

Innover dans l'évaluation de la production orale en langue étrangère. Étude exploratoire sur l'évaluation de l'allemand en Suisse romande

*Innovating in the assessment of oral production in a foreign
language. Exploratory study of the assessment of German in
French-speaking Switzerland*

Véronica Sánchez Abchi – veronica.sanchez@irdp.ch – <https://orcid.org/0000-0002-1544-9295>

IRD P Institut de Recherche et de documentation pédagogique, Suisse

Sophie Sieber Meylan – sophie.sieber@edu-vd.ch – <https://orcid.org/0009-0004-0711-5975>

Haute École Pédagogique Vaud, Suisse

Alina Matei – alina.matei@irdp.ch – <https://orcid.org/0000-0002-2630-5633>

IRD P Institut de Recherche et de documentation pédagogique, Suisse

Pour citer cet article : Sánchez Abchi, V., Sieber Meylan, S. et Matei, A. (2024). Innover dans l'évaluation de la production orale en langue étrangère. Étude exploratoire sur l'évaluation de l'allemand en Suisse romande. *Évaluer. Journal international de recherche en éducation et formation*, 10(2), 23-42. <https://doi.org/10.48782/e-jiref-10-2-23>

Résumé

La présente étude a pour but d'explorer le processus d'élaboration et de validation de contenu de tâches d'évaluation de la production orale en allemand langue étrangère chez des enfants de Suisse romande.

Les tâches, conçues pour évaluer les objectifs correspondants du plan d'études romand (PER), ont la particularité d'être des tâches numériques qui peuvent être réalisées de manière autonome par plusieurs élèves en même temps. Les tâches ont été analysées et validées par deux groupes d'expert·es – enseignant·es et didacticien·nes - puis prétestées sur un échantillon de 212 élèves de toute la Suisse romande, ce qui nous a permis d'évaluer l'impact des aspects numériques et de mieux définir les caractéristiques des tâches.

Les défis encore ouverts sont discutés, ainsi que l'apport des tâches en vue de les mettre à disposition des enseignant·es, sur une plateforme romande sur l'évaluation, en ligne (PistEval).

Mots-clés

Évaluation langues étrangères, évaluation production orale, évaluation numérique, validation de contenu

Abstract

The present study aims to explore the process of elaboration and content validation of tasks to assess oral production in German as a foreign language, produced by children from the French-speaking part of Switzerland.

The tasks designed to evaluate the corresponding objectives of the plan d'études romand, have the particularity of being digital tasks, which can be passed in autonomy and with several pupils at the same time. The tasks were analyzed and validated by two groups of experts - teachers and didacticians - and then pre-tested by a sample of 212 students from all over the French speaking part of Switzerland, in order to evaluate the impact of the digital aspects and to better define the characteristics of the tasks.

The challenges still open are discussed, as well as the contribution of the tasks to make them available to teachers, on an open platform of assessment resources (PistEval).

Keywords

Foreign language assessment, oral production assessment, digital assessment, content validation

1. Introduction

Cet article vise à explorer la pertinence et l'adéquation de deux tâches d'évaluation de production orale en allemand langue étrangère, qui ont été conçues pour des élèves de 8^e (fin du deuxième cycle de la scolarité obligatoire, niveau attendu A1) en Suisse romande. L'expérience a été développée dans le cadre d'un projet qui vise à fournir aux enseignant·es des cantons de Suisse romande des ressources d'évaluation validées dans différentes disciplines scolaires¹. Dans ce contexte, l'Institut de Recherche et de Documentation Pédagogique (IRDp) a été mandaté par les autorités éducatives pour coordonner ce processus de validation et de mise à disposition de tâches, sur la plateforme romande PistEval².

Dans les étapes précédentes, le projet s'est concentré sur la mise à disposition de tâches d'évaluation validées en mathématiques et en français L1 (Marc *et al.*, 2021 ; Sánchez Abchi *et al.*, 2016 ; Sánchez Abchi *et al.*, 2022 ; Roth et Ruf, 2023). Plus récemment, la nécessité d'intégrer des tâches d'évaluation de l'allemand L2³ a été soulevée, en donnant la priorité aux tâches qui permettraient d'évaluer les compétences orales chez les élèves de 8^e (11-12 ans).

Si, pour les étapes précédentes du projet, la procédure est partie de l'analyse de ressources fournies par les cantons romands⁴, pour les tâches d'évaluation de la production de l'oral qui font l'objet de cette contribution, la procédure a été différente. L'évaluation de la production orale en allemand ne fait pas toujours partie des compétences évaluées dans les épreuves cantonales de 8^e année en Suisse romande, pour des raisons notamment liées à des défis de gestion de temps et d'infrastructure disponible. Dans ce contexte, notre équipe a été sollicitée pour élaborer des tâches d'évaluation de la production orale, qui puissent ensuite être validées et mises à disposition pour les enseignant·es en Suisse romande.

Le projet se déroulant à un moment où l'éducation numérique est devenue une priorité en Suisse romande (CIIP, 2018), le besoin d'outils d'évaluation numérique s'est également fait sentir ; d'autant plus qu'ils pourraient contribuer à lever certaines contraintes de la passation en la facilitant. Il a ainsi été décidé de concevoir les tâches d'évaluation orale dans un format numérique, et la tablette s'est avérée le dispositif le mieux adapté pour ce faire.

La mise à disposition de tâches suppose, en premier lieu, un processus de validation. La validité - qui fait référence à la cohérence entre ce que la tâche prétend mesurer et ce qu'elle

¹ Projet EPROCOM : <https://www.irdp.ch/institut/eprocom-banque-items-3673.html>. Projet coordonné par l'Institut de recherche et de documentation pédagogique (IRDp), qui est la structure scientifique permanente de la Conférence intercantonale de l'instruction publique et culture Suisse Romande et Tessin.

² Les *Pistes pour l'évaluation* (PistEval), un site Internet sur lequel sont mises à disposition des enseignant·es romand·es des ressources évaluatives pour la 8^e année (élèves de 11-12 ans) en français et en mathématiques. Elles visent à les soutenir dans l'élaboration de leurs évaluations, qu'elles servent à des fins diagnostiques, formatives ou certificatives (Roth et Ruf, 2023).

³ La référence à L1, L2, etc. renvoie généralement à l'ordre d'acquisition de la langue. La langue 1 (L1) serait la première langue (ou les premières langues dans le cas d'un contexte bilingue) acquise par un locuteur ou une locutrice. La L2 est une langue « apprise » plus tard au cours du développement. Toutefois, dans le PER, la terminologie est utilisée de manière particulière. La L1 est la langue de scolarisation (ici le français), tandis que les L2 et L3 sont les langues nationales ou étrangères apprises à l'école primaire. Dans le contexte qui nous occupe, l'allemand est la L2 et l'anglais la L3. (CIIP, 2011, PER: Introduction Domaine des Langues).

⁴ La Suisse romande est formée par sept cantons : Berne (partie francophone), Fribourg (partie francophone), Jura, Genève, Neuchâtel, Valais et Vaud.

mesure effectivement - peut comporter différents aspects, parmi lesquels il y a la validité de contenu et la validité de mesure (Laveault et Grégoire, 2002). La validité de mesure implique d'analyser si l'outil de mesure utilisé permet de vérifier/contrôler réellement ce que l'on attend, à un niveau donné. Il s'agit souvent de démontrer la validité de l'utilisation de tests à l'aide de procédures statistiques complexes (avec des instruments comme la théorie des réponses aux items, TRI), méthodes qui sont prédominantes dans l'évaluation de la validité des tests à grande échelle (Gérard, 2005). Ces procédures sont moins applicables dans le contexte de l'évaluation en classe (Giraldo, 2020). Dans ce contexte, la notion de validité de contenu prend tout son sens. La validité de contenu consiste à examiner si la tâche et ses différentes composantes permettent d'évaluer de manière correcte l'objet que l'on souhaite évaluer : les compétences, les tâches ou le contenu stipulés dans un programme d'études. Il s'agit d'examiner le lien entre le curriculum et l'évaluation (Bonner, 2013). Dans notre cas, la validation de contenu suppose d'analyser si les tâches permettent d'évaluer les objectifs d'apprentissages inscrits, pour ce niveau de scolarité, dans le Plan d'études romand (PER, CIIP, 2001). La validation de contenu implique également d'observer si les différents éléments de la tâche (les consignes, les modalités de présentation de la tâche, les contraintes de temps et les modalités de réponse) ont un impact sur ce que l'on souhaite évaluer. La validation de contenu s'appuie sur les jugements de spécialistes du domaine.

Dans le cadre de notre travail, compte tenu des contraintes méthodologiques et de notre intérêt pour les aspects didactiques de la tâche, c'est la validation de contenu qui sera retenue. En effet, si les tâches d'évaluation ont bien été prétestées auprès des élèves dans notre expérimentation, c'est toujours dans le but de réaliser une validation de contenu, ce qui permet d'apporter des ajustements et d'améliorer les conditions d'évaluation.

Cette contribution vise ainsi à présenter le processus d'élaboration et de validation de contenu de deux tâches informatisées, ayant pour objectif d'évaluer la production en allemand langue étrangère à la fin du deuxième cycle de scolarité obligatoire, en Suisse romande.

Pour ce faire, nous partirons de l'analyse de la littérature disponible, qui a servi de base à l'élaboration des tâches. Ensuite, nous présenterons le processus de validation de contenu, qui a consisté, d'une part, à l'analyse des tâches par des expert·es et, d'autre part, à leur passation. Dans le cadre de ce travail, nous souhaitons répondre à deux questions de recherche en lien avec le processus de validation, et qui touchent à l'adéquation de tâches et à la comparabilité de la difficulté de deux tâches

Notre première question concerne ainsi l'adéquation de deux tâches pour évaluer la production de l'oral dans le contexte romand. De manière concrète, nous nous demandons si ces tâches permettent d'évaluer les *objectifs d'apprentissage* pour l'allemand L2, proposés pour ce niveau de scolarité dans le plan d'études romand (PER, 2000), ainsi que les descripteurs du Cadre Européen Commun de Référence pour les Langues (Conseil de l'Europe, 2001) pour le niveau A1, qui sont à la base des objectifs formulés dans le PER. Pour répondre à cette première question, les tâches seront soumises à un processus de validation qualitative, que nous détaillons dans la méthodologie (voir la section 2).

La seconde question, qui s'inscrit également dans la procédure de validation de contenu, vise à déterminer s'il existe des différences de difficulté entre les deux tâches mobilisées : les deux tâches sont-elles similaires en termes de difficulté ? Pour répondre à cette question,

les tâches seront prétestées de manière expérimentale auprès des élèves. Une analyse statistique sera effectuée sur la base de leurs résultats.

1.1. L'évaluation de la production orale : que retenir des expériences précédentes pour l'élaboration des tâches

L'évaluation de la production orale en langue seconde (L2) a été comparativement moins développée que l'évaluation d'autres compétences et, pour certaines épreuves internationales par exemple, son développement est relativement récent (Fulcher, 2015). En ce qui concerne l'évaluation de l'allemand comme langue étrangère, dans le contexte suisse, il existe peu d'études qui se sont penchées sur la question. La recherche de Peyer *et al.* (2023) constitue un travail pionnier dans ce sens, mais elle s'intéresse fondamentalement à l'évaluation dans le contexte de la classe. La perspective de l'évaluation externe de compétences en production orale en allemand a été moins abordée dans la littérature (Sieber *et al.*, 2015).

Dans les études sur l'évaluation de la production orale, la conception des tâches d'évaluation, ainsi que la définition du construit à évaluer, ont fait l'objet d'une attention particulière. Dans le cadre d'une perspective communicative et actionnelle⁵ (Conseil de l'Europe, 2001; Piccardo & North, 2019), l'évaluation de l'oral implique la mobilisation de tâches pertinentes qui permettent d'apprécier et juger les compétences de manière globale, et pas seulement certains aspects qui y sont liés. Ainsi, ces dernières décennies, il y a eu une évolution dans la conception des tâches d'évaluation de la production de l'oral (PO) : dans le passé, l'oralité pouvait, par exemple, être évaluée à l'aide d'une activité de lecture à haute voix, en focalisant l'attention seulement sur les aspects phonétiques. Maintenant, ce n'est plus le cas. Il s'agit, à l'heure actuelle, de proposer des tâches de production qui permettent d'évaluer réellement la compétence communicative et les ressources mobilisées à cette fin (Isaacs, 2016 ; Luoma, 2004). Les tâches que l'on trouve actuellement dans les évaluations à grande échelle visent ainsi la production - spontanée ou avec un temps de préparation - de discours différents, en interaction ou en continu, sur des thématiques qui peuvent être choisies ou imposées, etc. (North, 2005).

En outre, de nombreuses études (voir synthèse dans Luoma, 2004) ont examiné les différents facteurs pouvant avoir un impact sur la difficulté d'une tâche d'évaluation de la PO. Ces facteurs peuvent être les matériels mobilisés pour produire la tâche (images, vidéos, éléments pour construire une histoire, etc.), l'incidence du temps de planification, la complexité cognitive ou la familiarité avec la tâche. Bien que les études montrent parfois des résultats différents, en fonction des conditions d'expérimentation, de manière générale, le nombre d'informations ou d'évènements à inclure dans une production ainsi que la complexité de ces éléments pourraient influencer la difficulté.

Par ailleurs, le jugement de la qualité et la notation des productions elles-mêmes suscitent également des questions diverses (Brian, 2005). En effet, les études qui se sont concentrées sur les aspects liés à la correction se sont intéressées à différents aspects : le poids de la précision — accuracy — de la langue dans la tâche (Chavez, 2007), le rôle des aspects

⁵ Dans cette perspective, l'apprentissage des langues implique la réalisation d'"actions" concrètes qui mobilisent des activités communicatives de réception, de production, d'interaction et de médiation, dans un but spécifique. Cette perspective met l'accent sur l'autonomie et l'authenticité des ressources, des thèmes et des pratiques.

grammaticaux, l'influence de la prononciation et de l'intonation (Mi Sun Park, 2020), ainsi que l'importance de la fluidité du discours, face à l'intelligibilité (Isaacs, 2016), etc. Ces aspects, qui sont considérés de manière articulée et contextualisée, montrent la complexité de proposer des tâches d'évaluation de la production orale appropriées, pertinentes et valides.

Plus récemment, l'apport des nouvelles technologies a permis d'élargir les modalités d'évaluation de la production orale. Ainsi, la possibilité d'enregistrer les productions rend possible une pondération ultérieure. En outre, la modalité d'évaluation « semi-directe » (Quian, 2009) consiste en une évaluation assistée par ordinateur dans laquelle le/la candidat·e au test donne des réponses dans un dispositif d'enregistrement sans qu'un·e interlocuteur/interlocutrice humain ne soit présent. C'est cette modalité « semi-directe », sans phase d'interaction ni relance de la machine, qui a été retenue dans ce travail. Toutefois, si les études montrent l'apport des outils technologiques pour soutenir le processus d'apprentissage, leur contribution pour l'évaluation a été moins étudiée (Roussel, 2020 ; 2011). En plus, les systèmes d'évaluation linguistique assistée par ordinateur les plus connus n'incluent pas l'évaluation de la production orale (c'est le cas par exemple de DIALANG⁶), tandis que les expériences existantes⁷ s'adressent principalement à un public adulte (Criado de Diego, 2016). Dans notre cas, les tâches sont destinées à des enfants, de sorte que leur autonomie et leur motivation peuvent être radicalement différentes de celles d'un public adulte et, en ce sens, il n'est pas possible d'assurer que l'impact d'une épreuve autogérée sera le même dans notre expérience.

En outre, il faut prendre en considération tous les enjeux liés aux aspects informatiques de la tâche : il ne s'agit pas simplement d'une transformation en format numérique d'items originellement conçus pour format papier crayon. Dans cette optique, de nombreuses études se sont attachées à analyser les changements et les contributions apportés par les ordinateurs à l'évaluation des langues (Jamieson, 2005). Bessonneau *et al.* (2015) se sont intéressés justement à la comparabilité de la mesure obtenue avec des supports différents. En comparant des items qui évaluaient les compétences de base en français et en mathématiques, sur format papier-crayon et sur ordinateur, les auteurs ont observé que des variables différentes pouvaient avoir une conséquence sur la réussite sur un format ou sur l'autre. Ces facteurs étaient 1) la structure de l'item (la longueur des textes proposés, le nombre de documents, le type de documents, la mise en page et l'ergonomie intrinsèque), 2) le type d'activités mises en jeu (raisonnement nécessitant des étapes intermédiaires et capacité à « naviguer » dans le support numérique) et 3) les contraintes liées à la spécificité du support (utilisation d'outils différents : le brouillon, le tableur, le grapheur, etc.).

En ce qui concerne l'évaluation de la PO en langue étrangère, il n'y a pas d'expérience de recherche, à notre connaissance, qui permette de tirer des conclusions claires sur les différences de l'évaluation de cette compétence sur ordinateur ou en présence des évaluateurs. Toutefois, d'autres études se sont intéressées à l'impact que le format numérique peut avoir sur le construit de l'évaluation, en signalant que parfois le format

⁶ DIALANG *Language Assessment System* est un programme d'auto-évaluation en langues qui évalue le niveau de compétence selon le Cadre européen commun de référence pour les langues (CECRL). Les compétences évaluées sont la compréhension de l'écrit et la compréhension de l'oral (voir description dans Conseil de l'Europe, 2001). DIALANG est disponible pour plusieurs langues européennes.

⁷ Notamment le TOEFL IBT (*Test of English as a Foreign Language- Internet based test*) ou SIELE (*Servicio internacional de la evaluación en lengua española*).

peut changer ce qu'on est en train d'évaluer. Pour les mathématiques, par exemple, Grapin et Sayac (2022) ont observé que l'absence de travail avec la tablette dans l'enseignement peut avoir un impact sur la manière de réaliser une tâche d'évaluation sur tablette.

Dans notre travail, seul le format numérique sera testé, sans comparaison possible avec un support « traditionnel ». Nous essayerons pourtant de prendre en considération l'impact que le contexte numérique peut avoir dans le processus de validation de contenu des tâches de production orale sur support numérique.

Les différentes études présentées dans cette section ont servi de base pour l'élaboration et la validation des tâches dont il est question dans la présente contribution. De plus, les résultats observés dans ces enquêtes nous permettront d'analyser et de discuter nos propres observations, comme nous le verrons dans les sections suivantes.

2. Méthodologie

2.1. *Élaboration de tâches*

Pour l'élaboration des tâches, nous avons analysé les tâches existantes dans les examens disponibles (tâches d'évaluation du *Goethe Institut*, tâches d'évaluation des examens internationaux d'autres langues, tâches d'évaluation de la PO provenant d'épreuves cantonales pour la fin du cycle 2 et du cycle 3), ainsi que les tâches de production orale disponibles dans les moyens d'enseignement officiels (*Der Grüne Max* et *Junior*⁸).

Après avoir analysé la littérature spécialisée et le matériel existant, nous avons procédé à la conception des tâches en tenant compte, d'une part, des objectifs d'apprentissage du PER pour l'allemand langue étrangère, pour la fin du cycle 2, et, d'autre part, des descripteurs du CECR pour le niveau concerné (A1). Nous avons ensuite défini les spécifications de la tâche : la compétence à évaluer, la durée, la thématique appropriée, ainsi que les consignes.

2.2. *Caractéristiques des tâches*

Dans cette première expérimentation, deux tâches de production de l'oral en continu ont été proposées aux élèves. La modalité "en continu" consiste à s'exprimer d'une manière continue et ininterrompue, sans interaction, ce qui nous a paru plus simple pour le développement sur format numérique (voir annexe 1).

Les deux tâches devraient permettre d'évaluer l'objectif du PER « *Présentation de soi, de sa famille ou d'une tierce personne (nom, âge, provenance, domicile, école, emploi du temps, hobbies...)* », ce qui correspond au descripteur du CECR (A1.2) « *[l'apprenant·e] peut produire des expressions simples isolées sur les gens* ».

La première tâche (« Nouvelle école ») demandait à l'élève de se présenter et de présenter son école, dans le but de participer à un concours. Pour la deuxième tâche (« Au Parc »), l'élève devait se mettre dans la peau de quelqu'un qui s'est perdu dans un parc d'attractions et qui doit donner des informations personnelles et décrire la ou les personnes avec lesquelles elle ou il a fait l'excursion. Comme on le verra dans la procédure présentée ci-

⁸ <https://www.der-gruene-max.ch/> et https://junior-deutsch.ch/interactive_items_cat/j7-einheit-1/

dessous, les élèves devaient passer préalablement des tâches de compréhension de l'oral (CO) en allemand, qui pouvaient être liées à la thématique de la PO.

Les tâches étaient présentées aux élèves sur tablette. Les élèves pouvaient écouter de manière indépendante, autant de fois que nécessaire, les consignes et la contextualisation de la tâche, ce qui permettait d'enlever la difficulté potentielle de la lecture. Ces consignes sont données en français, qui est la langue de scolarisation, et non en allemand - langue cible - afin de s'assurer de leur bonne compréhension (Karges *et al.*, 2021). En plus, les élèves étaient libres de décider du moment pour lancer l'enregistrement de leurs productions. Les élèves pouvaient écouter leur production, décider de la valider et, si ce n'était pas le cas, choisir de continuer les essais. D'après les instructions données oralement, les élèves pouvaient réaliser jusqu'à trois enregistrements, sans contrainte de temps.

Afin d'ajuster les tâches et de juger l'impact des aspects techniques, les tâches ont été soumises, de manière expérimentale, à un échantillon de 14 élèves, avant la passation à grande échelle. Les élèves qui ont expérimenté ces premières versions ont également été interviewés par rapport à cette expérience. Les informations retenues lors de la passation expérimentale et lors des entretiens avec les élèves ont été utilisées ensuite pour améliorer les tâches. Des aspects liés à l'appropriation du dispositif d'enregistrement, de réception et d'interprétation autonome de consignes ont été également pris en considération et les observations ont été rapportées aux groupes de validation, dont nous parlerons ci-dessous, dans la section 2.5.

2.3. Caractéristiques de la passation

Les deux tâches retenues (nommées « Au Parc » et « Nouvelle école ») ont été administrées à un échantillon de 212 élèves de 8^e année (11-12 ans), provenant de 56 classes de toute la Suisse romande (environ 8 classes par canton). Les conditions de passation ont été identiques pour chaque élève. La passation des tâches de PO se faisait à la suite d'une série de tâches de CO en allemand sur tablette, qui étaient passées par la totalité des élèves de 56 classes. Avant la passation des tâches d'évaluation, les élèves ont réalisé des activités de « prise en main » pour se familiariser avec les fonctionnalités de l'application numérique. En ce qui concerne l'appropriation de la fonction « enregistrement », presque la totalité des élèves a réalisé les opérations sans problème et très rapidement. Seulement deux élèves ont eu besoin de plus de temps pour se familiariser avec cette fonctionnalité (plus de deux minutes).

Après avoir finalisé la partie de CO, 4 élèves de chaque classe, sélectionnés au hasard, recevaient une notification sur leurs tablettes pour réaliser une des deux tâches de PO. En tout, 104 élèves ont passé la tâche « Au Parc » et 108 élèves ont réalisé la tâche « Nouvelle école ».

Les élèves concernés sortaient de la salle de classe pour réaliser la tâche dans une salle séparée pour ne pas gêner les autres élèves qui poursuivaient le test de CO. Les élèves pouvaient réaliser une des deux tâches de PO de manière indépendante. Il n'y avait pas de limite de temps, mais les élèves ont passé en moyenne une dizaine de minutes entre la préparation et l'enregistrement des productions. Un adulte était toutefois présent dans la salle pour répondre à d'éventuelles questions techniques ou simplement pour rassurer les élèves.

2.4. Caractéristiques de la grille d'évaluation

Une grille analytique a été utilisée pour évaluer les PO des élèves. Celle-ci (voir l'annexe 6.2) prenait en considération trois critères liés :

- a) au contenu : ce critère prenait en considération la présence de tous les éléments de contenu et les informations sollicitées dans la consigne ;
- b) à l'étendue du vocabulaire et de la correction grammaticale : ce critère consistait à vérifier si le vocabulaire disponible était suffisant pour effectuer la tâche demandée, si les erreurs entravaient ou non la communication et si les structures utilisées étaient adéquates ;
- c) à la fluidité et la prononciation : ce critère vérifiait si la prononciation permettait ou non la compréhension, ainsi que la fréquence et la longueur des pauses.

La grille présentait trois valeurs correspondantes à différents niveaux de performance : le niveau le plus bas : score 0 ; niveau intermédiaire : score 1 ; le niveau le plus élevé : score 2.

Les productions des élèves ont été notées par 4 juges familiarisés avec la procédure d'évaluation de la production orale. Les transcriptions des productions servaient également de support pour la décision. En cas de désaccord entre les juges, les points problématiques ont été discutés afin de parvenir à un consensus.

2.5. Procédure de validation du contenu

La validation de contenu des tâches s'est déroulée en deux étapes. Une première étape de validation qualitative de contenu par des expert·es et une deuxième étape de comparaison des tâches.

2.5.1 Étape de validation qualitative de tâches

Pour répondre à la première question de recherche - à savoir si les deux tâches proposées sont adéquates pour évaluer la production de l'oral de manière valide dans le contexte romand - une procédure de validation de contenu, de nature qualitative, a été mise sur pied avant la passation. Comme expliqué dans l'introduction, le processus de validation de contenu suppose que des expert·es puissent analyser si les spécifications de chaque tâche (Fulcher, 2010) sont alignées avec les objectifs que l'on souhaite évaluer. Dans notre cas, il s'agit des objectifs du PER, formulés sur la base du CECR pour le niveau A1.

Cette analyse permet de fournir une évidence empirique pour la validation pour trois raisons : la première, parce que les spécificités de la tâche doivent se baser dans les habilités linguistiques du plan d'études. La deuxième, parce que l'analyse peut également mettre en évidence des problèmes en lien avec les tâches, qui impliquent que la tâche évalue des éléments qui ne sont pas présents dans le plan d'études et, par conséquent, ne constituent pas des objectifs que l'on souhaite évaluer. Et finalement, parce que, grâce au processus d'analyse, les tâches peuvent être modifiées afin de mieux refléter les objectifs du plan d'études que l'on souhaite évaluer (Fulcher, 2010; Bonner, 2013; Giraldo, 2020). Le but principal a été ainsi d'observer l'adéquation :

- des tâches pour évaluer les objectifs du PER pour l'allemand (L2) à la fin du cycle 2;
- des tâches pour évaluer un niveau A1.1 d'après les descripteurs du CECR;
- du format de la tâche et des thèmes au public cible.

La procédure consistait à soumettre les tâches à deux groupes d'accompagnement différents :

- un groupe de spécialistes en didactique de l'allemand constitué de trois formateurs et formatrices des enseignant·es de l'allemand langue étrangère. Ce groupe a réalisé une analyse des tâches qui prenait en compte à la fois les paramètres à considérer du point de vue de la littérature sur l'évaluation, les théories actuelles de l'enseignement de langues étrangères, les objectifs du PER et les descripteurs du CECR. Le contenu de la formation des enseignant·es, qui a une influence directe sur l'enseignement et l'évaluation, a été également pris en compte.
- un second groupe constitué de sept enseignant·es généralistes de 8^e année, avec plusieurs années d'expérience dans l'enseignement, chacun·e issu·e des cantons romands. Ce groupe évaluait la pertinence des tâches par rapport à ce que leurs élèves connaissent et ce qu'elles/ils sont en mesure de faire à partir de l'enseignement, en lien avec les objectifs du PER et les descripteurs du CECR. La présence de tous les cantons romands dans ce groupe a permis de garantir une certaine représentation par rapport à ce qui est enseigné en Suisse romande.

Ces deux groupes de validation devaient également se prononcer sur les autres aspects de la tâche qui pouvaient avoir une incidence sur l'évaluation (formulation de consignes, modalité de réponse, modalité de présentation de la tâche, modalités de passation).

2.5.2 Étape de comparaison des tâches

Pour la présente étude, nous souhaitons élaborer des tâches de difficulté comparable. Bien que ces deux tâches soient jugées comme équivalentes en termes de difficulté de contenu, dans une des deux tâches, les élèves étaient censés prendre en compte plus d'éléments (présentation personnelle et, en plus, caractérisation d'une autre personne). Le nombre d'éléments à prendre en compte ainsi que la proximité de la thématique de la tâche à la réalité des élèves auraient pu avoir ainsi un impact sur sa difficulté (voir synthèse dans Luoma, 2004). C'est pourquoi nous avons fait l'hypothèse que la tâche « Au parc » serait plus compliquée à réaliser que « Nouvelle école ». En plus, la thématique de la description de l'école est beaucoup plus présente et développée dans les moyens d'enseignement romand que la description, notamment physique, d'une personne. Cet aspect pouvait donc également avoir un impact sur la complexité de la tâche « Au parc ».

Pour répondre à la seconde question de recherche, concernant la difficulté des tâches, des analyses statistiques ont été réalisées, suite à la passation de deux tâches (cf. section 3.4 Caractéristiques de la passation). Pour ce faire, les productions orales des élèves ont été analysées notamment à l'aide de la grille analytique, qui a permis d'associer un score à chaque élève pour chaque tâche. Comme la grille analytique le prévoit, trois types de scores ont été utilisés : score bas (valeur 0), intermédiaire (valeur 1) et élevé (valeur 2). Ensuite, nous avons considéré les résultats des deux tâches ensemble, et nous avons utilisé une analyse des correspondances multiples qui permet une comparaison des résultats obtenus pour les deux tâches (cf. section 3.2).

3. Résultats et Discussion

Les résultats s'organisent en deux sections ci-dessous. Dans la première section (3.1) nous présentons les résultats concernant l'étape de validation qualitative des tâches.

La seconde section (3.2) présente les résultats concernant la comparaison de tâches, permettant ainsi d'avoir des informations plus achevées sur les tâches et sur la grille d'évaluation, pour compléter le processus de validation de leurs contenus.

3.1. Résultats concernant la validation qualitative des tâches

Pour répondre à notre première question — si les tâches proposées permettaient d'évaluer les *objectifs d'apprentissage* du PER - nous avons mis sur pied un système de validation qualitative du contenu des tâches, comme mentionné dans la section 2.5.1. Les deux groupes d'accompagnement impliqués dans le processus ont jugé les deux tâches valides, parce que celles-ci permettaient de bien saisir les objectifs d'apprentissage proposés dans le plan d'études. En outre, ces deux groupes ont apporté des propositions de modifications, afin de mieux tenir compte des objectifs du plan d'étude et de mieux cibler les différents aspects pouvant avoir une incidence sur le construit d'évaluation.

D'après la décision de ces deux groupes de validation, les deux tâches, avec les ajustements suggérés, permettaient de bien évaluer les objectifs du PER retenus pour la fin du deuxième cycle de scolarité obligatoire. Le processus de validation qualitative a donc confirmé que les tâches permettaient d'évaluer les objectifs du PER.

3.2. Résultats concernant la comparaison de tâches

Comme déjà précisé, pour la présente étude, nous souhaitons élaborer des tâches de difficulté comparable. Pour analyser les différents facteurs pouvant avoir un impact sur la difficulté, nous devons comparer les deux tâches lors d'une passation auprès des élèves. Ainsi, pour répondre à notre seconde question de recherche, les scores obtenus (grâce à la grille analytique) par les 212 élèves, lors de la passation des deux tâches, ont été pris en considération.

Trois variables, qui correspondent aux scores des trois critères de la grille analytique (voir la section 2.4), ont été utilisées : la variable « Contenu » (pour les scores du critère a), la variable « Vocabulaire et grammaire », désormais « VocGr » (pour les scores du critère b) et la variable « Prononciation » (pour les scores du critère c). Les trois variables sont qualitatives, avec chacune trois catégories 0, 1 et 2 (le niveau le plus bas : 0 ; niveau intermédiaire : 1 ; le niveau le plus élevé : 2, voir la section 2.4). Les corrélations de Kendall entre ces trois variables sur l'ensemble des données provenant de la passation de deux tâches sont respectivement : 0,66 (entre « Contenu » et « VocGr »), 0,62 (entre « VocGr » et « Prononciation ») et 0,44 (entre « Contenu » et « Prononciation »). Ces corrélations sont statistiquement significatives au seuil de 5% et montrent l'existence d'un lien assez important entre les trois variables provenant de la grille analytique.

Une analyse des correspondances multiples (ACM)⁹ a été effectuée sur les trois variables et l'ensemble des données. Cette analyse permet d'explorer plus en profondeur les relations entre les catégories de ces trois variables. Elle constitue une étape préliminaire à l'analyse par clustering, que nous avons réalisée afin de regrouper les élèves selon leurs performances et de comparer les résultats des deux tâches. L'ACM résume l'information contenue dans les trois variables (« Contenu », « VocGr », « Prononciation ») et génère deux nouvelles variables (ou dimensions). Celles-ci sont représentées dans la **Figure 1** par deux

⁹ Pour plus de détails sur l'ACM, voir par exemple Escofier et Pagès, 2023

axes perpendiculaires, appelés Dim1 et Dim2. La figure 1 montre que les deux premières dimensions de cette analyse expliquent environ 60 % de la variance totale des données (la somme des valeurs affichées sur les axes, soit 37,44 % et 23,31 %), grâce aux liens existant entre les trois variables initiales. Pour ce type de données, un taux de 60 % constitue un pourcentage relativement élevé, ce qui atteste de la bonne qualité de la représentation fournie par l'ACM.

Suite à l'ACM, nous avons pu mettre en évidence trois clusters, qui rassemblent, à l'aide d'un algorithme, les élèves en trois groupes d'après leur niveau de réussite. La **Figure 2** montre ces trois clusters.

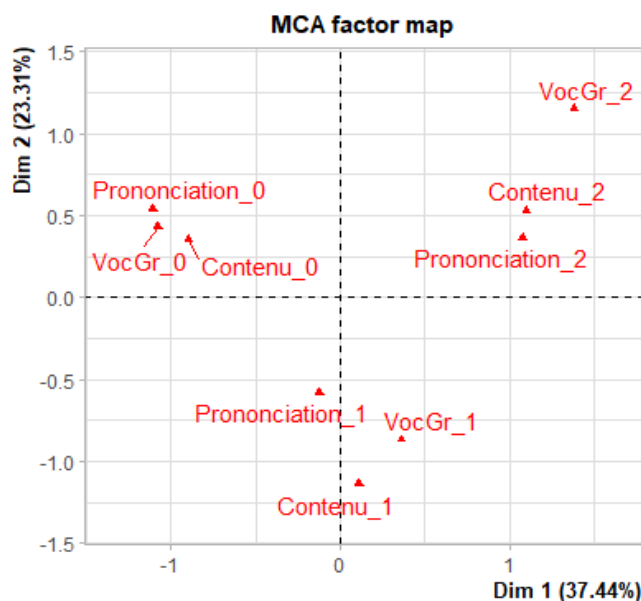


Figure 1 : Carte de l'analyse des correspondances multiples pour les variables Contenu, Étendue de vocabulaire et Prononciation

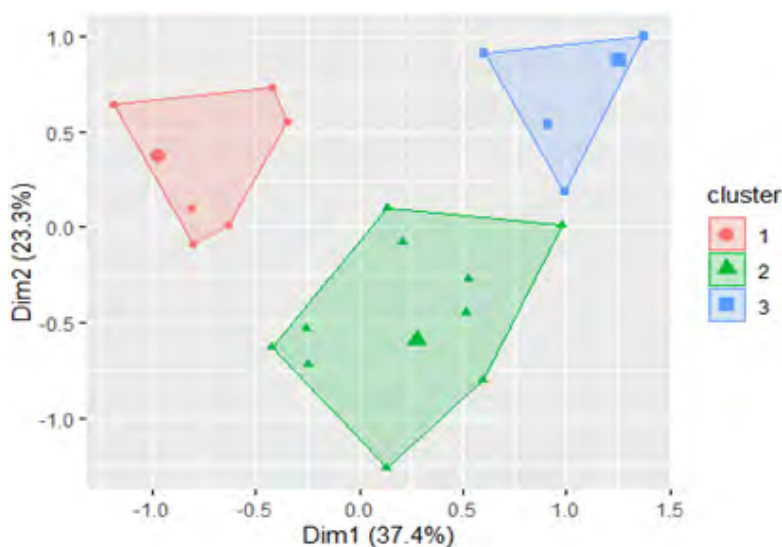


Figure 2 : Clusters construits sur la base de l'analyse des correspondances multiples

Conformément aux figures 1 et 2, l'interprétation des clusters est la suivante : le cluster 1 regroupe les élèves de performance faible (75 élèves), le cluster 2 ceux de performance intermédiaire (101 élèves) et le cluster 3 ceux de performance élevée (36 élèves). Comme déjà indiqué, cette interprétation des clusters est réalisée sur la base de l'interprétation ci-dessus de la carte montrée dans la figure 1.

En considérant ainsi les trois critères de la grille analytique ensemble et les données de l'échantillon en entier de 212 élèves, nous avons ensuite calculé le pourcentage d'élèves qui ont résolu une tâche et qui se retrouvent dans chaque cluster. La **Figure 3** présente la répartition des élèves par tâche et par cluster. Nous observons que les pourcentages qui correspondent à chaque cluster sont similaires pour les deux tâches (pour la tâche « Nouvelle école », 35 % des élèves sont dans le cluster 1 versus 36 % pour la tâche « Au parc » ; 50 % versus 45 % dans le cluster 2 et 15 % versus 19 % dans le cluster 3).

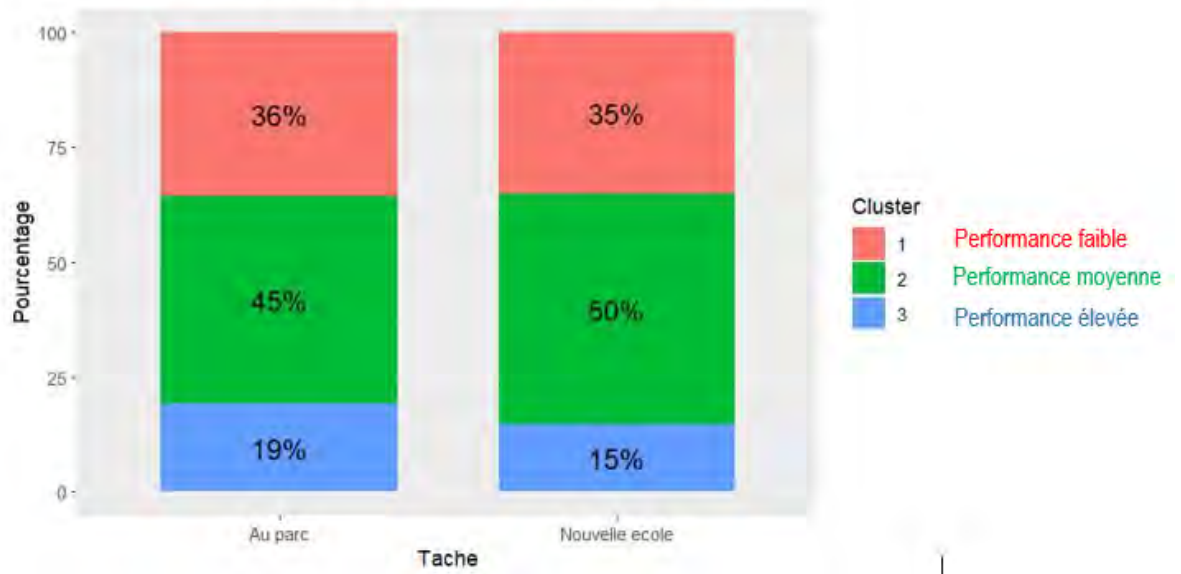


Figure 3 : Répartition par tâche et par cluster

Nous avons testé ensuite le lien entre la variable « tâche » (avec deux catégories, les deux tâches) et la variable « cluster » (avec trois catégories, les trois clusters) à l'aide d'un test d'indépendance (un test de khi-deux). Nous avons constaté qu'il n'existe pas de lien significatif entre ces deux variables au seuil de 5 %, c'est-à-dire que la répartition des élèves dans les clusters ne dépend pas de la tâche résolue.

À priori, lors de l'élaboration des tâches, et en fonction de ce qui avait été observé dans la littérature (Luoma, 2004) nous avons estimé qu'il pourrait y avoir une différence de difficulté entre les deux tâches, en lien avec deux facteurs. Le premier facteur était la quantité d'éléments de contenu à intégrer dans la production des élèves. En effet, dans les deux tâches les élèves devaient se présenter, mais dans la deuxième tâche, « Au parc », les élèves devaient en plus donner des informations sur une autre personne.

Un second facteur qui aurait pu avoir un impact sur la difficulté était la familiarité de la thématique de la tâche « Nouvelle école ». En lien avec cet aspect-là, nous avons fait l'hypothèse que la première tâche pouvait être plus simple, car elle était plus proche de la réalité des élèves et impliquait la mobilisation d'un univers et d'un vocabulaire qui était

souvent travaillé dans les moyens d'enseignement officiel¹⁰. Malgré nos hypothèses, qui auraient laissé penser que la tâche « Au parc » était plus difficile et aurait pu être moins bien réussie, les résultats obtenus suggèrent que les deux tâches sont de difficulté comparable.

Une explication possible de ces résultats serait que les facteurs de difficulté n'ont tout simplement pas assez d'impact. Une autre possibilité serait le fait que, comme les études le montrent, la difficulté d'une tâche peut être le résultat de l'interaction de plusieurs facteurs, tels que le niveau de maîtrise de la langue de la part des élèves. Cette interaction pourrait également expliquer les résultats obtenus.

En effet, les apprenant·es qui ont des difficultés avec l'allemand éprouvent des problèmes avec l'une ou l'autre tâche, sans que cela ait un impact sur les résultats. Au contraire, les apprenant·es qui sont bons en allemand peuvent parvenir à réaliser des performances adéquates indépendamment de la tâche résolue.

Il s'agissait d'une situation d'évaluation pour laquelle les apprenant·es ne pouvaient pas se préparer avant la situation d'évaluation. L'absence de planification et de préparation peut également avoir accentué l'effet de différence entre les apprenant·es, et donc éclipsé les éventuelles différences entre les tâches.

4. Conclusion

La présente étude montre la complexité liée à la conception de tâches d'évaluation valides et pertinentes pour un contexte particulier. En effet, notre procédure de validation de contenu, constituée de deux étapes - validation qualitative et procédure statistique - nous a permis d'examiner les différents facteurs concernés. L'étape de validation qualitative a été nécessaire pour mieux saisir la manière dans laquelle les tâches reflétaient les objectifs du PER et permettaient ainsi d'évaluer précisément le construit que l'on souhaitait évaluer, ainsi que l'impact des facteurs tels que la formulation de la consigne, la présentation des informations et les implications liées au support. L'étape de comparaison des tâches via la procédure statistique, après la passation, a fourni, de son côté, des informations importantes sur les caractéristiques des tâches et a conclu que les deux tâches semblaient être comparables en termes de difficulté de contenu. Cette double procédure de validation a montré que les tâches sont valides et qu'il est possible de les mettre à disposition des enseignant·es sur la plateforme romande PistEval, agrémentées d'un étayage didactique concernant leur conception et leur élaboration.

D'autre part, le fait qu'il s'agisse de tâches numériques présente d'autres aspects qui méritent d'être pris en considération.

Tout d'abord, d'après nos observations, l'interaction avec la tablette semble avoir un impact positif sur la motivation des élèves à effectuer la tâche.

Par ailleurs, le fait de pouvoir enregistrer et stocker les productions des élèves se traduit non seulement par un outil très précieux pour soutenir l'activité de pondération des textes, mais nous permet également de disposer d'un matériel utile pour des recherches futures :

¹⁰ En Suisse, à l'école obligatoire, l'allemand L2 est enseigné à l'aide d'un seul et même moyen d'enseignement officiel, qui est partagé par tous les élèves de la même année: « *Der grüne Max* » pour la 5^e et 6^e année, et « *Junior* » pour la 7^e et 8^e année. Tous les ouvrages ont été élaborés par l'éditorial Klett et adaptés aux objectifs du PER.

les différentes tentatives enregistrées offrent des pistes sur le processus de production des élèves, qui pourraient également être utilisées pour soutenir l'enseignement.

Par ailleurs, si l'on considère que l'enseignement et l'évaluation de la production orale se font très peu dans le contexte scolaire, parce que les enseignant·es trouvent que c'est très peu pratique à mettre en place du fait que la passation est individuelle et qu'elle est difficile à réaliser dans une classe avec beaucoup d'élèves, la possibilité de compter avec des tâches d'évaluation validées, qui peuvent être passées sur ordinateur de manière autonome et qui évaluent plusieurs élèves en parallèle, devrait permettre de surmonter cette difficulté. En effet, cette modalité de passation se traduit par une optimisation des ressources, autant pour l'évaluation en classe que pour les pratiques d'évaluation standardisée.

Pourtant, certains aspects de la passation restent critiques. Tout d'abord, s'il est vrai que l'enregistrement en autonomie donne aux élèves la possibilité de décider à quel moment le débiter, cette fonctionnalité ne remplace cependant pas nécessairement le temps de préparation et de planification préalable habituellement consacré à une production. En plus, l'absence de relance ou d'interaction avec un adulte peut péjorer la performance, surtout chez les enfants et les jeunes.

En outre, certains aspects liés à la nature numérique des tâches peuvent avoir un impact sur les performances des apprenant·es. Bien évidemment, les tâches ont été testées de manière expérimentale avant la passation, pour ajuster les aspects techniques. Par ailleurs, lors de la passation sur 212 élèves, les élèves devaient faire un exercice de « prise en main » pour s'appropriier les fonctionnalités, afin d'éviter que les aspects techniques - écoute autonome des consignes, fonction d'enregistrement - puissent entraver la réalisation de la tâche de production orale. Pourtant, les aspects liés à l'appropriation de l'outil et l'habitude - ou l'absence d'habitude - de travailler avec ce genre de tâches sur l'ordinateur pourraient constituer un défi pour les élèves qui devraient, idéalement, être familiarisés avec un tel usage déjà lors de la phase d'apprentissage. Dans des situations similaires, par exemple les examens sur ordinateur SIELE, les candidat·es à l'épreuve se préparent pendant quatre mois pour se familiariser avec les caractéristiques des tâches et de l'application (Criado de Diego, 2016). Même si on assume que les défis de l'application utilisée dans notre étude restent à la portée de n'importe quel·le élève de 11 ans, on ne peut pas assurer que chaque participant·e ait le même niveau de familiarité avec la problématique de la tâche.

En effet, des études sur l'évaluation en format numérique de la compréhension de l'oral en allemand ont montré que la passation en autonomie pouvait impliquer une situation de surcharge cognitive, qui péjorait les élèves en difficulté. Pour les élèves avec de faibles compétences, le soutien et l'étayage de l'enseignant·e, qui leur allège la prise de décision stratégique de déterminer quand commencer ou s'arrêter, peut-être plus bénéfique que l'autonomie (Roussel, 2020).

En ce sens, il serait intéressant d'explorer l'impact que la gestion autonome de la production orale peut avoir au niveau psychologique. Cet aspect, qui avait été soulevé par les groupes de validation, a été compensé par la présence d'un adulte dans la salle, qui pouvait par sa présence rassurer l'élève lors de la réalisation de la tâche. Pourtant, pour mieux estimer l'impact de la gestion autonome face à la tablette, il faudrait pouvoir comparer la même tâche dans des conditions de passation non numérique.

Il convient de noter que l'étude présentée ici concerne une expérience à petite échelle et que les résultats, la portée et les limites à plus grande échelle doivent encore être testés.

En attendant, nous considérons que cette expérience nous a permis, d'une part, de mettre sur pied un processus de validation innovant pour le contexte romand et, d'autre part, de mettre en lumière les aspects à prendre en compte dans le développement d'outils numériques d'évaluation dans le domaine des langues, en particulier pour la production de l'oral.

5. Références

- Álvarez, M. F. (2016). Language testing in the digital era. Dans M.-M., I. Elorza et B. García Riaza (dir.), *Technology-Enhanced Language Learning for Specialized Domains* (83–94). Routledge.
- Bonner, S. M. (2013). Validity in classroom assessment: Purposes, properties, and principles. Dans J. McMillan (dir.), *SAGE handbook of research on classroom assessment*, (87-106). Sage.
- Chavez, M. (2007). Students' and teachers' assessments of the need for accuracy in the oral production of German as a foreign language. *The Modern Language Journal*, 91(4), 537–563.
- Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP). (2010). Plan d'études romand (PER). <http://www.plandetudes.ch/>
- Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin CIIP (2018) Plan d'action Numérique [document non publié].
- Criado de Diego, C. (2016). E-Oral, una herramienta para evaluar la expresión oral y la comprensión auditiva a distancia. *Foro de Profesores de E/LE*, 12, 47-54. <https://doi.org/10.7203/foroel.1.9166>
- Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues ; apprendre, enseigner, évaluer*. Les Éditions Didier.
- Escofier, B., et Pagès, J (2023). *Analyses factorielles simples et multiples* (5e éd.). Dunod.
- Fulcher, G. (2010). *Practical language testing*. Hodder Education.
- Fulcher, G. (2015). Assessing second language speaking. *Language Teaching*, 48(2), 198–216.
- Gérard, F.-M. (2005, 24-26 octobre). *L'évaluation des compétences par des situations complexes* [communication orale]. Colloque de l'Admee-Europe, IUFM Champagne-Ardenne, Reims, France.
- Giraldo, F. (2020). Validity and classroom language testing: a practical approach. *Colombian Applied Linguistics Journal*, 22(2), 194-206.
- Goh, C. C. (2016). Teaching speaking. Dans Willy A. Renandya et Handoyo Puji Widodo (dir.) *English language teaching today: Linking theory and practice* (143–159). Springer.
- Grapin, N., et Sayac, N. (2022). *From paper-pencil to tablet-based assessment: a comparative study at the end of primary school* [communication orale]. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12).
- Isaacs, T. (2016). Assessing speaking. Dans D. Tsagari et J. Banerjee (dir.) *Handbook of second language assessment* (131–146). De Gruyter Mouton.
- Jamieson, J. (2005). Trends in computer-based second language assessment. *Annual Review of Applied Linguistics*, 25, 228–242.
- Karges, K., Barras, M., et Lenz, P. (2021). *Leseverstehen in einer Fremdsprache testen. Bericht des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit*. Institut für Mehrsprachigkeit.

- Laveault, D., et Grégoire, J. (2014). *Introduction aux théories des tests en sciences humaines*. De Boeck Université.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Marc, V. (dir), Roth, M., Ruf, I., Sánchez Abchi, V., Soussi, A., & Weiss, L. (2021, août). *EpRoCom : dispositif romand de mutualisation de tâches évaluatives* [document inédit]. IRDP.
- Mi Sun Park (2020) Rater Effects on L2 Oral Assessment: Focusing on Accent Familiarity of L2 Teachers, *Language Assessment Quarterly*, 17(3), 231-243.
<https://doi.org/10.1080/15434303.2020.1731752>
- North, B. (2005). Assessing Spoken Performance in relation to the Common European Framework of Reference. *Babylonia*, 2, 22-29.
- Peyer, E., Lüthi, G., et Ravazzini, N. (2023). Formative Beurteilung des Sprechens im DaF-Unterricht: Praktiken und Überzeugungen von Lehrpersonen der Sekundarstufe I. *Informationen Deutsch als Fremdsprache*, 50(5), 492-504. <https://doi.org/10.1515/infodaf-2023-0082>
- Piccardo, E., et North, B. (2019). *The action-oriented approach: A dynamic vision of language education* (Vol. 72). Multilingual Matters.
- Roth, M., et Ruf, I. (2023). Des ressources évaluatives pour les enseignant·es romand·es : une démarche intercantonale. *La Revue LEE*, 8. <https://doi.org/10.48325/rleee.008.08>
- Roussel, S. (2020). *Apport du numérique à l'enseignement-apprentissage des langues*. Cnesco
- Sánchez Abchi, V., De Pietro, J. F., et Roth, M. (2016). *Évaluer en français : comment prendre en compte la difficulté des items et des textes*. Institut de recherche et de documentation pédagogique (IRDP).
- Sánchez Abchi, V., Roth, M., et Matei, A. (2022). Estimer la difficulté des questions en compréhension de l'écrit en français. Vérification empirique d'un modèle théorique. *e-JIREF*, 8(1), 29-46.
<https://doi.org/10.48782/e-jiref-8-1-29>
- Sieber, J., Elmiger, D., et Wokusch, S. (2015). *Évaluation des compétences en allemand de 6^e année de la scolarité obligatoire vaudoise*. Unité de recherche pour le pilotage des systèmes pédagogiques. Département de la formation, de la jeunesse et de la culture. Canton de Vaud.

6. Annexes

6.1. Annexe 1: Tâches

Tâche PO.1 "Nouvelle École"

Spécifications de la tâche

Descripteur évalué	PER : L'élève fournit des informations essentielles sur sa famille ou sur sa classe (A1.2) CECR : Peut produire des expressions simples isolées sur les gens.
Fonctions discursives (actes de langage)	Présentation avec description
Objectif(s) à maîtriser	PER : L2 24 Présentation de soi, de sa famille ou d'une tierce personne (nom, âge, provenance, domicile, école, emploi du temps, hobbies, ...). Au cours, mais au plus tard à la fin du cycle, l'élève... fournit des informations essentielles sur sa famille ou sur sa classe. L2 26 Au cours, mais au plus tard à la fin du cycle, l'élève... se fait comprendre au moyen d'un répertoire limité d'expressions et de mots mémorisés.
Type de tâche / Format	PO en continu. Format ouvert. Sans préparation.
Support	Mots clés dans la consigne
Thématique	Présentation personnelle. École
Consignes	écrites et orales (enregistrées) dans la langue de scolarisation (français).

Consigne:

Tu participes à un concours pour lequel tu dois faire une vidéo pour présenter ton école.

- Prépare ton texte en parlant au moins des éléments 1 à 5 ci-dessous. Cela peut être vrai ou inventé.
- Enregistre-toi.
- Tu peux t'écouter et te réenregistrer si nécessaire. Attention, si tu te réenregistres, l'enregistrement précédent disparaît.
- Donne au minimum :
 - o Ton nom et 1-2 informations (*âge, hobby, langue...*).
 - o Le nom de ta classe et/ou de ton école.
 - o Ton avis sur ta classe/ton école.
 - o Au moins 3 choses que tu aimes dans ton école (*branche, salle, prof, activités extrascolaires, ...*).
 - o Un élément que tu n'aimes pas.

Tâche PO.2 Tâche "Au Parc"

Spécifications de la tâche

Descripteur évalué	PER : L'élève fournit des informations essentielles sur sa famille ou sur sa classe (A1.2). CECR : Peut produire des expressions simples isolées sur les gens.
Fonctions discursives (actes de langage)	Présentation avec description
Objectif(s) à maîtriser	PER: L2 24 Présentation de soi, de sa famille ou d'une tierce personne (nom, âge, provenance, domicile, école, emploi du temps, hobbies,...). Au cours, mais au plus tard à la fin du cycle, l'élève... fournit des informations essentielles sur sa famille ou sur sa classe. L2 26 Au cours, mais au plus tard à la fin du cycle, l'élève... se fait comprendre au moyen d'un répertoire limité d'expressions et de mots mémorisés.
Type de tâche / Format	PO en continu. Format ouvert. Sans préparation.
Support	Mots clés dans la consigne
Thématique	Présentation personnelle. École
Consignes	écrites et orales (enregistrées) dans la langue de scolarisation (français).

Consigne :

Tu es au parc d'attractions avec ta classe, mais tu ne trouves plus tes camarades. Tu t'adresses au stand d'information pour que l'on t'aide. La personne au stand parle seulement allemand.

- Prépare ton texte en tenant compte des éléments ci-dessous. Cela peut être vrai ou inventé.
- Enregistre-toi.
- Tu peux t'écouter et te réenregistrer si nécessaire. Attention, si tu te réenregistres, l'enregistrement précédent disparaît.
- Tu commences ainsi :

Hallo, ich finde meine Klasse nicht ... können Sie mir helfen ?

- Donne :
 - o Autant d'informations que possible sur toi (*nom, âge, école, ville, langue, classe, ...*).
 - o La description d'un·e prof ou d'un·e ami·e qui (*nom, âge, quelques traits physiques, téléphone...*)

6.2. Annexe 2 : Grille d'analyse

Critère/ Score	- (0)	(1)	(2)
Contenu (tâche)	La production correspond partiellement à la consigne (plusieurs informations (plus de 3) manquent) ou le message ne répond pas à la situation de communication (informations incohérentes).	La production correspond à la consigne. Il peut manquer quelques informations (1 à 3). Message cohérent en lien avec la situation de communication, malgré quelques maladresses.	La production correspond à la consigne. L'élève donne tous les éléments demandés, voire plus. Message cohérent en lien avec la situation de communication
Étendue du vocabulaire et correction grammaticale	Ne dispose pas d'un vocabulaire suffisant pour effectuer correctement la tâche et recourt à d'autres langues. Erreurs fréquentes sur les structures simples.	Peut mobiliser le lexique nécessaire à la réalisation de la tâche ou à une partie de la tâche. Certains mots peuvent manquer. Quelques erreurs de structure subsistent sans gêner la compréhension.	Maitrise bien les structures courantes, voire fait un effort pour mobiliser un lexique plus varié, sans répétition fréquente de structures identiques ou plus précis.
Utilisation correcte de la prononciation et de l'intonation. Fluidité	Souvent incorrecte, gênant fréquemment la compréhension. Pausés longues et fréquentes.	Prononciation correcte, malgré quelques erreurs. Pausés longues et fréquentes.	Prononciation correcte malgré quelques erreurs, Effort pour adopter une intonation authentique. Rythme fluide malgré quelques hésitations