numériques calculées qui a été implémenté en tant que variable dépendante aux modèles de régression linéaire présentés ci-dessous pour observer et évaluer d'éventuelles progressions différentes entre les groupes (avec et sans remédiation). Pour chacune des deux langues successivement, deux modèles ont été calculés : pour le progrès en langue, puis pour le progrès en autonomie d'apprentissage.

2.2.1. Pour l'anglais

Concernant l'anglais, une comparaison superficielle des données calculées montre que, pour le groupe n'ayant pas suivi de cours de remédiation, le progrès moyen entre la passation du test de langue de septembre et celle de juin est de 1.34 point, tandis que le progrès moyen pour le groupe ayant suivi la remédiation est de 0.28 point. Pour le degré d'autonomie d'apprentissage, on observe un progrès notable de 16.6 points pour le groupe ayant suivi la remédiation, mais une régression pour le groupe sans remédiation (-1.17 point). Les Figures 1 et 2 illustrent la distribution de ces données dans les effectifs des groupes, avec, en abscisse, le groupe, et en ordonnée, le progrès en valeur numérique. Sur ces graphiques, le petit carré noir indique la moyenne, la boîte est une « boîte à moustaches » et représente les quartiles de la distribution (25 % des valeurs dans la partie inférieure de la boîte centrale, 25 % dans la partie supérieure, et 25 % pour chaque trait en-dessous et au-dessus de la boîte, avec la médiane indiquée par un trait horizontal dans la boîte). L'ombre blanche à l'arrière est un diagramme « en violon » et permet d'affiner encore la visualisation des distributions.

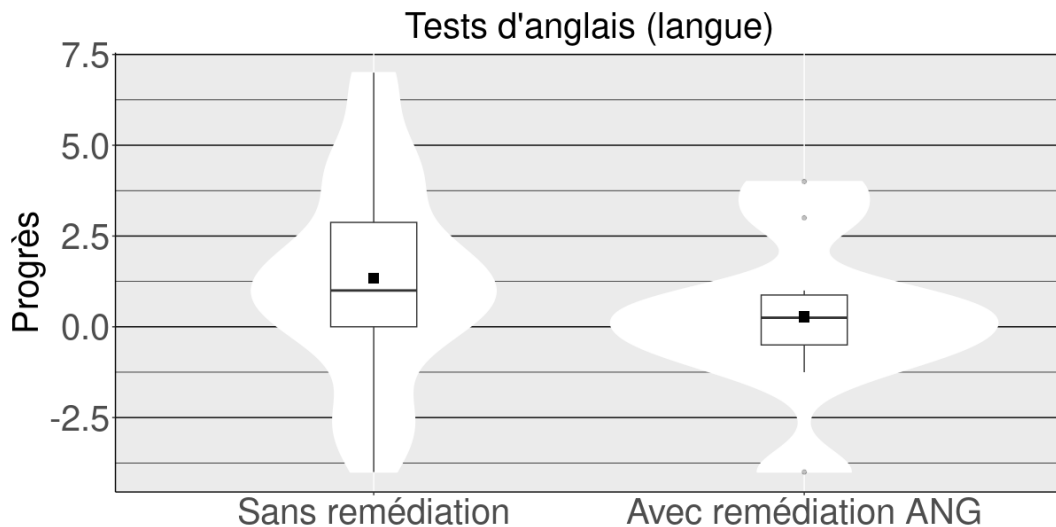


Figure 1. Progrès en langue (anglais)

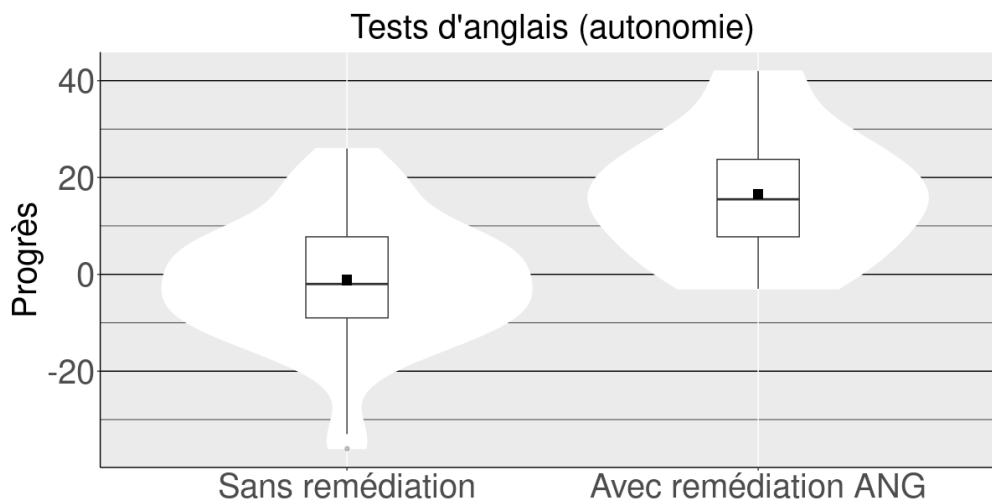


Figure 2. Progrès en autonomie d'apprentissage (anglais)

Des modèles calculés sur ces données permettent d'en évaluer la portée généralisable (Tableau 1). Le premier modèle calculé pour le niveau de langue indique que la différence de moyennes entre les deux groupes n'est pas significative, et que si le progrès est numériquement supérieur chez les étudiant-e-s n'ayant pas suivi la remédiation, cette tendance n'est pas généralisable, au vu des données, à un échantillon plus large de la population. À l'inverse, le second modèle calculé pour le degré d'autonomie d'apprentissage montre que la différence observée entre les deux groupes est cette fois significative (les étoiles dans la dernière colonne indiquent une significativité statistique). D'après la distribution des données observées, la tendance (meilleur score pour les étudiant-e-s ayant suivi la remédiation) est généralisable à une population plus large malgré les différences d'effectifs entre les deux groupes.

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Langue	0.28	1.34	0.25246	
Autonomie	16.60	-1.17	0.000251	***

Tableau 1. Résultats des modèles généraux pour l'anglais⁸

2.2.2. Pour l'espagnol

Concernant l'espagnol, les Figures 3 et 4 illustrent le progrès réalisé en langue et en autonomie d'apprentissage par les groupes avec et sans remédiation.

⁸ Une version de ce tableau, ainsi que de ceux qui suivent, comportant les valeurs brutes des modèles par rapport au groupe de référence (groupe sans remédiation) est disponible ici : <https://osf.io/f4kvu>.

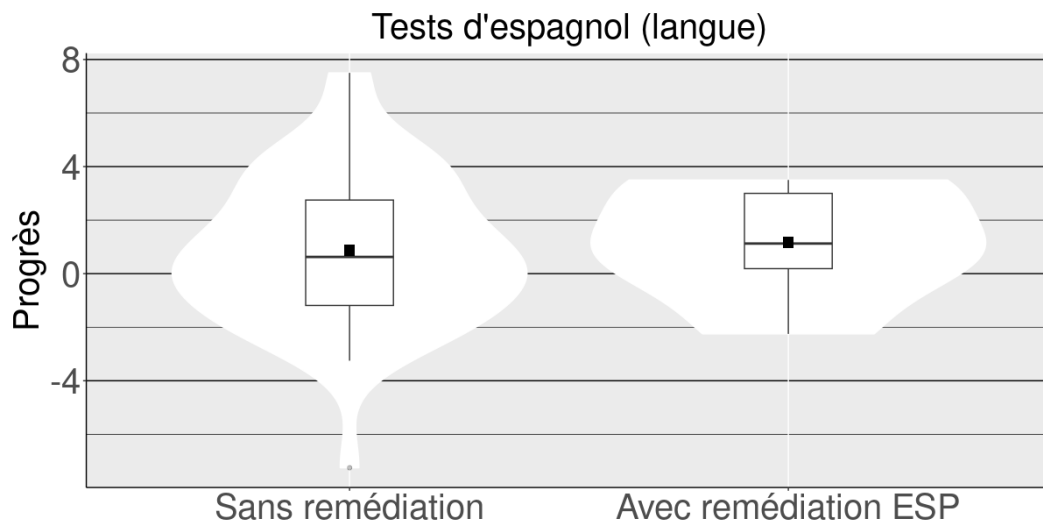


Figure 3. Progrès en langue (espagnol)

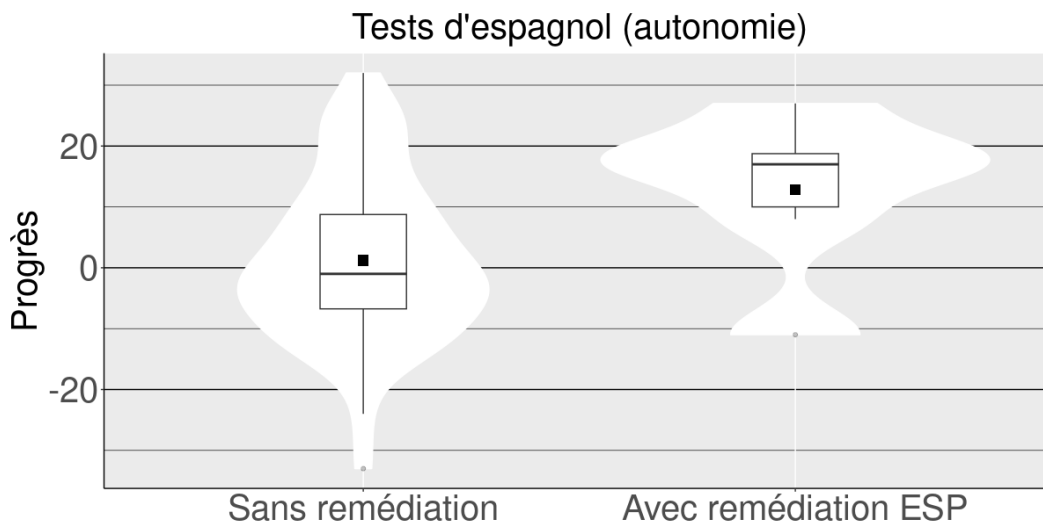


Figure 4. Progrès en autonomie d'apprentissage (espagnol)

Ces graphiques semblent indiquer un patron identique à l'anglais. Des modèles calculés sur ces données tendraient à le confirmer, mais c'est une affirmation à nuancer (Tableau 2). Si les données sur le progrès en langue espagnole sont semblables à celles de l'anglais (différence entre moyennes non significative), celles sur le degré d'autonomie sont un peu différentes. En effet, pour l'espagnol, la différence de moyenne entre les groupes n'est ici pas significative (on parle de « tendance statistique » pour une valeur p comprise entre 0.05 et 0.10). Une valeur extrême (basse), qu'on aperçoit sur le Figure 4, et le trop petit nombre d'observations pour ce groupe (N=6), empêchent de conclure de façon solide à un effet du groupe sur le progrès moyen ici. Ceci illustre la nécessité de procéder à des analyses

inférentielles et de dépasser la simple description visuelle des résultats pour s'assurer de la robustesse des résultats d'un dispositif pédagogique.

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Langue	1.17	0.88	0.8111	
Autonomie	12.83	1.22	0.0558	.

Tableau 2. Résultats des modèles généraux pour l'espagnol

3. Résultats par compétences langagières et dimensions d'autonomie

Avec la même démarche, il est possible d'évaluer si, pour chaque compétence et dimension des deux tests, les progrès observés sont différents ou non en fonction des groupes.

3.1. En langue

Le Figure 5 présente les progrès en langue anglaise réalisés par les groupes ayant suivi le dispositif de remédiation (ligne pointillée et carrés pleins) ou non (ligne pleine et cercles vides). Le progrès en valeur numérique est en ordonnée, et, en abscisse, se trouvent les trois compétences langagières testées. Les barres autour des cercles/carrés sont les barres d'erreur, indiquant la dispersion des valeurs observées autour de la valeur moyenne au sein de chaque groupe. Plus ces barres sont éloignées autour de la moyenne, plus les valeurs sont dispersées. Plus les barres d'erreurs d'un groupe sont éloignées de celles de l'autre groupe, plus il y a de chances que la différence entre les deux groupes soit importante et généralisable. Néanmoins, pour s'en assurer, le calcul d'un modèle de régression linéaire est, de nouveau, indispensable.

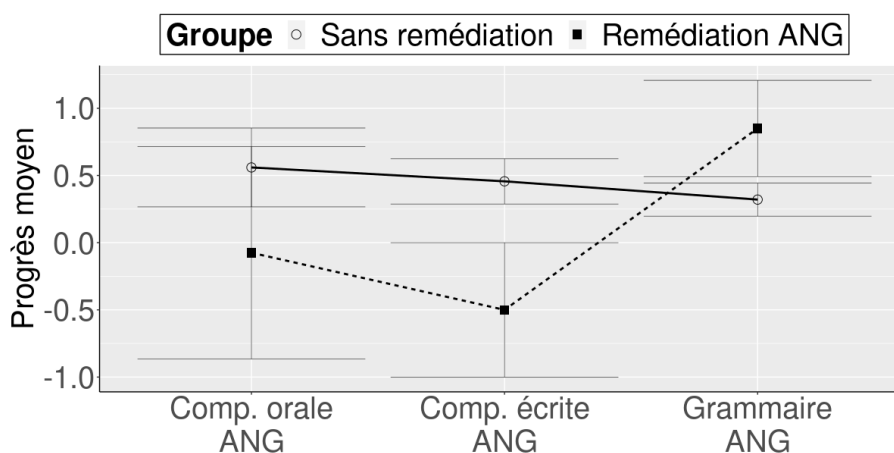


Figure 5. Progrès en langue par compétence langagière (anglais)

Ces modèles, calculés pour chacune des trois compétences langagières testées, permettent de dépasser l'interprétation qui se dégagerait d'une simple inspection visuelle du graphique (meilleure réussite des étudiant-e-s n'ayant pas suivi la remédiation concernant les compétences de compréhension de l'oral/écrit, l'inverse en grammaire). Le Tableau 3 montre que seule la différence observée en compréhension écrite est significative. Les autres différences ne peuvent être généralisées, ni dans un sens ni dans l'autre.

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Comp. orale	-0.08	0.56	0.386	
Comp. écrite	-0.5	0.46	0.0297	*
Grammaire	0.85	0.32	0.0943	.

Tableau 3. Différences entre les groupes et significativité en langue (anglais)

Le Tableau 4 permet de mettre en parallèle les résultats obtenus pour l'espagnol sur ce même test de langue. On n'y observe cette fois aucune différence de moyennes qui semble importante numériquement selon le groupe, et les modèles de régression calculés confirment qu'aucune différence n'est significative entre le groupe ayant suivi la remédiation et le groupe ne l'ayant pas suivie.

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Comp. orale	-0.04	-0.04	0.996	
Comp. écrite	0.54	0.64	0.9142	
Grammaire	0.67	0.28	0.526	

Tableau 4. Différences entre les groupes et significativité en langue (espagnol)

De manière générale, on observe donc que les progrès en langue sont modérés quelle que soit la compétence langagière observée et quelle que soit la langue considérée, que les étudiant-e-s aient suivi le dispositif de remédiation ou non.

3.2. En autonomie d'apprentissage

Le Figure 6 illustre le progrès moyen observé dans les cinq dimensions du questionnaire d'autonomie d'apprentissage en anglais pour le groupe ayant suivi la remédiation et celui ne l'ayant pas suivie. On y observe que, de manière générale, le degré d'autonomie d'apprentissage du groupe ayant suivi la remédiation semble souvent plus important que celui du groupe ne l'ayant pas suivie. Un modèle calculé pour chacune de ces dimensions confirme ce patron général pour les dimensions 1 (se fixer des objectifs), 2 (organiser son apprentissage), 4 (suivre ses avancées) et 5 (s'auto-évaluer), et ce, de façon robuste statistiquement, comme le résume le Tableau 5. La différence observée pour la

dimension 3 (trouver des ressources/activités/exercices) n'est, elle, pas significative d'un point de vue statistique ; collecter plus de données serait nécessaire pour cette dimension.

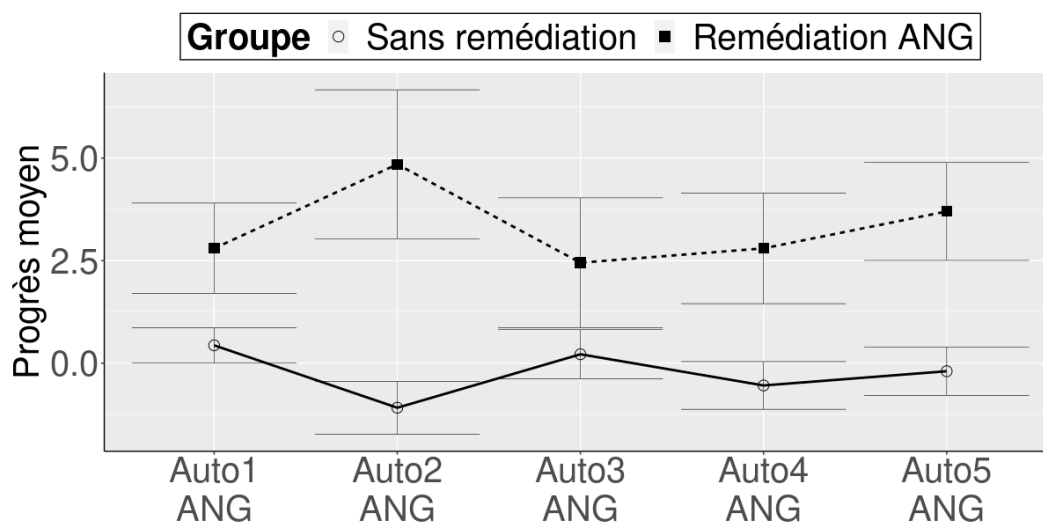


Figure 6. Progrès en autonomie d'apprentissage par dimension (anglais)

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Dim. 1	2.8	0.43	0.0274	*
Dim. 2	4.85	-1.09	0.00053	***
Dim. 3	2.45	0.22	0.138	
Dim. 4	2.8	-0.54	0.020	*
Dim. 5	3.7	-0.20	0.00691	**

Tableau 5. Différences entre les groupes et significativité en autonomie d'apprentissage (anglais)

Le Tableau 6 permet de faire le parallèle entre les données de l'anglais et celles de l'espagnol. On y observe un patron pour l'espagnol qui semble numériquement assez similaire à celui observé en anglais (le degré d'autonomie d'apprentissage chez le groupe ayant suivi la remédiation semble plus important que chez le groupe ne l'ayant pas suivie). Les modèles calculés pour chacune de ces dimensions confirment ce patron général pour la dimension 5 uniquement (s'auto-évaluer), en termes de significativité. Les différences observées pour les dimensions 3 (trouver des ressources/activités/exercices) et 4 (suivre ses avancées) ne sont, elles, pas significatives – probablement du fait de la présence d'une valeur extrême (basse) dans les données, déjà peu nombreuses, pour le groupe ayant suivi la remédiation en espagnol. À nouveau, la nécessité de recourir à l'inférence statistique est ici illustrée : la tendance observée sur ce questionnaire d'autonomie d'apprentissage en espagnol devrait être confirmée avec plus de données.

	Progrès moyen (remed)	Progrès moyen (sans remed)	Valeur P	Signif.
Dim. 1	2.67	1.43	0.494	
Dim. 2	-0.5	-0.67	0.929	
Dim. 3	4	0.48	0.0847	.
Dim. 4	2.5	-0.57	0.0673	.
Dim. 5	4.17	0.54	0.0344	*

Tableau 6. Différences entre les groupes et significativité en autonomie d'apprentissage (espagnol)

En résumé, de manière générale, on observe un progrès plus important au sein du questionnaire d'autonomie d'apprentissage pour le groupe ayant suivi le dispositif de remédiation, et c'est principalement le cas pour le groupe d'anglais. En espagnol, la tendance observée, similaire, n'est pas significative statistiquement et invite à une collecte plus importante de données pour s'en assurer. Les résultats exposés dans la présente publication sont donc préliminaires ; nous espérons obtenir des tendances plus affirmées au terme des deux prochaines sessions de collecte de données (au sein des cohortes d'étudiant·e·s 2022-2023 et 2023-2024), et ainsi pouvoir discuter ces tendances plus finement.

Conclusion

Dans cette contribution, nous avons souhaité revenir sur la façon dont nous avons choisi d'évaluer l'efficacité d'un dispositif de remédiation anglais/espagnol sur le développement de compétences langagières et d'autonomie d'apprentissage dans ces deux langues chez des étudiant·e·s en première année de licence LEA. Nous avons opté pour une démarche « pré-tests / post-tests » avec une prise de mesures pour chacun de ces deux indicateurs avant et après le déroulé du cours et auprès d'étudiant·e·s ayant suivi le cours et d'étudiant·e·s ne l'ayant pas suivi.

Au terme d'une étude pilote conduite auprès d'une première cohorte d'étudiant·e·s, nous avons trouvé que, de manière générale, le dispositif de remédiation anglais/espagnol a un impact plutôt positif sur le développement des compétences d'autonomie d'apprentissage dans ces deux langues. De manière plus fine, au sein de chacune des 5 dimensions composant ces compétences, les résultats sont plus difficilement appréciables en raison des faibles effectifs composant notre échantillonnage (la plupart de ces dimensions sont affectées positivement par le dispositif en anglais mais il est difficile de se prononcer pour l'espagnol). Concernant le niveau de langue, le dispositif de remédiation anglais/espagnol ne semble pas avoir d'impact – que ce soit de manière générale ou de manière plus fine au sein des trois compétences langagières observées (compréhension écrite/orale et grammaire). Seul le groupe n'ayant pas suivi le dispositif de remédiation semble témoigner de progrès

notables en compréhension de l'écrit en anglais par rapport au groupe ayant suivi la remédiation anglais ; mais ici aussi, il serait intéressant d'avoir plus de participant·e·s.

Cette démarche quantitative et inférentielle a ainsi permis de comparer les scores obtenus par chacun·e de nos participant·e·s aux deux passations des tests de langues et du questionnaire d'autonomie d'apprentissage de manière objective et systématique – non pas à l'œil nu en s'arrêtant aux valeurs elles-mêmes ni en se limitant à des comparaisons de moyennes entre les groupes. Nous pensons ainsi que cette démarche méthodologique est prometteuse pour mesurer l'efficacité de notre dispositif de remédiation anglais/espagnol, à savoir son impact sur le développement de compétences langagières et d'autonomie d'apprentissage dans ces deux langues pour lequel il a été conçu. Dans cette optique, nous comptons reproduire l'intégralité de cette procédure méthodologique auprès des cohortes d'étudiant·e·s inscrit·e·s en remédiation anglais/espagnol des deux prochaines années universitaires afin de pouvoir s'appuyer sur un vivier de participant·e·s et de données plus conséquent. La collaboration et la mutualisation des compétences de chacun d'entre nous pour réaliser cette étude quantitative nous ont donc permis d'atteindre un haut niveau de rigueur scientifique, qui nous paraît essentiel pour la recherche en didactique des langues.

BIBLIOGRAPHIE

- Abd-El-Fattah, S. M. (2010). Garrison's Model of Self-Directed Learning: Preliminary Validation and Relationship to Academic Achievement. *Spanish Journal of Psychology*, 13(2), 586-596. <https://doi.org/10.1017/S1138741600002262>
- Ceylan, N. O. (2015). Fostering learner autonomy. Dans I. H. Mirici, I. H. Erten, et H. Oz (dir.), *Proceedings of the 1st Globelt Conference on Teaching and Learning English as an Additional Language*, Amsterdam: Elsevier Science Bv, 85-93.
- Cheng, S.-F., Kuo, C.-L., Lin, K.-C., et Lee-Hsieh, J. (2010). Development and preliminary testing of a self-rating instrument to measure self-directed learning ability of nursing students. *International Journal of Nursing Studies*, 47(9), 1152-1158. <https://doi.org/10.1016/j.ijnurstu.2010.02.002>
- Fisher, M., et King, J. (2010). The Self-Directed Learning Readiness Scale for nursing education revisited: A confirmatory factor analysis. *Nurse Education Today*, 30(1), 44-48. <https://doi.org/10.1016/j.nedt.2009.05.020>
- Holec, H. (1990). Qu'est-ce qu'apprendre à apprendre ? *Mélanges Pédagogiques*, 75-87.
- Holec, H. (1994). L'apprenant autonome : Quelques repères conceptuels. *Babylonia, revue pour l'enseignement et l'apprentissage des langues*, 2, 10-14.
- Kuznetsova, A., Brockhoff, P.B. et Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26. <https://doi.org/10.18637/jss.v082.i13>.
- Manoïlov, P. (2019). *Les acquis des élèves en langues vivantes étrangères* (Cnesco).
- Murase, F. (2015). Measuring Language Learner Autonomy: Problems and Possibilities. Dans C. J. Everhard et L. Murphy (dir.), *Assessment and Autonomy in Language Learning*, Basingstoke: Palgrave Macmillan, 35-63. https://doi.org/10.1057/9781137414380_3
- Paris, J. (2015). *Lumière sur le développement de la production de langage non-littéral en L2. Pour une comparaison avec l'acquisition des langues maternelles* [thèse de doctorat, Sorbonne Nouvelle]. <https://tel.archives-ouvertes.fr/tel-01522893>
- Posit Team (2022). RStudio: Integrated Development Environment for R (version 2022.12.0.353). Posit Software, PBC, Boston, MA. <http://www.posit.co/>.
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Tassinari, M. G. (2010). *Autonomes Fremdsprachenlernen : Komponenten, Kompetenzen, Strategien*, Bern: Peter Lang.
- Tassinari, M. G. (2015). Assessing Learner Autonomy: A Dynamic Model. Dans C. J. Everhard & L. Murphy (dir.), *Assessment and Autonomy in Language Learning*, Basingstoke: Palgrave Macmillan, 64-88. https://doi.org/10.1057/9781137414380_4
- Tassinari, M. G. (2016). Assessment for learning; assessment for autonomy. Dans C. Gitsaki & C. Coombe (dir.), *Current issues in language evaluation, assessment and testing. Research and practice* (Cambridge Scholars Publishing, p. 118-136).
- Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21(12), 1-20. <http://www.jstatsoft.org/v21/i12/>.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York. <https://ggplot2.tidyverse.org>.
- Wickham, H. et Bryan, J. (2022). *readxl: Read Excel Files* [R package version 1.4.1]. <https://CRAN.R-project.org/package=readxl>.
- Wickham, H., François, R., Henry, L. et Müller, K. (2022). *Dplyr: A grammar of Data Manipulation* [R package version 1.0.10]. <https://CRAN.R-project.org/package=dplyr>.
- Williamson, S. N. (2007). Development of a self-rating scale of self-directed learning. *Nurse Researcher*, 14(2), 66-83. <https://doi.org/10.7748/nr2007.01.14.2.66.c6022>