

Utiliser une grille critériée pour évaluer les explications orales de ses pairs : quels fonctionnements et quels effets ?

Using rubrics for peer assessment of oral explanation skills: how does it work and what results ?

Christine Wiertz – christine.wiertz@uclouvain.be

Université catholique de Louvain - Faculté de psychologie et des sciences de l'éducation, Belgique

Benoît Blondeau – Benoit.blondeau@vinci.be

Haute Ecole Léonard de Vinci, Belgique

Eve Francotte – eve.francotte@uclouvain.be

Université catholique de Louvain - Faculté de psychologie et des sciences de l'éducation, Belgique

Benoit Galand – benoit.galand@uclouvain.be

Université catholique de Louvain - Faculté de psychologie et des sciences de l'éducation, Belgique

Stéphane Colognesi – Stephane.colognesi@uclouvain.be

Université catholique de Louvain - Faculté de psychologie et des sciences de l'éducation, Belgique

Pour citer cet article : Wiertz, C., Blondeau, B., Francotte, E., Galand, B., & Colognesi, S. (2022). Utiliser une grille critériée pour évaluer les explications orales de ses pairs : quels fonctionnements et quels effets ?, 8(2), 51-88. <https://doi.org/10.48782/e-jiref-8-2-51>

Résumé

L'évaluation de l'oral est souvent mentionnée par les enseignants comme étant un obstacle à son enseignement. Les grilles critériées et l'évaluation par les pairs sont deux leviers identifiés dans les écrits scientifiques pour faciliter ce travail d'évaluation. La combinaison de ces deux méthodes est censée apporter certains avantages, mais ces derniers dépendent des caractéristiques de la situation d'évaluation. Cette étude vise, d'une part, à tester les effets de recours à une grille critériée dans un contexte d'évaluation par les pairs sur la compétence des élèves à communiquer oralement, et, d'autre part, à mieux comprendre les processus sous-jacents à son utilisation. Pour ce faire, une méthode mixte en milieu naturel a été suivie. L'échantillon est constitué de 47 élèves âgés de 10 à 12 ans partagés en trois conditions : une condition où les élèves s'appuient sur une grille critériée pour donner des rétroactions avec des consignes strictes et imposées, une condition où les élèves s'appuient sur une grille critériée avec une totale liberté d'action et une condition d'évaluation par les pairs sans grille. Les résultats quantitatifs indiquent une amélioration des composantes orales quelle que soit la condition. Les résultats qualitatifs révèlent des rétroactions plus précises lorsque les élèves utilisent la grille librement. À l'inverse, l'utilisation imposée de la grille implique un plus grand nombre de rétroactions,

une posture plus active des élèves évalués et des interactions de groupe plus nombreuses. Par ailleurs, les résultats mettent en évidence la nécessité de former les élèves à donner des rétroactions de qualité.

Mots-clés

Évaluation de l'oral, Grille critériée, Évaluation par les pairs.

Abstract

Teachers often mention the difficulty to assess oral communication skills as an obstacle to their teachings. Rubrics and peer assessment have been suggested as useful tools to remove this obstacle. The combination of these two methods is also supposed to bring some benefits but these seem to depend strongly on the context. This study aims to test the effects of the use of rubrics in peer assessment on students' oral communication skills, and to better understand how this use shapes peer interactions. To achieve this, a mixed method approach in a natural setting was used. The sample consisted of 47 students aged 10 to 12 years old divided into three conditions: two conditions in which the students used rubrics to give feedback (either with strict and imposed instructions or with total freedom of action) and a last condition without a grid. The quantitative results indicate a significant improvement in oral components whatever the condition. The qualitative results show a better use of the grid when students use it freely. Conversely, the imposed use of the grid implies a greater number of feedbacks given, a more active posture of the evaluated students and more group interactions. Furthermore, the results highlight the need to train students to give effective feedback.

Keywords

Oral evaluation, rubrics assessment, peer assessment.

1. Introduction

Les recherches sur l'oral dans les différents pays francophones mettent en évidence un même constat : les enseignants consacrent peu de temps à son enseignement (Colognesi & Deschepper, 2019 ; Dumais et al., 2017 ; Gagnon et al., 2017 ; Sénéchal, 2017). Ceci malgré l'influence positive de l'enseignement explicite de l'oral sur les compétences orales des élèves (Colognesi & Dolz, 2020 ; Dolz & Gagnon, 2008 ; Dumais et al., 2015 ; Sénéchal, 2016). Un des freins à cet enseignement est la difficulté pour les enseignants de l'évaluer (Nonnon, 2016). En effet, l'oral est indissociable de la personne que représente l'élève (Alrabadi, 2010 ; Berdal-Masuy & Renard, 2015 ; Maurer, 2001), ce qui peut provoquer une certaine subjectivité dans son évaluation (Garcia-Debanco, 1999). De plus, le fait de ne pas maîtriser les composantes de l'oral complique la création de critères d'évaluation à la fois adaptés à l'âge des élèves (Nonnon, 2016) et reprenant toutes les caractéristiques de l'oral (Lafontaine & Préfontaine, 2007). L'oral, pour être conservé, nécessite également la mise en place d'un matériel et d'un dispositif spécifiques (Stordeur & Colognesi, 2020), ce qui est considéré comme énergivore et chronophage par les enseignants (Garcia-Debanco, 1999 ; Wiertz et al., 2020).

Pour répondre à ces difficultés, deux leviers sont mis en évidence dans les écrits scientifiques. Premièrement, l'utilisation de grilles critériées (Dunbar et al., 2006) : les critères aident à identifier ce qui peut être évalué (Berthiaume et al., 2011), ce qui amène une évaluation plus analytique et équitable de l'oral (Dunbar et al., 2006 ; Pomplun et al., 1998). En second lieu, l'évaluation par les pairs, qui est reconnue comme étant une pratique efficace d'enseignement (Ko et al., 2014), décharge l'enseignant en donnant certaines responsabilités aux apprenants (Allal, 2020 ; Huisman et al., 2019). Une meilleure compréhension des objectifs est aussi observée chez les élèves suite à ces pratiques (Allal & Mottier Lopez, 2005 ; Hattie & Timperley, 2007).

Combiner ces deux méthodes offrirait également certains avantages (Panadero et al., 2013) notamment dans le cadre de l'oral (Colognesi et al., 2020 ; Dumais, 2012 ; Vassart et al., 2022). Cependant, certaines études nuancent ces constats (Double et al., 2020 ; Peters et al., 2018 ; Laveault & Miles, 2008), les bénéfices étant fort liés au contexte d'utilisation des grilles en termes d'entraînement (Ashton & Davies, 2015) ou de niveau de scolarité des élèves (Leenknecht & Prins, 2018). Par ailleurs, beaucoup d'études utilisent des données auto-rapportées, ce qui permet difficilement de comprendre les effets de différentes modalités d'évaluation par les pairs (Double et al., 2020).

Par conséquent, nous souhaitons, d'une part, tester l'efficacité d'une grille spécialement conçue pour évaluer les composantes de l'oral dans un contexte d'évaluation par les pairs en primaire, et, d'autre part, mieux comprendre les processus sous-jacents à son utilisation. Pour y arriver, nous avons mis en place trois conditions différentes d'évaluation par les pairs. Dans la condition « *grille imposée* », les élèves utilisent une grille critériée pour donner des rétroactions à leurs pairs, en employant obligatoirement tous les critères de la grille. Dans la condition « *grille libre* », les élèves utilisent la même grille pour réaliser leurs rétroactions mais cette fois-ci ils sont libres de l'exploiter comme ils le souhaitent. Enfin, dans la condition « *sans grille* », les élèves n'ont pas de grille critériée à leur disposition. Au vu des difficultés pour les élèves de savoir ce qu'il est attendu d'eux lors des évaluations de l'oral (Lafontaine &

Messier, 2009), nous pensons que l'augmentation de la visibilité des objectifs imputée par l'utilisation d'une grille critériée (Dumais, 2012 ; Panadero et al., 2013) aura une influence positive sur les performances des élèves.

Sur cette base, notre première question de recherche est la suivante : (1) Le fait de construire une évaluation par les pairs sur la base d'une grille utilisée librement ou de manière imposée impacte-t-il positivement le développement des composantes orales ? Afin de pouvoir finement comprendre l'évolution de ces composantes de l'oral en fonction de la modalité d'utilisation d'une grille critériée, nous souhaitons également analyser comment les élèves s'approprient celle-ci lorsqu'ils sont invités à évaluer leurs pairs, ce qui nous amène à notre deuxième question de recherche : (2) Comment se passe l'évaluation par les pairs lorsque les élèves n'utilisent pas de grille, ou lorsqu'ils l'emploient librement ou de manière imposée ? Nous avons utilisé des analyses quantitatives pour répondre à la première question et une analyse qualitative pour la seconde.

2. Cadre conceptuel

2.1. L'oral et ses composantes

Pour pouvoir analyser finement l'oral, il est primordial de pouvoir identifier ses différentes composantes. Selon Dolz et ses collègues (1993), l'oral est constitué de : (1) la *capacité d'action*, qui correspond à la capacité de l'élève à respecter les codes du genre discursif utilisé et à adapter son discours à la situation de communication ; (2) la *capacité discursive* qui se rapporte à la capacité à produire un contenu et à l'organiser de manière cohérente ; (3) la *capacité linguistico-discursive* qui représente la capacité à utiliser une syntaxe, une grammaire ainsi qu'un vocabulaire corrects et adéquats. Plus récemment, Colognesi et Hanin (2020) ont croisé les travaux de Dolz et al. (1993), Dumais (2016) et les compétences du prescrit belge (2013) pour identifier six composantes, qui seront utilisées dans la suite de ce texte (voir *tableau 1*) : (1) l'intention de communication, (2) l'élaboration du contenu, (3) l'organisation du message, (4) les aspects grammaticaux et lexicaux, (5) les aspects paraverbaux et (6) les aspects non verbaux.

Tableau 1. Parallèle entre les capacités langagières de Dolz et al. (1993) et le socle des compétences belges (2013), repris de Colognesi et Hanin (2020)

Les capacités langagières (Dolz et al., 1993)	Les compétences orales du socle des compétences belge (2013)	Les 6 composantes de l'oral (Colognesi & Hanin, 2020 ; Colognesi et al., 2020)	Un exemple par composante de l'oral (Colognesi & Hanin, 2020) pour le genre « explication orale »
Capacité d'action	C1. Orienter sa parole et son écoute en fonction de la situation de communication	Intention de communication	Ex : Adaptation du discours en fonction du niveau d'expertise du public cible
Capacité discursive	C2. Élaborer des significations C3. Assurer et dégager l'organisation et la cohérence du message...	Élaboration du contenu Organisation du message	Ex : Sélection des informations transmises Ex : Présence de connecteurs
Capacité linguistico-discursive	C3. ... en utilisant le lexique et la morphosyntaxe C4. Utiliser et identifier les moyens verbaux (et non verbaux)	Aspects grammaticaux et lexicaux Aspects paraverbaux Aspects non verbaux	Ex : Syntaxe correcte et vocabulaire précis Ex : Débit approprié : ni trop lent ni trop rapide Ex : Absence de gestes parasites

2.2. Les grilles critériées dans l'évaluation de l'oral

Les grilles critériées ont l'avantage de présenter des critères et des échelles d'appréciation, spécifiques à chaque critère, qui informent les évaluateurs des points d'attention à avoir, et des seuils de réussite qu'ils peuvent attendre (Berthiaume et al., 2011 ; Bernal-Masuy & Renard, 2015). Les critères permettent également d'obtenir une évaluation plus équitable, fiable, valide (Dunbar et al., 2006 ; Lavoie & Bouchard, 2017 ; Pomplun et al., 1998) et moins subjective (Garcia-Debanco, 1999). Grâce à ces critères, les enseignants ont également une meilleure représentation des attendus possibles en termes de performance (Bernal-Masuy & Renard, 2015), ce qui est particulièrement intéressant pour un objet aussi flou et méconnu que l'oral (Colognesi & Deschepper, 2019 ; Colognesi & Dolz, 2017 ; Lafontaine & Préfontaine, 2007). Ainsi, l'enseignant perçoit plus clairement les objectifs finaux, ce qui peut le guider dans la

création de ses séquences d'apprentissage (Berthiaume et al., 2011) et lui permettre d'améliorer l'alignement pédagogique (Biggs, 1996). En plus, il s'agit d'un support non négligeable pour donner aux élèves des rétroactions fines et complètes (Balan & Jönsson, 2018 ; Bernal-Masuy & Renard, 2015). Une meilleure compréhension des attendus est également observée chez les élèves lorsqu'une grille critériée leur est présentée (Bernal-Masuy & Renard, 2015 ; Dumais, 2012). Par ailleurs, en utilisant une grille, les enseignants considèrent moins l'oral comme un tout et adoptent une évaluation plus analytique (Berthiaume et al., 2011 ; Lavoie & Bouchard, 2017).

Bien que les avantages de l'utilisation d'une grille critériée semblent prometteurs (Dunbar et al., 2006), certains obstacles existent également. Notamment le fait que les expériences et le cadre de référence des enseignants influencent leur utilisation de la grille (Coen & Pellegrini, 2011) et que les critères employés ne sont pas toujours compréhensibles par les évaluateurs non formés (Balan & Jönsson, 2018). Par ailleurs, représenter la réalité et ses possibilités uniquement grâce à des critères, n'est pas chose aisée (Bouwer et al., 2016). Effectivement, l'utilisation d'une grille critériée se confronte à l'idée que « le tout fait plus que la somme des parties ». Ainsi, les appréciations données par l'évaluateur pour chaque critère ne permettent pas toujours de faire ressortir l'impression globale d'une performance. Dans le même ordre d'idée, l'échelle de mesure des critères peut difficilement prendre en compte l'ensemble des réponses possibles de la personne évaluée. Il est alors parfois difficile pour l'évaluateur de choisir entre deux échelons de réponses (Bouwer et al., 2016). Pour finir, bien que certains auteurs utilisent déjà des grilles critériées pour évaluer l'oral (Colognesi et al., 2020 ; Berdal-Masuy & Renard, 2015 ; Dunbar et al., 2006 ; Gregg & Mather, 2002), ces grilles n'évaluent pas toujours l'ensemble de ses composantes et elles sont encore peu disponibles et utilisées sur le terrain (Alrabadi, 2011 ; Lafontaine & Messier, 2009).

2.3. L'évaluation par les pairs

Un des principaux avantages de l'évaluation par les pairs est sa potentielle influence positive sur les performances académiques des élèves (Double et al., 2020 ; Huisman et al., 2019). En effet, l'élève va non seulement réguler ses productions sur la base des rétroactions qu'il reçoit, mais également grâce aux commentaires qu'il donne (Colognesi & Deschepper, 2018 ; Dumais, 2012), puisque construire des commentaires permet à l'élève de mieux s'approprier les objectifs fixés par l'enseignant (Allal & Mottier Lopez, 2005 ; Hattie & Timperley, 2007). L'évaluation par les pairs permet également de décharger l'enseignant car les élèves reçoivent des rétroactions sans que l'enseignant doive lui-même évaluer les productions et formuler des commentaires (Double et al., 2020 ; Huisman et al., 2019). De plus, l'évaluation par les pairs étant interpersonnelle (Van Gennip et al., 2009), elle impacte également les compétences de collaboration et de communication des élèves (Gielen et al., 2011).

Néanmoins, l'efficacité de l'évaluation par les pairs dépend beaucoup de la qualité de la rétroaction : au plus celle-ci est élevée, au plus elle est bénéfique pour l'élève (Althausser & Darnall, 2001 ; Van Gennip et al., 2009). Par exemple, Lafontaine (2007) met en évidence trois formes d'évaluation par les pairs: (1) la forme normative où les élèves doivent donner une note à la production évaluée ; (2) la forme qualitative où les élèves donnent une rétroaction commentée sur la performance du pair évalué et (3) la forme descriptive où les élèves identifient les forces et faiblesses de la production au moyen d'une grille critériée, puis

transmettent ces différentes informations à l'élève évalué. Cette dernière forme semble être la plus efficace (Brookhart, 2008). L'utilisation d'une note scorée est particulièrement peu efficace pour les élèves du primaire (Double et al., 2020). Ces rétroactions de formes qualitatives et descriptives peuvent également être de natures différentes (Rodet, 2000). Tout d'abord, les rétroactions informatives sont des rétroactions où l'évaluateur indique ce qui est correct ou non. Ensuite, pour les rétroactions prescriptives, l'évaluateur donne des injonctions en utilisant des expressions du type « *il faut* » ou « *tu devrais* ». Pour finir, les rétroactions suggestives caractérisent des rétroactions où l'évaluateur fournit des conseils et des pistes de solutions pour améliorer la production évaluée. Selon Rodet (2000), les rétroactions suggestives seraient plus efficaces que les informatives qui elles-mêmes fonctionneraient mieux que les prescriptives. Notons que l'utilisation de rétroactions centrées sur la personne et non sur ses performances paraît contre-productive (Brookhart, 2008). Par ailleurs, le style de feedback peut également varier : les rétroactions négatives pointent les éléments à améliorer mais risquent de créer un blocage chez l'élève évalué alors que les rétroactions positives sont plus agréables à écouter mais ne permettent pas forcément la progression (Rodet, 2000). De plus, pour l'évaluation des compétences orales, plus de progrès sont constatés lorsque le feedback est donné oralement (ce qui favorise les rétroactions suggestives) plutôt qu'à l'écrit où l'on retrouve plus de rétroactions de nature prescriptive (Colognesi et al., 2020). Pour finir, les perceptions de la personne évaluée sont également essentielles pour que l'évaluation par les pairs soit efficace (Berndt et al., 2018) d'autant plus que de nombreux élèves sont méfiants par rapport à la pertinence des commentaires reçus (Strijbos et al., 2010), même si leur considération de l'évaluation par les pairs tend à s'améliorer après l'avoir testée (Dumais, 2012).

Pour s'assurer de la qualité de ces rétroactions, il semble essentiel d'enseigner aux élèves comment formuler une rétroaction formative (Van Gennip et al., 2009) et de leur permettre de s'exercer (Double et al., 2020). En effet, dans leur revue systématique de la littérature, Van Gennip et ses collègues (2009) mettent en évidence que des élèves qui reçoivent des informations théoriques sur ce qu'est l'évaluation par les pairs et qui ont l'occasion de s'exercer à en réaliser, modifient plus facilement et en plus grande quantité leur production après les rétroactions reçues que les étudiants qui n'ont reçu aucune information et qui n'ont pas eu l'occasion de s'exercer ou les étudiants qui reçoivent des informations sur le processus d'évaluation par les pairs sans pouvoir s'exercer. Ces résultats s'expliquent notamment par le fait que l'entraînement permet aux élèves d'être plus confortables et plus confiants face au processus d'évaluation par les pairs (Dumais 2012 ; Van Gennip, 2009).

2.4. L'utilisation de grilles critériées dans le cadre de l'évaluation par les pairs

Au vu des avantages de l'utilisation des grilles critériées et de l'évaluation par les pairs, il semble pertinent d'envisager l'utilisation conjointe de ces deux possibilités. Dumais, en 2012, en a réalisé l'expérience. Il ressort que les compétences orales des élèves augmentent suite à la mise en place de l'évaluation par les pairs, en partie parce que les critères de la grille permettent aux élèves de mieux comprendre les objectifs à atteindre et aux évaluateurs d'être plus confiants dans l'évaluation qu'ils donnent (Dumais, 2012 ; Panadero et al., 2013). Par ailleurs, il a été montré que fournir aux élèves un guide à l'évaluation (comme une grille critériée) augmente le nombre et la qualité des commentaires donnés (Peters et al., 2018) et permet une évaluation plus valide, fiable et précise (Panadero et al., 2013).

Cependant, la méta-analyse de Double et ses collègues (2020) indique que l'utilisation d'une grille n'aurait pas d'influence sur l'efficacité de l'évaluation par les pairs. Par ailleurs, les bénéfices de la grille divergent d'un élève à l'autre (Panadero et al., 2013 ; Laveault & Miles, 2008) : les élèves qui s'engagent activement dans leurs apprentissages montrent de meilleurs résultats (Van Gennip et al., 2009) et les élèves plus performants scolairement vont être plus précis et plus sévères dans leur utilisation d'une grille (Laveault & Miles, 2008). Il a aussi été montré que les élèves de l'école primaire ont un style d'évaluation plus sommatif que formatif avec les grilles (Leenknecht & Prins, 2018). Enfin, l'appropriation et la compréhension des critères par les élèves ne sont pas chose aisée. Ainsi, former les élèves à l'utilisation de la grille employée semble important pour que les rétroactions soient efficaces (Laveault & Miles, 2008), plus précises et moins soumises au biais de tendance centrale (Ashton & Davies, 2015). Pour ce faire, il est important que l'enseignant puisse prendre le temps d'expliquer chaque critère et le fonctionnement de chacun de ceux-ci pour que l'élève comprenne comment il doit les interpréter et les utiliser (Laveault & Miles, 2009). Une meilleure compréhension de ces critères par les élèves peut également se faire en engageant activement les élèves dans la création de ceux-ci (Lafontaine, 2007 ; Laveault & Miles, 2009).

3. Méthodologie

Pour approfondir notre compréhension de l'influence, dans un contexte d'enseignement de l'oral, d'une grille critériée dans l'évaluation par les pairs, nous nous sommes saisis de deux questions de recherche : (1) *Le fait de construire une évaluation par les pairs sur la base d'une grille utilisée librement ou de manière imposée impacte-t-il positivement le développement des composantes orales ?* (2) *Comment se passe l'évaluation par les pairs lorsque les élèves n'utilisent pas de grille, ou lorsqu'ils l'emploient librement ou de manière imposée ?* Pour y répondre, nous avons réalisé une étude en milieu naturel (Denzin & Lincoln, 1994) en suivant une méthode mixte de type séquentiel (Creswell, 1999). Plus précisément, nous avons dans un premier temps réalisé des analyses quantitatives pour identifier l'évolution des composantes orales de l'élève à travers le dispositif. Pour mieux comprendre cette évolution et le rôle de l'utilisation de la grille critériée dans celle-ci, nous avons, dans un second temps, effectué une analyse qualitative sur les échanges réalisés entre élèves lors du processus d'évaluation par les pairs. Les participants à l'étude, l'intervention mise en place ainsi que le processus de récolte et d'analyse des données seront présentés dans cette section.

3.1. Participants

Un total de 54 élèves, âgés entre dix et douze ans, provenant de trois classes d'un même établissement de la Fédération Wallonie-Bruxelles de Belgique, a participé à cette étude. Afin d'obtenir les données les plus valides possibles, nous avons décidé de garder uniquement les élèves qui ont suivi l'ensemble du processus : 47 élèves (20 garçons, 27 filles).

La condition *grille imposée* a été mise en place au sein d'une classe de sixième primaire auprès de dix filles et quatre garçons de 11-12 ans. La condition *grille libre* a été attribuée à une classe de cinquième primaire (élèves de 10-11 ans) composée de six filles et neuf garçons. Pour finir, la condition *sans grille* a été remplie par onze filles et sept garçons d'une classe de sixième primaire

(élèves de 11 à 12 ans)¹. Cette dernière condition, nécessitant une intervention un peu plus courte que les autres, a été choisie par une des enseignantes participantes au projet. Les deux autres conditions ont été attribuées aux classes par la chercheuse de référence.

3.2. Déroulement de l'intervention

La *figure 1* présente les étapes suivies dans les trois conditions, qui sont inspirées du dispositif *Itinéraires* (Colognesi & Dolz, 2017 ; Colognesi & Lucchini, 2018a). Le principe est de permettre aux élèves de produire plusieurs versions d'une même production orale pour qu'ils puissent s'améliorer ; c'est la "réoralisation" (Colognesi & Dolz, 2017). Plusieurs interventions sont mises en place entre les versions pour aider les élèves à améliorer leur production : des étayages apportés par l'enseignant (Bucheton & Soulé, 2009 ; Colognesi & Hanin, 2020) sur les caractéristiques et la structure de l'explication, des temps d'évaluation par les pairs (Colognesi & Deschepper, 2018) et des médiations métacognitives (Colognesi et al., 2019, 2020 ; Veenman, 2012). Pour ces dernières, un journal de bord a été distribué à chaque élève au début du dispositif. Au sein de ce journal, les élèves devaient, d'une part, avant chaque étape du projet, y inscrire leurs objectifs attendus et questionnements et, d'autre part, après chacune de ces mêmes étapes, indiquer leurs forces et faiblesses ainsi que ce qu'ils retenaient de l'activité.

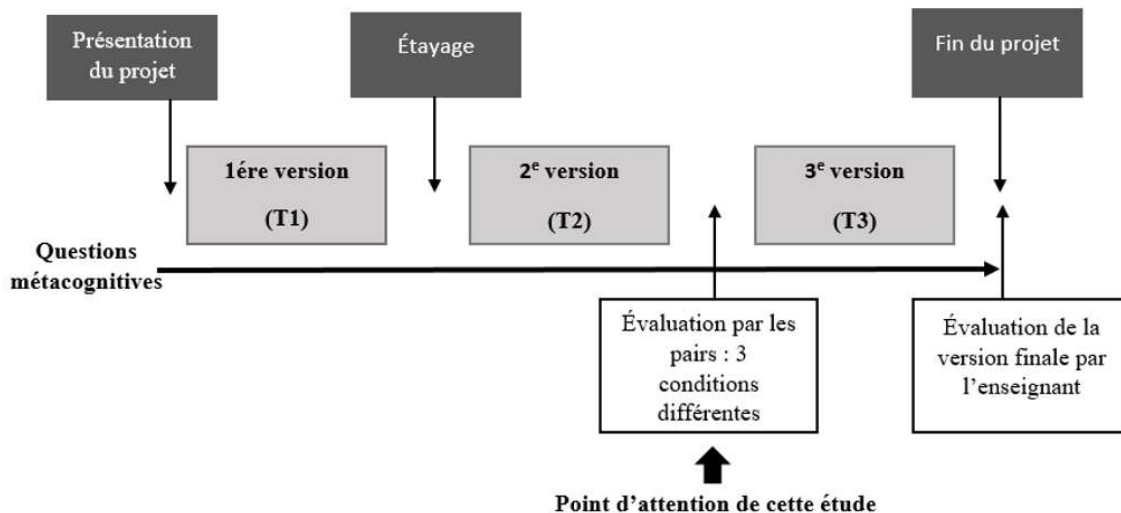


Figure 1. Schématisation du dispositif d'enseignement mis en place pour l'étude

¹ Initialement, nous avons prévu de travailler avec deux classes de cinquième primaire et deux classes de sixième primaire afin d'analyser les différences et ressemblances entre ces deux niveaux. Malheureusement, le contexte de la crise sanitaire, ne nous a pas permis de mettre ce dispositif en place, ce qui explique pourquoi nous avons collaboré avec une seule classe de cinquième primaire.

Plus précisément, l'ensemble des élèves des trois classes a été amené à créer une vidéo qui explique un jeu de société pour un public d'élèves plus jeunes. C'est cette vidéo qui a fait l'objet de l'évaluation des composantes de l'oral. La thématique du jeu de société a été choisie en collaboration avec les enseignants participant au projet. Trois raisons sous-tendent ce choix : (1) le dispositif itinéraire préconise de mettre les élèves dans un projet qui les intéresse (Colognesi & Lucchini, 2018a), le côté ludique du jeu de société permet donc d'atteindre cette recommandation ; (2) les vidéos réalisées sont à destination des élèves plus jeunes de l'école. Les enseignants de ces élèves plus jeunes ont souhaité que cette thématique soit abordée pour qu'ils puissent eux-mêmes intégrer ces vidéos dans leurs séquences d'apprentissage ; (3) cette mise en projet permet de travailler le genre explication avec les élèves. L'explication est un oral hétérogène : il peut être spontané ou préparé et encore monologal (avec un locuteur) ou dialogal (avec plusieurs locuteurs) (Grandaty, 1998). Le point commun de ces différents oraux explicatifs est de répondre à un questionnement de l'interlocuteur et de lui fournir les clés de compréhension pour y répondre (Grandaty, 1998 ; Plane 2001). Dans le cadre de cette étude, nous proposons aux élèves de réaliser un oral préparé monologal permettant de répondre à la question « comment joue-t-on à ce jeu ? ». Par ailleurs, le choix de l'explication d'un jeu permet de travailler plus spécifiquement la compétence des élèves à adapter leur discours pour des élèves plus jeunes qu'eux, à identifier les différentes catégories d'informations pouvant être données (matériel, règles du jeu, but du jeu, ...) et à structurer ces informations de manière intelligible pour l'interlocuteur tout en manipulant un support matériel (pion, plateau...).

Après un temps où les paramètres de la situation de communication sont discutés, chacun a réalisé une première version orale, filmée par un autre élève à l'aide d'une tablette (Temps 1)². À la suite de cette première version, un enseignement des caractéristiques principales d'une explication est réalisé suivant les préconisations de Colognesi et Lucchini (2018b). Cet étayage permet essentiellement de travailler trois composantes de l'oral : *l'intention de communication*, *l'élaboration du contenu* et *l'organisation du message*.

Suite à cela, les élèves ont fait une deuxième version de leur explication (Temps 2). Cette version a été filmée de la même façon et a fait l'objet de l'évaluation par les pairs, dans les trois conditions présentées plus haut.

Ainsi, dans la classe condition grille libre et dans la classe condition grille imposée, une grille d'évaluation est présentée aux élèves. Cette grille (*disponible en annexe 1*) est le fruit d'un travail réalisé lors d'une recherche collaborative (Wiertz et al., 2020) et est structurée selon les composantes de l'oral déjà développées *supra* et adaptées à l'explication orale. La grille est constituée de deux parties. La première est un espace de prise de notes qui permet à l'évaluateur de retranscrire les informations entendues. La seconde partie rassemble une série de critères (repris dans le *tableau 2*) pour lesquels les évaluateurs doivent se positionner. Les différents critères ont été créés sur la base des spécificités de chacune des six composantes de l'oral présentées *supra* (Colognesi & Hanin, 2020). Ainsi, chaque composante est constituée du nombre de critères nécessaires pour l'analyser finement et entièrement. Le groupe collaboratif a également veillé à ce que l'échelle de mesure employée soit appropriée pour chaque critère, ce qui implique une variation en termes d'échelon (Wiertz et al., 2020).

² Pour toutes les versions de l'explication, les élèves ont pu préparer leur explication pendant 10 minutes en s'aidant d'une feuille de brouillon.

Suivant les préconisations de Laveault et Miles (2008), un temps de travail sur la grille a été réalisé avec les élèves, d'une durée de deux périodes de 50 minutes, comprenant la présentation de son fonctionnement et de ses critères ainsi qu'un entraînement à son utilisation. Lors de ces deux périodes, les élèves ont eu l'occasion de poser toutes leurs questions relatives à la compréhension de la grille.

C'est suite à ce travail que les élèves, par sous-groupes de 4 ou 5 (définis par l'enseignant au début du projet), ont évalué la deuxième version de leurs pairs. Pour ce faire, l'ensemble du sous-groupe a regardé la vidéo de chaque membre tout en prenant des notes sur une grille ou une feuille de brouillon. Pour les trois conditions, l'ensemble des membres du sous-groupe a donné, après chaque visionnage et sur la base des notes prises, des rétroactions à l'orateur. Comme évoqué *supra*, dans la condition *grille imposée*, les élèves ont dû utiliser l'ensemble de la grille, pour formuler leur rétroaction ; dans la condition *grille libre*, ils l'utilisaient comme ils le souhaitent pour formuler leurs commentaires. Dans la condition *sans grille*, les élèves n'ont pas reçu de grille et se sont appuyés uniquement sur une feuille de brouillon.

Enfin, à partir de ces retours et d'un rappel de ce qui a été vu, les élèves ont réalisé une troisième et dernière version orale de leur explication (Temps 3). C'est cette version, toujours filmée, qui a été présentée aux destinataires.

L'ensemble de l'intervention s'est étalé sur trois semaines et a duré 10 périodes de 50 minutes dans la condition *sans grille* et 12 périodes pour les deux conditions avec grille (incluant les 2 périodes de travail sur l'instrument). Dans une démarche collaborative (Desgagné & Larouche, 2010 ; Van Nieuwenhoven & Colognesi, 2015), un des enseignants a participé à la réalisation du dispositif pour une meilleure adaptation de celui-ci au terrain ; il a pu former ses collègues aux étapes du dispositif. Par ailleurs, les enseignants des trois classes ont reçu un vade-mecum reprenant en détail le dispositif. Deux rencontres ont également été organisées avec les trois enseignants et la chercheuse de référence afin d'accompagner la mise en place du projet.

3.3. Récolte et analyse des données

3.3.1. Volet quantitatif

Pour répondre à notre première question de recherche, développée *supra*, l'ensemble des vidéos, pour les trois versions, a été visionné et codé. Cela correspond à 141 vidéos allant de 30 secondes à quelques minutes. Pour obtenir une mesure des six composantes de l'oral, la chercheuse de référence a complété la grille d'évaluation utilisée par les élèves, en référence à un vade-mecum reprenant les consignes d'évaluation pour chaque critère. Ensuite, les annotations prises au moyen de la grille ont été encodées et les items demandant des modifications ont été transformés sur la base du *tableau 2*.

Tableau 2. Critères utilisés pour mesurer les composantes de l’oral

Composantes de l’oral	Critères	Type de mesure du critère
Adaptation à la situation de communication (/6)	a. L’orateur fait attention à son ou ses interlocuteurs b. Manière de mener l’interaction	0 (non) ou 1 (oui) est attribué à ce critère Nombre d’occurrences où l’orateur utilise une méthode différente pour interagir avec son interlocuteur. <i>Les résultats ont été transformés en échelle de Likert (voir Annexe 1)</i> Échelle allant de 0 (non) à 2 (oui)
Élaboration d’un contenu (/14)	c. L’explication est adaptée au public d. L’explication est compréhensible e. Les idées sont élaborées (accompagnées d’exemple, ...) f. Nombre de catégories (différentes) abordées g. Nombre d’idées globales (différentes) énoncées h. Nombre d’informations erronées énoncées (<i>item renversé</i>)	Échelle allant de 0 (non) à 2 (oui) Nombre d’occurrences. <i>Les résultats ont été transformés en échelle de Likert (voir Annexe 1)</i>
Organisation du message (/8)	i. Qualité de l’introduction j. Les différents éléments apparaissent les uns après les autres de manière cohérente k. Qualité de la conclusion	Échelle allant de 0 (faible) à 3 (excellent) Échelle allant de 0 (non) à 2 (oui) Échelle allant de 0 (faible) à 3 (excellent)
Aspects grammaticaux et lexicaux (/5)	l. L’orateur utilise un vocabulaire adapté m. L’orateur commence ses phrases sans les terminer n. L’orateur a des tics de langage	Échelle allant de 0 (pas adapté) à 1 (adapté) Échelle allant de 0 (souvent) à 2 (jamais)
Éléments para verbaux (/12)	o. Le volume est adapté p. Le débit de parole est adapté q. L’articulation est distincte r. L’expressivité est adéquate	Échelle allant de 0 (pas adapté) à 3 (excellent)
Éléments non verbaux (/9)	s. Le regard accompagne le discours t. Les gestes sont de qualité u. Utilisation du matériel mis à disposition de l’élève	Échelle allant de 0 à 3 Échelle allant de 0 à 4 Échelle allant de 0 à 2

Tous les critères demandant à l’évaluateur d’inscrire des occurrences ont été transformés en échelle. Pour cela, quatre niveaux ont été créés sur la base du maximum obtenu. Un exemple est présenté dans le *tableau 3* et l’ensemble des transformations est disponible en *annexe 2*. Ensuite, pour chaque temps de mesure, les critères ont été additionnés pour former six scores représentant chacun une des composantes de l’oral. Pour finir, ces scores ont été remis sur 10 puis additionnés pour donner un score total sur 60.

Tableau 3. Exemple de transformation d’échelle

Critère	Mesure	Maximum	Échelle transformée
Nombre d’idées globales énoncées	Occurrences	22	0 : 0-5
			1 : 6-11
			2 : 12-17
			3 : 18 et +

Une analyse ANOVA a permis d’identifier les différences de moyenne pour les six composantes de l’oral entre les conditions pour le temps 1 et le temps 3. Par ailleurs, des analyses ANOVA à mesures répétées ont été réalisées sur l’ensemble des données pour identifier l’effet du dispositif et des conditions sur les composantes de l’oral. Pour les différents tests post hoc réalisés, le test de Bonferroni a été utilisé afin de contrôler le taux d’erreur global. En accord avec Green et Salkind (2017), nous considérons une taille d’effet partielle (η^2) comme faible lorsqu’elle est supérieure à .01, comme modérée lorsqu’elle dépasse .09 et comme grande au-delà de .14.

3.3.2. Volet qualitatif

Pour répondre à notre seconde question de recherche, tous les échanges réalisés lors des évaluations par les pairs ont été enregistrés puis retranscrits. Cela correspond à 18 643 mots pour la condition *grille imposée*, 5018 mots pour la condition *grille libre* et 2339 mots pour la condition *sans grille*. Une analyse de contenu (L'Ecuyer, 1990) en suivant un modèle mixte (utilisation de catégories issues de la littérature ainsi qu'émergence de nouvelles catégories) a été effectuée sur l'ensemble du matériau. Les moments de discussion où les élèves n'étaient pas directement centrés sur la tâche n'ont pas été pris en compte. En tout, six catégories ont été mises en évidence, 1697 unités de sens ont été identifiées dans la condition *grille imposée*, 624 pour la condition *grille libre* et 439 pour la condition *sans grille*.

4. Résultats

Dans cette section, nous allons commencer par identifier les progrès des élèves pour les différentes composantes de l'oral à travers le dispositif en fonction de la condition mobilisée. Une compréhension plus fine de ces résultats est proposée dans un second temps où les différences de fonctionnement entre les conditions seront explicitées.

4.1. Q1 : Le fait de construire une évaluation par les pairs sur la base d'une grille utilisée librement ou de manière imposée impacte-t-il positivement le développement des composantes orales ?

4.1.1. État de la compétence à communiquer oralement des élèves au départ et à l'arrivée

Les résultats de l'ANOVA (*tableau 4*) indiquent que les classes ne sont pas, au T1, équivalentes pour quatre composantes de l'oral ainsi que pour le score total. En effet, les élèves de la condition *grille libre* présentent des scores moyens significativement moins élevés que les élèves des deux autres conditions pour l'*élaboration du contenu* ($F(2,43)=11.83, p<0.001$), les *aspects non verbaux* ($F(2,43)=6.85, p<0.01$) ainsi que pour le *score total* ($F(2,43)=5.26, p<0.01$) et significativement moins élevés que les élèves de la condition *grille imposée* pour les *aspects para verbaux* ($F(2,43)=4.09, p < 0.05$). Ces différences peuvent s'expliquer par le fait que ces élèves sont d'un niveau scolaire inférieur (P5) à celles des deux autres classes (P6). Par ailleurs, les élèves dans la condition *grille imposée* ont des résultats plus faibles que ceux dans la condition *sans grille* ($F(2,43)=3.86, p < 0.05$) pour les *aspects grammaticaux et lexicaux*.

Tableau 4. Analyses descriptives, test ANOVA I et test post-hoc pour les composantes orales au T1 pour chacune des classes

Moyennes	Intention de communication /10	Élaboration du contenu /10	Organisation du message /10	Aspects grammaticaux et lexicaux /10	Aspects para verbaux /10	Aspects non verbaux /10	Total /60
1.Sans grille	6,37 (2,37)	6,65 (1,45)	6,99 (1,79)	8,62 (1,58)	8,48 (1,26)	7,65 (1,66)	44,77 (6,47)
2. Grille libre	5,95 (2,14)	4,29 (1,27)	7,22 (1,36)	9,17 (1,08)	7,32 (1,82)	6,03 (2,45)	39,98 (4,34)
3.Grille imposée	7 (1,57)	6,8 (1,9)	7,33 (1,38)	7,44 (2,26)	8,67 (0,93)	8,44 (1,01)	45,69 (3,74)
ANOVA F(2,43)		11,83 ***		3,86 *	4,09 *	6,85 **	5,26 **
Résultats du test de Bonferonni	N.S.		N.S.				
		2 < 1 et 3		3 < 2	2 < 3	2 < 1 et 3	2 < 1 et 3

Note. * : résultat significatif au seuil $p < .05$; ** : résultat significatif au seuil $p < .01$; *** : résultat significatif au seuil $p < .001$

En revanche, les résultats de l'ANOVA (tableau 5) au temps 3 ne mettent pas en évidence des scores moyens significativement différents entre les trois conditions, sauf pour l'organisation du message où, dans la condition *sans grille*, des résultats inférieurs aux deux autres conditions sont à relever ($F(2,44)=4.21, p < 0.05$). Cela semble montrer que l'écart entre les conditions s'est réduit à la fin du dispositif. Cependant, certains élèves atteignent le maximum des points pour certaines composantes dès la première version. Ces résultats amènent à considérer l'existence d'un plafond dû à la prise de mesure qui ne permet pas d'analyser finement la progression des élèves.

Tableau 5. Analyses descriptives, test ANOVA I et test post-hoc pour les composantes orales au T3 pour chacune des classes

Moyennes	Intention de communication /10	Élaboration du contenu /10	Organisation du message /10	Aspects grammaticaux et lexicaux /10	Aspects para verbaux /10	Aspects non verbaux /10	Total /60
1.Sans grille	7,35 (1,77)	6,29 (1,36)	7,38 (0,87)	8,43 (1,5)	8,92 (1,01)	8,23 (1,47)	46,62 (4,15)
2.Grille libre	6,78 (2,57)	6,77 (0,92)	7,94 (1,05)	8,45 (1,22)	8,33 (1,7)	7,61 (2,42)	46,64 (4,91)
3.Grille imposée	7,6 (1,71)	6,87 (1,36)	8,26 (0,7)	9,17 (1,05)	8,64 (1,05)	8,4 (1,57)	48,96 (3,5)
ANOVA F(2,44)	N.S.	N.S.	4,21 *	N.S.	N.S.	N.S.	N.S.
Résultats du test de Bonferonni			1 < 3				

Note. * : résultat significatif au seuil $p < .05$; ** : résultat significatif au seuil $p < .01$; *** : résultat significatif au seuil $p < .001$

4.1.2. Effets du dispositif et de la condition sur le développement des composantes orales

Les résultats de l'ANOVA à mesures répétées, présentés dans le *tableau 6*, permettent d'observer l'effet du dispositif (effet intra sujet) ainsi que des différentes conditions (effet inter sujets) sur le développement des composantes de l'oral.

Pour l'ensemble des élèves, il ressort une augmentation significative de l'*intention de communication* ($F(2,84)=4.16, p < 0.05, \eta^2=0.09$) et des *aspects non verbaux* ($F(2,84)=6.65, p < 0.01, \eta^2=0.14$). D'après le test de Bonferroni, ces augmentations sont uniquement significatives entre le T1 et le T3 ($MD=0.97, p < 0.05$; $MD=0.89, p < 0.01$). Par ailleurs, aucune différence entre les conditions n'a été mise en évidence. Ainsi, le dispositif a eu un effet, indépendamment de la condition, sur le développement de ces compétences chez les élèves.

Pour les quatre autres compétences, aucune progression significative à travers les trois versions n'est identifiée. Par contre, le niveau moyen des élèves pour l'*élaboration du contenu* est significativement différent entre les conditions ($F(2,42)=14.11, p < 0.001, \eta^2=0.4$). Le test de Bonferroni indique que les scores moyens de la condition *grille libre* sont significativement inférieurs à ceux de la condition *grille imposée* ($MD = -1.79, p < .001$) et de la condition *sans grille* ($MD = -1.38, p < 0.01$). Une différence entre conditions est également observée pour l'*organisation du message* ($F(2,42)=3.86, p < 0.05, \eta^2=0.15$), les résultats moyens de la condition *sans grille* étant significativement inférieurs à la condition *grille imposée* ($MD = -.68 ; p < 0.05$). Ces

résultats confirment, avec une grande taille d'effet, que le niveau des élèves entre les conditions pour ces deux composantes n'est pas équivalent et ce indépendamment du dispositif mis en place.

Tableau 6. Mesure ANOVA à mesures répétées sur l'effet du dispositif, de la modalité et de son interaction pour chacune des composantes de l'oral ainsi que pour le total de celles-ci

		Sommes des carrés	dl	Carré moyen	F	η^2
Intention de communication	Effet intra sujet	21,38	2	10.69	4.16*	.09
	Interaction	3,22	4	.8	.31	.01
	Effet classe	14.04	2	7.02	.88	.04
Elaboration du contenu	Effet intra sujet	13.23	2	6.62	3.08	.07
	Interaction	42.05	4	10.51	4.89	.19
	Effet classe	72.58	2	36.29	14.11***	.40
Organisation du message	Effet intra sujet	9.87	2	4.93	2.8	.06
	Interaction	1.45	4	.36	.21	.06
	Effet classe	11.75	2	5.87	3.86*	.15
Aspects grammaticaux et lexicaux	Effet intra sujet	3.75	2	1.87	1.13	.03
	Interaction	35.39	4	8.85	5.35*	.2
	Effet classe	2.8	2	1.4	.62	.03
Aspects paraverbaux	Effet intra sujet	5.61	1.66	3.38	2.9	.06
	Interaction	7.17	3.32	2.16	1.85	.08
	Effet classe	13.14	2	6.57	2.01	.09
Aspects non verbaux	Effet intra sujet	21.67	2	10.84	6.65**	.14
	Interaction	13.36	4	3.34	2.05	.09
	Effet classe	30.61	2	15.3	2.86	.12
Total	Effet intra sujet	326.118	2	181.06	16.58***	.28
	Interaction	85	4	21.25	1.95	.08
	Effet classe	441.46	2	220.73	4.64*	.18

Note. * : résultat significatif au seuil $p < .05$; ** : résultat significatif au seuil $p < .01$; *** : résultat significatif au seuil $p < .001$

Pour les *aspects grammaticaux et lexicaux*, un effet d'interaction significatif apparaît ($F(4,84)=5.35$ $p>0.01$, $\eta^2=0.2$) avec une grande taille d'effet. De fait, pour les deux conditions avec *grille*, les résultats moyens augmentent après l'évaluation par les pairs, alors qu'ils diminuent pour la condition *sans grille* (voir *figure 2*). Par ailleurs, pour la condition grille libre et sans grille, les résultats finaux sont moins bons que les résultats au T1.

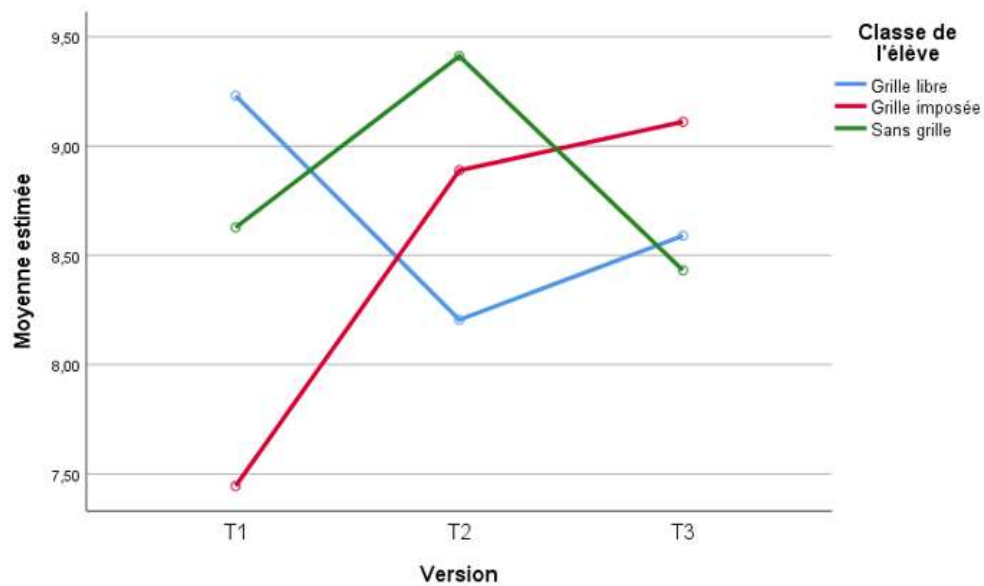


Figure 2. Diagramme à trois lignes brisées présentant l'évolution de la composante « aspect grammaticaux et lexicaux » en fonction de la condition

Pour le score total, les analyses montrent une augmentation moyenne significative des productions des élèves à travers le dispositif ($F(2,84)=16.58, p < 0.001, \eta^2=0.28$). Le test de Bonferroni indique que les résultats au T3 sont significativement plus élevés qu'au T1 ($MD=3.97 ; p < 0.001$) et qu'au T2 ($MD=2.6 ; p < 0.001$). Une différence entre les conditions est également observée ($F(2,42)=4.54, p < 0.05, \eta^2=0.18$) : la condition *grille libre* obtient des scores totaux moyens inférieurs à ceux de la condition *grille imposée* ($MD=-4.59 ; p < 0.05$). Cependant, aucun effet significatif d'interaction n'a été trouvé (voir *figure 3*), ce qui indique que l'augmentation du score total à travers le dispositif ne peut pas être imputable aux conditions.

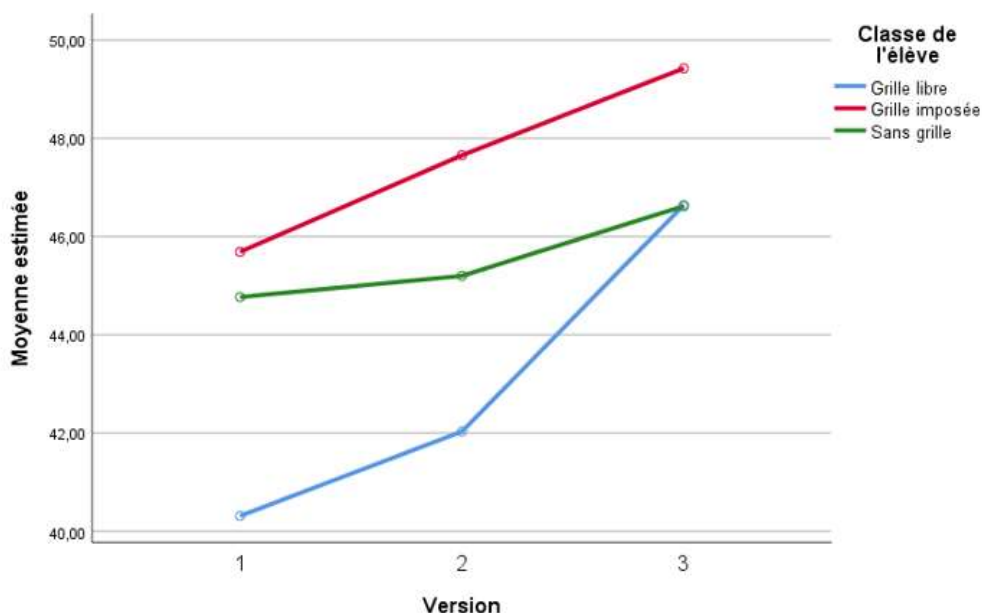


Figure 3. Diagramme à trois lignes brisées présentant l'évolution du total des composantes orales en fonction de la condition

En somme, bien que des différences entre conditions soient observées et que les résultats confirment l'influence du dispositif dans son ensemble sur le développement du total des composantes ainsi que sur deux de celles-ci, il n'y a pas d'impact positif de l'utilisation d'une grille critériée, aussi bien imposée qu'utilisée librement.

Les résultats de l'analyse qualitative, présentés dans la section suivante, amènent à mieux comprendre pourquoi notre hypothèse ne s'est pas vérifiée.

4.2.Q2: Comment se passe l'évaluation par les pairs lorsque les élèves n'utilisent pas de grille, ou lorsqu'ils l'emploient librement ou de manière imposée ?

L'analyse réalisée a permis de faire émerger trois thèmes : les **interactions sociales**, les **commentaires donnés par les évaluateurs** et les **réactions de l'évalué**. Le *tableau 7* offre une vue d'ensemble de ces thèmes, des catégories qui les composent et du nombre d'unités de sens répertorié pour chaque condition. La colonne « pourcentage » de ce tableau permet d'identifier, au sein d'une même condition, la répartition des unités de sens au travers des différentes catégories. Des tableaux similaires sont repris en *annexe 3* pour chacune des sous-catégories identifiées. Dans la section suivante, nous présentons chacun de ces thèmes en détail.

Tableau 7. Synthèse des catégories identifiées dans l'analyse qualitative

Thèmes	Catégories	Grille imposée		Grille libre		Sans grille	
		Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Les interactions sociales	Modalités d'échange au sein du groupe	343	20,21%	79	12,66%	68	15,49%
	Modalités pour faire passer le message	97	5,71%	62	9,93%	84	19,13%
Les commentaires donnés	Nature des commentaires	330	19,45%	279	44,71%	180	41 %
	Explication du commentaire	59	3,48%	38	6,1%	42	9,58%
	Utilisation de notes pour formuler les commentaires	669	39,42%	124	19,87%	13	2,96%
L'orateur	Réaction de l'orateur aux commentaires	199	11,73%	42	6,73%	52	11,84%
Total		1697	100%	624	100%	439	100%

4.2.1. Les interactions sociales

Au vu de la nature interpersonnelle de l'évaluation par les pairs (Van Gennip et al., 2009), il est naturel que les échanges ne se soient pas uniquement centrés sur le contenu des rétroactions et que certaines unités de sens concernent les interactions sociales vécues au sein du groupe. Deux catégories composent ces interactions : les *modalités d'échange au sein du groupe* et les *modalités pour faire passer le message*.

4.2.1.1. Modalités d'échange au sein du groupe

Quand ils se trouvent en sous-groupes, les élèves prennent le temps d'explicitier leur fonctionnement et de réagir aux commentaires donnés par les autres membres. Il est à noter que plus d'unités de sens sont trouvées dans cette catégorie pour la condition *grille imposée*³.

Ainsi, dans les trois conditions, les élèves prévoient l'*organisation de la prise de parole*⁴ en (s') attribuant la parole et en indiquant où en est la discussion (*extrait 1*). Ils gèrent aussi la parole en faisant des *rappels à l'ordre* (*extrait 2*).

³ Les unités de sens pour chaque catégorie ont été transformées en pourcentages. Ainsi, pour l'ensemble de cette partie, lorsque nous comparons le nombre d'unités de sens d'une modalité à l'autre, nous avons vérifié que la différence se marque également au niveau des pourcentages. Plus d'informations en ce sens sont disponibles dans l'*annexe 3*.

Extrait 1 – S3 « Ok, maintenant on va faire du coup les commentaires de X »⁵

Extrait 2 – I2 « Tu feras ça après, c'est pas ton tour »

En plus de la gestion des prises de parole, d'autres aspects ressortent, liés aux interactions entre les élèves. Ainsi, dans les discussions, les élèves reviennent *sur un avis antérieur* pour l'approuver ou le rediscuter (*extrait 3*), ils *reformulent* les commentaires donnés par un autre élève, ils *interrogent* les autres membres du groupe (*extraits 4*) et font preuve d'*humour* entre eux (*extrait 5*).

Extrait 3 – S1 : « Je pense la même chose »

Extrait 4 – I3 : « Tout le monde est d'accord ? »

Extrait 5 – I1 : « X elle dit oui à tout (l'élève fait un clin d'œil à X) »

4.2.1.2. Modalités pour faire passer le message

En plus d'échanger sur leur fonctionnement de groupe, les élèves font également attention à la manière dont ils divulguent leur opinion sur la prestation des autres.

Ainsi, les élèves font précéder leur commentaire par une expression explicitant que ce n'est que leur *opinion personnelle* et non pas une vérité absolue (*extrait 6*). Ils utilisent aussi des termes spécifiques pour nuancer les propos : des *modérateurs de fréquence ou d'intensité*, comme *parfois, un petit peu*, pour atténuer un commentaire donné ; des *amplificateurs* en ponctuant leur phrase par des expressions comme *tout le temps, jamais* ; des *contre balancements* pour apporter un petit élément négatif après un commentaire positif ou inversement (*extrait 7*).

Extrait 6– L1 : « **Moi je trouve** que tu devrais parler un peu plus fort »

Extrait 7 – I2 : « Je trouvais que c'était bien **mais** tu parlais pas assez fort »

Certains élèves utilisent aussi des stratégies de *déresponsabilisation* de l'orateur qui permettent de mettre en exergue que ce n'est pas de la faute de l'orateur si tout n'est pas parfait actuellement (*extrait 8*). D'autres tentent de *dédramatiser* en précisant que ce n'est pas grave de faire des erreurs (*extrait 9*) ou *reconnaissent la difficulté* en expliquant qu'ils sont conscients que ce n'est pas évident de faire mieux (*extrait 10*).

Extrait 8 – S2 : « ça c'est pas sa faute mais la tablette elle est pas terrible »

Extrait 9 – L3 : « Tu t'es un peu trompé dans les gestes et des trucs comme ça, **bon ça c'est pas très grave** »

⁴ Pour l'ensemble des catégories trouvées, les sous-catégories qui ont émergé ont été mises en italique dans le texte.

⁵ Les verbatims seront précédés d'un numéro correspondant au sous-groupe dans lequel ils ont été émis ainsi que de la lettre « I » si l'extrait provient de la condition *grille imposée*, la lettre « L » pour la condition *grille libre* et la lettre « S » pour la condition *sans grille*.

Extrait 10 – S3 : « Même si je sais que c'est un peu stressant »

4.2.2. Les commentaires donnés

Au-delà de ces modalités de fonctionnement et de diffusion du message, il y a le message en lui-même. Ainsi, ce thème reprend les éléments relatifs au contenu des rétroactions : la *nature des commentaires*, l'*explication du commentaire* et l'*utilisation de notes pour donner un commentaire*.

4.2.2.1. Nature des commentaires

La nature des commentaires, en accord avec Rodet (2000), est informative, prescriptive ou suggestive.

Les *commentaires informatifs* sont ceux le plus souvent trouvés dans les trois conditions. Les élèves y énoncent des aspects déjà acquis par l'orateur, ou au contraire qui n'ont pas été bien réalisés à leurs yeux (*extrait 11*). Plus de commentaires positifs sont trouvés dans la condition *grille imposée* (175 unités positives contre 105 négatives) et *grille libre* (185 unités positives contre 54 négatives) et inversement pour la condition *sans grille* (33 unités positives contre 102 négatives). De plus, dans les trois conditions, la majorité des commentaires informatifs peut être reliée à une des six composantes de l'oral (comme l'*extrait 12* qui peut être associé à l'*intention de communication*). À l'inverse, surtout dans la condition *sans grille*, certains commentaires sont soit centrés sur l'orateur lui-même (*extraits 13*) soit correspondent à des aspects non maîtrisables de l'explication (*extrait 14*).

Extrait 11 – I1 : « C'est arrivé genre 2,3 fois que tu commences des phrases sans les terminer »

Extrait 12 – L3 : « Tu expliques bien pour des plus petits que nous »

Extrait 13 – L3 : « T'es super marrante »

Extrait 14 – L2 « J'aime bien tout **sauf les bruits de fond** »

Pour leurs *commentaires prescriptifs*, les moins efficaces selon Rodet (2000), les élèves donnent des injonctions à l'orateur (*extrait 15*). Ils sont un peu plus présents dans la condition *sans grille*.

Extrait 15 – S1 « Tu devrais parler un peu plus fort ».

Quant aux *commentaires suggestifs*, à privilégier selon Rodet (2000), ils correspondent aux commentaires où les élèves proposent des pistes de solution afin d'améliorer l'explication de l'orateur (*extrait 16*). Les commentaires suggestifs sont plus donnés dans la condition *grille imposée* que dans les deux autres conditions.

Extrait 16 : « C'est peut-être mieux de dire d'abord je m'appelle [...] »

4.2.2.2. Explication du commentaire

Dans cette catégorie sont regroupées les unités de sens où les élèves expliquent leur commentaire. Comparativement aux autres catégories, très peu d'unités de sens sont identifiées

ici. Il semblerait que, spontanément, les élèves justifient peu leurs commentaires. Néanmoins, lorsqu'ils le font, sept façons de faire ressortent de l'analyse.

Ainsi, pour justifier leur jugement, les élèves évoquent un *extrait de la présentation* (extrait 17) ; font part d'une *expérience subjective* en exprimant leur ressenti (extrait 18) ; *comparent* les possibilités de l'orateur à celles d'autres élèves (extrait 19) ; font référence *aux règles du jeu* dont il est question pour mettre en évidence les erreurs de l'orateur ; interprètent la *réaction possible des interlocuteurs* (extrait 20) ou encore s'appuient sur *l'environnement de l'orateur* (extrait 21). En plus, les élèves des conditions *avec grille* donnent des *exemples fictifs*, en mettant en évidence des éléments concrets imaginés qui n'ont pas été donnés par l'orateur lors de l'explication (Extrait 22).

Extrait 17 – L2 : « Quand c'est le roi tu montres bien le roi et tout »

Extrait 18 - I1 : « Parce que là même moi j'ai pas tout compris »

Extrait 19 – S1 : « Moi par exemple je fais très beaucoup de tics de langage mais c'est parce que je fais de la logopédie, alors que toi je sais très bien que tu pourrais ne pas faire de tics de langage »

Extrait 20 – S3 « Parce que les petits ils ont vraiment envie d'avoir des grands très gentils »

Extrait 21 – L4 « Du coup fais bien attention de bien parler fort surtout quand on est dehors **parce qu'il y a plein de gens qui passent et tout** »

Extrait 22 – I3 - « Ben oui tu as un vocabulaire adapté, **tu n'utilises pas beaucoup machin truc chose bidule** »

4.2.2.3. Utilisation de notes pour formuler les commentaires

Dans cette catégorie sont regroupées toutes les unités de sens où les élèves, pour donner leurs commentaires, s'appuient sur des notes. Ainsi, soit ils utilisent explicitement la grille d'évaluation de l'oral (conditions *grille imposée* et *grille libre*) soit leur feuille de brouillon (condition *sans grille*). Pour la condition sans grille, treize unités de sens font *référence à la feuille de brouillon* (extrait 23). Ainsi, très peu d'élèves explicitent cette utilisation.

Extrait 23 – S1 : « **Ben moi j'ai mis**, il n'y a pas de démonstration du jeu »

Par contre, les élèves des conditions *avec grille*, surtout dans la condition *imposée*, s'appuient beaucoup sur la grille pour formuler leur jugement. Dans la condition *grille imposée*, les élèves vont majoritairement *indiquer la réponse* qu'ils y ont cochée sans pour autant traduire ce que cela signifie pour l'orateur (extrait 24). À l'inverse, les élèves de la condition *grille libre* utilisent principalement *la grille comme checklist* pour guider la conversation. Ici, ils lisent un critère de la grille pour engager la discussion sur ce sujet, ce qui leur permet de discuter des différents critères proposés, sans se restreindre à uniquement énoncer les réponses qu'ils ont annotées sur la grille (extrait 25). Les élèves de la condition *grille libre* utilisent donc la grille de manière plus pertinente pour donner des commentaires plus complets.

Extrait 24 - I2 : « Pour l'explication est compréhensible, j'ai mis un peu ».

Extrait 25 – S1 : « Du coup pour **l'explication est adaptée au public ben oui parce que** tu regardes la caméra et t'expliques ça bien pour des plus petits que nous »

De plus, au vu de la complexité de la grille, des discussions sur l'endroit où se situe le groupe sur la grille (*extrait 26*) ainsi que sur la *signification d'un item* (*extrait 27*) sont identifiées, surtout dans la condition *grille imposée*.

Extrait 26 - I1 : « Euh X on est pas là, on est ici »

Extrait 27 – I1 : « Parce que c'est pas ça l'intonation, c'est le point d'exclamation, point d'interrogation »

4.2.3. Les réactions de l'évalué

La réaction de l'orateur aux commentaires reçus est également essentielle pour comprendre les effets de l'utilisation d'une grille critériée. Ainsi, cette catégorie reprend 8 types de réactions possibles chez l'orateur.

Dans la condition *grille imposée*, les élèves ont tendance à être d'*accord* avec le commentaire reçu (*extrait 28*) alors que, dans les deux autres conditions, les élèves vont plus souvent mettre en évidence qu'ils sont en *désaccord* avec le commentaire reçu et/ou qu'ils n'auraient pas pu faire autrement (*extrait 30*).

Extrait 28 – I2 : « *Ouais je sais* »

Extrait 29 – I4 : « Ouais ben en même temps il y avait énormément de bruit »

Les orateurs ont des réactions plus proactives dans la condition *grille imposée*. Effectivement, ils vont réaliser des *évaluations formatrices* en mettant en évidence des éléments qu'ils devraient *améliorer* (*extrait 30*), *poser des questions* (*extrait 31*) et proposer des *pistes d'amélioration* pour répondre aux commentaires reçus (*extrait 32*).

Extrait 30 – I1 « Le regard alors moi je dis c'est bof parce que de temps en temps je regarde ma feuille je suis tout le temps gnuuu (en regardant vers le bas) »

Extrait 31 – S1 : « Tu trouvais que j'articulais bien ? »

Extrait 32 – I2 « Je devrais moins l'utiliser je devrais plus parler »

Dans les trois conditions, les orateurs vont également *remercier* les autres élèves pour les commentaires ou partager un *ressenti émotionnel* (*extrait 33*). Dans les conditions avec les grilles, les élèves vont également réagir avec *humour* aux commentaires donnés (*extrait 34*).

Extrait 33 – S3 « Oh non je n'aime pas les commentaires »

Extrait 34 – I3 « Tu ne m'entendais pas ? ALORS BONJOUR (l'orateur augmente exagérément le son) »

En définitive, les résultats des analyses qualitatives nous permettent de mieux comprendre les résultats de notre première question de recherche. Effectivement, dans les conditions avec

grille, et surtout dans la condition *grille imposée*, les élèves lisent certains items sans véritablement fournir de conseils à l'orateur. Dans le même ordre d'idée, les élèves n'expliquent pas plus leurs commentaires dans les deux conditions avec grilles que dans celle sans grille. Malgré ces difficultés, les résultats mettent en évidence d'autres apports des grilles. Par exemple, plus de rétroactions ainsi qu'une posture plus proactive chez l'orateur sont présentes dans la condition *grille imposée*.

5. Discussion et conclusion

Dans cette contribution, nous avons mis en place un dispositif de réoralisation (Colognesi & Dolz, 2017), afin d'observer si l'utilisation d'une grille critériée impacte positivement l'évaluation par les pairs ainsi que la compétence à communiquer oralement des élèves de fin du primaire. Plus précisément, deux questions ont guidé nos investigations : (1) *Le fait de construire une évaluation par les pairs sur la base d'une grille utilisée librement ou de manière imposée impacte-t-il positivement le développement des composantes orales ?* (2) *Comment se passe l'évaluation par les pairs lorsque les élèves n'utilisent pas de grille, ou lorsqu'ils l'emploient librement ou de manière imposée ?*

Pour notre première question, les résultats des analyses quantitatives mettent en évidence un effet positif du dispositif pour l'*intention de communication*, les *aspects non verbaux* ainsi que pour le *score total* des six composantes de l'oral. Cependant, l'utilisation d'une grille, qu'elle soit imposée ou employée librement, n'a pas influencé les performances orales des élèves.

Ces résultats semblent contradictoires avec les études menées par Dumais (2012) et Colognesi et al. (2020) qui présentent une amélioration de la compétence à communiquer oralement suite à l'utilisation d'une grille critériée lors d'évaluations par les pairs. Mais dans ces études, l'ensemble des élèves a utilisé ce support, il était donc difficile de différencier les effets de l'évaluation par les pairs de ceux de l'utilisation d'une telle grille. C'est ce que nous avons pu faire dans notre travail, dont les résultats rejoignent ceux de la méta-analyse de Double et ses collègues (2020) : l'utilisation d'une grille ne permet pas d'augmenter les effets significatifs apportés par l'évaluation par les pairs.

La littérature précise, d'une part, que le contexte d'utilisation d'une grille critériée nuance les bénéfices sur les performances des élèves (Ashton & Davies, 2015) et, d'autre part, que son utilisation apporte d'autres avantages (Panadero et al., 2013 ; Peters et al., 2018). Ainsi, notre seconde question de recherche a permis d'investiguer comment les élèves utilisent une grille critériée pour faire de l'évaluation par les pairs, et ce que cette utilisation apporte à ce processus. Cinq aspects ressortent de nos analyses qualitatives.

Premièrement, en adéquation avec les résultats de Peters et ses collègues (2018), un plus grand nombre de rétroactions est observé lorsque les élèves utilisent une grille critériée, surtout dans la condition *grille imposée*. Ce constat peut s'expliquer par le fait que, contrairement à la condition *grille libre*, les élèves de la condition *grille imposée* s'appuient à de nombreuses reprises sur la partie prise de notes de la grille, ce qui ajoute du contenu aux échanges. Ainsi, imposer l'utilisation complète d'une grille renforce l'adoption d'une évaluation plus complète et analytique (Berthiaume et al., 2011 ; Lavoie & Bouchard, 2017), ce qui est un avantage pour un objet abstrait et subjectif comme l'oral (Dunbar et al., 2006).

Deuxièmement, pour les trois conditions testées, seule une minorité des rétroactions sont accompagnées d'explications. Or, en utiliser est une des caractéristiques à remplir pour fournir une rétroaction de qualité (Ashton & Davies, 2015). Dans le même ordre d'idée, la grande majorité des rétroactions était de nature informative (Rodet, 2000), ce qui est cohérent avec les résultats d'une étude antérieure (Colognesi et al., 2020). Il semblerait donc que, spontanément, les élèves donnent des commentaires qui s'inscrivent plus dans une dynamique d'énonciation d'observations plutôt que dans celle de proposer des suggestions, et ce indépendamment de l'utilisation d'une grille. Par contre, c'est dans la condition *sans grille* que se trouvent le plus de commentaires prescriptifs, à éviter selon Rodet (2020). Les commentaires suggestifs (à privilégier selon cet auteur) sont, quant à eux, plus présents dans la condition *grille imposée*, ce qui est un avantage en faveur de cette modalité de travail. Par ailleurs, dans les deux conditions *avec grille*, les commentaires donnés mettent plus en évidence les aspects positifs de la production. Peut-être que l'utilisation d'une grille fournit aux élèves une meilleure vue d'ensemble des composantes de l'oral (Berthiaume et al., 2011 ; Lavoie & Bouchard, 2017) et de ce qui peut être évalué (Bernal-Masuy & Renard, 2015 ; Dumais, 2012), ce qui leur permettrait ne pas se centrer uniquement sur les éléments qui posent problème. Ces résultats semblent donc en faveur de l'utilisation d'une grille critériée.

Pourtant, aucun apport significatif de celle-ci sur les composantes de l'oral n'a été identifié. Une première piste d'explication est que se centrer sur le positif protège les ressentis de la personne évaluée mais n'invite pas forcément aux progrès (Rodet, 2000). Une autre explication possible est que, dans les conditions *avec grille*, les commentaires des élèves sont cadencés aux critères apportés par la grille alors que, dans la condition *sans grille*, certaines rétroactions se centrent sur des éléments qui n'ont pas été repris par les critères. Une certaine liberté des propos est donc rendue possible par cette condition, ce qui pourrait avoir un avantage sur les performances des élèves. Cependant, c'est dans cette condition *sans grille* qu'il y a le plus de commentaires centrés sur l'orateur et sur des éléments non maîtrisables par celui-ci, ce qui est à éviter pour fournir des rétroactions efficaces (Brookhart, 2008) ; cette liberté est donc à double tranchant.

Troisièmement, il apparaît que, dans la condition *grille libre*, les élèves ont plutôt utilisé la grille comme une checklist pour guider leurs interactions et identifier les aspects à discuter. Cette stratégie permet aux élèves de ne pas oublier d'éléments importants et de discuter en sous-groupe chacun de ceux-ci alors que, dans la condition *grille imposée*, les élèves utilisent majoritairement la grille pour donner les réponses qu'ils y ont inscrites, sans pour autant faire le lien avec la production évaluée. Une dynamique plus sommative est ainsi suivie par ces élèves, comme souvent pour des élèves du primaire (Leenknecht & Prins, 2018). De ce fait, les élèves de la condition *grille libre* semblent mieux optimiser leur utilisation de la grille. Il est possible qu'au vu de la complexité de ce support, son utilisation complète dans la condition *imposée* provoque une surcharge cognitive chez les évaluateurs (Sweller et al., 2011) qui cherchent alors à employer la grille le plus simplement possible pour diminuer cette charge. À l'inverse, les élèves, dans la condition *grille libre*, peuvent contourner cette surcharge cognitive en se concentrant sur moins d'éléments, mais en augmentant la qualité de l'utilisation de ceux-ci. Il est possible qu'une augmentation du temps de formation et d'entraînement avec la grille réduise cette surcharge cognitive (Sweller et al., 2011) et permette aux élèves de l'utiliser de manière plus qualitative et complète (Ashton & Davies, 2015 ; Laveault & Miles, 2008). Impliquer les élèves dans la construction de la grille pourrait également permettre une meilleure compréhension et donc une meilleure maîtrise des critères par les élèves (Lafontaine, 2007).

Quatrièmement, dans la condition *grille imposée*, les orateurs ont montré une posture plus proactive : ils acceptent et identifient plus facilement les éléments à améliorer, posent plus de questions et proposent plus de pistes de solution. Ces résultats pourraient être expliqués par une meilleure compréhension chez les élèves des objectifs attendus et une plus grande confiance aux commentaires donnés par autrui (Allal & Mottier Lopez, 2005 ; Brendt et al., 2018 ; Hattie & Timperley, 2007), induite par l'obligation d'utiliser l'ensemble de la grille critériée. Ces résultats plaident en faveur d'une utilisation *imposée* du support. Cependant, certains élèves tendent à poser plus de questions, à être plus souvent d'accord ou justement en désaccord avec les commentaires donnés. Cette diversité de réactions rejoint l'idée que l'efficacité d'une grille est très variable d'un individu à l'autre (Panadero et al., 2013) et qu'elle dépend, entre autres, de la perception de la personne évaluée (Berndt et al., 2018).

Pour finir, un temps très conséquent a été utilisé dans la condition *grille imposée* pour gérer les interactions sociales au sein du sous-groupe et organiser son fonctionnement, ce qui pourrait avoir un effet positif sur les compétences communicationnelles des élèves (Gielen et al., 2011). Ces constatations appuient l'importance mise en évidence par Van Gennip et ses collègues (2015) de prendre en compte les variables interpersonnelles dans l'étude de l'évaluation par les pairs.

6. Limites et perspectives

Plusieurs limites de cette étude sont à mettre en évidence. Premièrement, le peu de participants complique l'observation d'effets significatifs et la généralisabilité des résultats ; et le fait que les trois groupes n'étaient pas du même niveau scolaire au départ ne permet pas une comparaison idéale, d'autant que la condition expérimentale est confondue avec l'effet enseignant. Répliquer cette étude avec plus d'élèves de compétences équivalentes au départ semble ainsi nécessaire. Deuxièmement, relativement aux résultats quantitatifs, certains élèves atteignent pour certaines composantes le maximum des points dès le premier temps de mesure. Ainsi, il est possible qu'un effet plafond n'ait pas permis de constater certains progrès chez les élèves. Affiner l'échelle d'appréciation pourrait être pertinent pour obtenir des données plus fines. Troisièmement, toutes les composantes de l'oral n'ont pas été travaillées en classe alors qu'elles ont été évaluées et utilisées lors de l'évaluation par les pairs, ce qui ne permet pas un bon alignement pédagogique (Biggs, 1996). Se centrer sur certaines composantes permettrait de diminuer, d'une part, le temps consacré à l'évaluation par les pairs et, d'autre part, la surcharge cognitive des élèves. Quatrièmement, les élèves ont eu peu de temps pour s'approprier et s'entraîner à l'utilisation de la grille. Refaire une étude dans laquelle les élèves ont déjà utilisé cette grille à plusieurs reprises permettrait de mieux saisir l'impact qu'elle peut avoir sur les performances orales des élèves lorsqu'ils la maîtrisent.

7. Implications pratiques

Notre étude soutient l'intérêt potentiel de l'évaluation par les pairs pour le développement des compétences orales des élèves de fin de primaire. Cependant, il n'est pas certain que l'utilisation d'une grille critériée ait une valeur ajoutée. Par ailleurs, lorsque les élèves sont obligés d'utiliser une grille, ils risquent, si elle est trop complexe et qu'ils ne la maîtrisent pas suffisamment, d'énoncer leur réponse sans transformer leurs annotations en réels commentaires utilisables par l'élève évalué. L'augmentation de l'autonomie donnée aux élèves implique une diminution du nombre de commentaires donnés mais permet une utilisation plus qualitative de la grille. De plus, que ce soit avec ou sans grille, il semble utile de former les élèves à expliquer leurs commentaires et à donner des rétroactions de nature suggestive.

L'utilisation imposée d'une grille dans son entièreté apporte néanmoins certains avantages non négligeables comme un plus grand nombre de rétroactions, notamment de nature suggestive, une posture plus active de l'orateur lors de la réception des commentaires et des interactions de groupe plus nombreuses.

8. Références

- Allal, L., & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage: revue de publications en langue française. *OCDE, L'évaluation formative: pour un meilleur apprentissage dans les classes secondaires*, 265-290.
- Allal, L. (2020). Assessment and the co-regulation of learning in the classroom, *Assessment in Education: Principles, Policy & Practice*, (27)4, 332-349. <https://doi.org/10.1080/0969594X.2019.1609411>
- Alrabadi, E. (2010). Quelle méthodologie faut-il adopter pour l'enseignement/apprentissage de l'oral ? *Didáctica. Lengua y Literatura*, 23, 15-34. https://doi.org/10.5209/rev_DIDA.2011.v23.36308
- Althaus, R., & Darnall, K. (2001). Enhancing critical reading and writing through peer reviews: An exploration of assisted performance. *Teaching Sociology*, 23-35.
- Ashton, S., & Davies, R. S. (2015). Using scaffolded rubrics to improve peer assessment in a MOOC writing course. *Distance education*, 36(3), 312-334. <https://doi.org/10.1080/01587919.2015.1081733>
- Balan, A., & Jönsson, A. (2018). Increased explicitness of assessment criteria: Effects on student motivation and performance. *Frontiers in education*, 81 (3). <https://doi.org/10.3389/feduc.2018.00081>
- Berdal-Masuy, F., & Renard C. (2015) Comment évaluer l'impact des pratiques théâtrales sur les progrès en langue cible ? Vers un nouveau dispositif d'évaluation de l'oral en FLE. *Lidil*, 52, 153-174. <https://doi.org/10.4000/lidil.3872>
- Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, 33(1), 31-49. <https://doi.org/10.1007/s10212-017-0343-z>
- Berthiaume, D., David, J., & David, T. (2011). Réduire la subjectivité lors de l'évaluation des apprentissages à l'aide d'une grille critériée : repères théoriques et applications à un enseignement interdisciplinaire. *Revue internationale de pédagogie de l'enseignement supérieur*, 27(2). Repéré à <http://ripes.revues.org/524>
- Biggs, J.B. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- Bouwer, R., Koster, M., & Van den Bergh, H. (2016). Benchmark rating procedure, best of both worlds? *Comparing procedures to rate text quality in a reliable and valid manner*. Manuscript submitted for publication.
- Brookhart, S. M. (2008). Feedback that fits. Engaging the whole child: Reflections on best practices in *learning, teaching, and leadership*, 65(4), 54-59.
- Bucheton, D., & Soulé, Y. (2009). Les gestes professionnels et le jeu des postures de l'enseignant dans la classe : un multi-agenda de préoccupations enchâssées. *Éducation et Didactique*, vol. 3, 3, 29-48. <https://doi.org/10.4000/educationdidactique.543>
- Coen, P-F., & Pellegrini, S. (2011). *Quand plusieurs correcteurs évaluent les mêmes tâches complexes : regards sur la concordance inter-juges et point de vue des évaluateurs* [communication orale]. 23^e colloque de l'ADMEE-Europe. Paris, France.
- Colognesi, S. & Dolz-Mestre, J. (2017). Faire construire des scénarios pour développer les capacités orales des élèves du primaire. Dans J.-F. de Pietro, Carole Fisher et R. Gagnon (dir.), *L'oral aujourd'hui : perspectives didactiques* (p. 177-196). Presses universitaires de Namur. <https://archive-ouverte.unige.ch/unige:97894>
- Colognesi, S., & Deschepper, C. (2018). La relecture collaborative comme levier de réécriture et de soutien aux corrections des textes. *Le français aujourd'hui*, 203(4), 63-72. <https://doi.org/10.3917/lfa.203.0063>.
- Colognesi, S., & Lucchini, S. (2018a). Enseigner l'écriture: l'impact des étayages et des interactions entre pairs sur le développement de la compétence scripturale. *Canadian Journal of Education*, 41(1), 514-540.

- Colognesi, S., & Lucchini, S. (2018b). Aider les élèves à organiser leurs écrits: les effets d'un étayage destiné à travailler la superstructure textuelle. *Repères. Recherches en didactique du français langue maternelle*, 57, 143-162.
- Colognesi, S., & Deschepper, C. (2019). Les pratiques déclarées de l'enseignement et de l'évaluation de l'oral au primaire : qu'en est-il en Belgique francophone ? *Language & Literacy*, 21(1), 1-18. <https://doi.org/10.20360/langandlit29365>
- Colognesi, S., & Hanin, V. (2020). Quelles pratiques efficaces pour enseigner l'oral? Expérimentations dans 8 classes du primaire et suivi de 16 futurs enseignants. *Recherches*, 73, 35-54. <https://doi.org/10.26822/iejee.2019257657>
- Colognesi, S., Hanin, V., Still, A., & Van Nieuwenhoven, C. (2019). The Impact of Metacognitive Mediation on 12-Year-Old Students' Self-Efficacy Beliefs for Performing Complex Tasks. *International Electronic Journal Of Elementary Education*, 12(2), 127-136. <https://doi.org/10.26822/iejee.2019257657>.
- Colognesi, S., Piret, C., Demorsy, S., & Barbier, E. (2020). Teaching Writing—With or Without Metacognition?: An Exploratory Study of 11- to 12-Year-Old Students Writing A Book Review. *International Electronic Journal Of Elementary Education*, 12(5), 459-470. <https://doi.org/10.26822/iejee.2020562136>
- Colognesi, S., Vassart, C., Blondeau, B., & Coertjens, L. (2020). Formative peer assessment to enhance primary school pupils' oral skills: Comparison of written feedback without discussion or oral feedback during a discussion. *Studies in educational evaluation*, 67, 100917. <https://doi.org/10.1016/j.stueduc.2020.100917>
- Creswell, J. W. (1999). Mixed-method research: Introduction and application. In *Handbook of educational policy*, 455-472.
- Denzin, N. K., & Y. S. Lincoln. (1994). Introduction: Entering the Field of Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 1–17). Sage.
- Desgagné, S., & Larouche, H. (2010). Quand la collaboration de recherche sert la légitimation d'un savoir d'expérience. *Recherches en éducation*, 1, 7-18.
- Dolz, J., Pasquier, A., & Bronckart, J.P. (1993). L'acquisition des discours : Émergence d'une compétence ou apprentissage de capacités langagières diverses ? *Études de linguistique appliquée*, 92, 23-37.
- Dolz, J., & Gagnon, R. (2008). Le genre du texte, un outil didactique pour développer le langage oral et écrit. *Pratiques*, 137/138, 179-198.
- Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, 32, 481-509.
- Dumais, C. (2012). L'évaluation de l'oral et la grille d'observation progressive. In *Curriculum et progression en français : Actes du 11e colloque de l'AIRDF* (pp. 363-381).
- Dumais, C., Lafontaine, L., & Pharand, J. (2015). Enseigner et évaluer l'oral en milieu défavorisé au Québec : premiers résultats d'une recherche-action-formation au 3e cycle du primaire. *Langage et littérature*, 17(3), 5-27.
- Dumais, C., Lafontaine, L., & Pharand, J. (2017). « J'enseigne et j'évalue l'oral » : pratiques effectives au 3e cycle du primaire. In J.-F. de Pietro, C. Fisher & R. Gagnon (dir.). *L'oral aujourd'hui : Perspectives didactiques* (151-174). Presses Universitaires de Namur.
- Dumais, C. (2016). Proposition d'une typologie des objets d'enseignement/apprentissage de l'oral. *Les dossiers des sciences de l'éducation*, 36, 37-56.
- Dunbar, N. E., Brooks, C. F., & Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31(2), 115-128. <https://doi.org/10.1007/s10755-006-9012-x>

- Gagnon, R., de Pietro, F., & Ficher, C. (2017). Introduction. In J-F. de Pietro, C. Fisher & R. Gagnon (dir.), *L'oral aujourd'hui : perspectives didactiques* (pp. 11-42). Presses universitaires de Namur.
- Garcia-Debanc, C. (1999). Evaluer l'oral. *Pratiques*, 103-104, 193-212.
- Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education*, 36(2), 137-155. <https://doi.org/10.1080/02602930903221444>
- Grandaty, M. (1998). Elaboration à plusieurs d'une conduite d'explication en sciences, au cycle 2. *Repères. Recherches en didactique du français langue maternelle*, 17(1), 109-125. <https://doi.org/10.3406/reper.1998.2251>
- Green, S., & Salkind, N. (2017). *Using SPSS for windows and macintosh* (8th ed.). Pearson.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Huisman, B., Saab, N., van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: a Meta-Analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863-880. <https://doi.org/10.1080/02602938.2018.1545896>
- Ko, J., Sammons, P., & Bakkum, L. (2014). Effective teaching, education development trust. *Reading, United Kingdom*.
- Lafontaine, L. (2007). *Enseigner l'oral au secondaire. Séquences didactiques intégrées et outils d'évaluation*. Chenelière Education.
- Lafontaine, L., & Préfontaine, C. (2007). Modèle didactique descriptif de la production orale en classe de français langue première au secondaire. *Revue des sciences de l'éducation*, 33(1), 47-66.
- Lafontaine, L., & Messier, G. (2009). Les représentations de l'enseignement et de l'évaluation de l'oral chez des enseignants et des élèves du secondaire en français langue d'enseignement. *Revue du Nouvel-Ontario*, 34, 119-144.
- Laveault, D., & Miles, C. (2008). Utilité des échelles descriptives et différences individuelles dans l'autoévaluation de l'écrit. *Mesure et évaluation en éducation*, 31(1), 1-29. <https://doi.org/10.7202/1025011ar>
- Lavoie, C., & Bouchard, É. (2017). Formation universitaire à l'évaluation de l'oral : regard sur la capacité d'autoévaluation de futurs enseignants. In J.-F., De Pietro, C. Fisher & R. Gagnon (Dir.), *L'oral aujourd'hui : perspectives didactiques* (pp. 259-274.). Presses universitaires de Namur.
- L'Écuyer, R. (1990). *Méthodologie de l'analyse développementale de contenu. Méthode GPS et concept de soi*. Presses de l'université du Québec.
- Leenknecht, M. J., & Prins, F. J. (2018). Formative peer assessment in primary school: the effects of involving pupils in setting assessment criteria on their appraisal and feedback style. *European Journal of Psychology of Education*, 33(1), 101-116.
- Maurer, B. (2001). *Une didactique de l'oral. Du primaire au lycée*. Éditions Bertrand-Lacoste.
- Nonnon, E. (2016). 40 ans de discours sur l'enseignement de l'oral. Enseignement/apprentissage de la langue, des textes et des discours. *Pratiques [En ligne]*, 169-170. <https://doi.org/10.4000/pratiques.3115>.
- Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, 39(4), 195-203. <https://doi.org/10.1016/j.stueduc.2013.10.005>
- Peters, O., Körndle, H., & Narciss, S. (2018). Effects of a formative assessment script on how vocational students generate formative feedback to a peer's or their own performance. *European Journal of Psychology of Education*, 33(1), 117-143. <https://doi.org/10.1007/s10212-017-0344-y>
- Plane, S. (2001). Quelques repères pour la description d'explications. Fonctionnements et dysfonctionnements de l'explication orale. *Repères. Recherches en didactique du français langue maternelle*, 24(1), 113-136. <https://doi.org/10.3406/reper.2001.2373>

- Pomplun, M., Capps, L., & Sundbye, N. (1998). Criteria teachers use to score performance items. *Educational Assessment*, 5(2), 95-110. https://doi.org/10.1207/s15326977ea0502_2
- Rodet, J. (2000). *La rétroaction, support d'apprentissage ? DistanceS*. Accessed Montreal : TELUQ/UQAM. <https://edutice.archives-ouvertes.fr/>
- Sénéchal, K. (2016). Expérimentation et validation de séquences didactiques produites selon une ingénierie didactique collaborative: L'enseignement de la discussion et de l'exposé critique au secondaire [Thèse de doctorat non publiée]. Ulaval, Québec.
- Sénéchal, K. (2017). Quels obstacles à la mise en œuvre de séquences didactiques destinées à l'enseignement de l'oral ? Dans J.-F. De Pietro, C. Fisher, & R. Gagnon (Dir.). *L'oral aujourd'hui : perspectives didactiques*. (pp.177-199). Presses Universitaires de Namur
- Stordeur, M. F., & Colognesi, S. (2020). Transformer l'exposé oral classique en exposé oral enregistré au primaire: quelles modalités de travail de l'enseignant et quels effets pour les élèves?. *Formation et profession: revue scientifique internationale en éducation*, 28(4), 1-12. <https://doi.org/10.18162/fp.2020.683>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load*. Springer.
- Vassart, C., Blondeau, B. & Colognesi, S. (2022). Dans les coulisses de l'évaluation de l'oral par les pairs au primaire. *Éducation et francophonie*, 50(1). <https://doi.org/10.7202/1088549ar>
- Van Gennip, N. A., Segers, M. S., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, 4(1), 41-54. <https://doi.org/10.1016/j.edurev.2008.11.002>
- Van Nieuwenhoven, C., & Colognesi, S. (2015). Une recherche collaborative sur l'accompagnement des futurs instituteurs : un levier de développement professionnel pour les maîtres de stage. *e-JIREF*, 1(2), 103-121. https://dial.uclouvain.be/pr/boreal/object/boreal:165239/datastream/PDF_01/view
- Veenman, M. (2012). Metacognition in science education: Definitions, constituents, and their intricate relation with cognition. In A. Zohar, & Y.-J. Dori (Eds.). *Contemporary trends and issues in science education metacognition and learning* (pp. 21–36). Springer.
- Wiertz, C., Van Mosnenck, S., Galand, B., & Colognesi, S. (2020). Évaluer l'oral quand on est enseignant ou chercheur: points de discussion et prises de décision dans la coconception d'une grille critériée. *Mesure et évaluation en éducation*, 43(3), 1-37. <https://doi.org/10.7202/1083006ar>

Annexe 1. Grille critériée utilisée dans cette étude (première partie de la grille)

Phrase pour introduire :

Corps du discours : Informations (écris les différentes idées avec des mots-clés)	L'information est-elle correcte ?	L'information est-elle élaborée ? (est-elle accompagnée d'exemples, d'adjectifs, ...)
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI
	NON - OUI	NON - OUI

Ci-dessus, colorie de la même couleur les informations qui appartiennent à la même catégorie, au même groupe d'idée.

Lorsque tu regardes l'ensemble de l'explication, est-ce que **l'orateur** NON - OUI
fait attention à son public ?







Si oui, comment ? (tu peux entourer plusieurs réponses)







En posant des questions - En clarifiant certaines phrases - En reformulant certaines phrases – En souriant – En faisant des gestes
















Autres :

Phrase pour conclure :

Annexe 1. Grille critériée utilisée dans cette étude (deuxième partie de la grille)

1 – J'adapte ma parole à la situation de communication			
L'explication est adaptée au public (Le vocabulaire et la formulation des phrases)	Non 	Un peu 	Oui 
2 – J'élabore des significations			
L'explication est compréhensible	Non 	Un peu 	Oui 

3. J'utilise le lexique et la morphosyntaxe			
L'orateur utilise un vocabulaire ...	Pas adapté <i>(ex. : machin, truc, chose, ...)</i>		Adapté à la thématique de l'explication
L'orateur commence ses phrases sans les terminer .	Souvent 	Parfois 	Jamais 
L'orateur a des tics de langage (euh, ben, en fait, du coup, voilà, et tout ça, ...)	Souvent 	Parfois 	Jamais 

4. Utiliser et identifier les moyens verbaux et non verbaux			
Volume			
Débit (vitesse)			
Articulation			
Intonation			
Regards			
Présence de gestes	Non		Oui
Si oui , les gestes sont...	Parasites		Adéquats
Présence d'un support	Non		Oui
Si oui , le support...	Ne sert à rien	Est trop utilisé par l'orateur	Est bien utilisé

Annexe 2. Tableau reprenant les transformations d'occurrences en échelles.

Critère	Mesure	Maximum	Échelle transformée
Manière de mener l'interaction	Occurrences	7	0 : 0-1
			1 : 2-3
			2 : 4-5
			3 : 6-7
Nombre de catégories (différentes) abordées	Occurrences	5	0 : 1
			1 : 2
			2 : 3
			3 : 4-5
Nombre d'idées globales énoncées	Occurrences	22	0 : 0-5
			1 : 6-11
			2 : 12-17
			3 : 18 et +

Annexe 3. Tableaux synthèse des sous-catégories identifiées dans l'analyse qualitative

Synthèse de la catégorie « Modalités d'échange au sein du groupe »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Organisation de la prise de parole	68	19,82%	39	49,37%	30	44,12%
Revenir sur un avis antérieur	116	33,81%	16	20,25%	14	20,58%
Reformulation	15	4,37%	4	5,06%	5	7,35 %
Humour	15	3,37%	3	3,8%	4	5,88%
Rappel à l'ordre	37	10,78%	12	15,19%	4	5,88%
Interrogation	92	26,82%	5	6,33%	11	16,17%

Synthèse de la catégorie « Modalités pour faire passer le message »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Opinion personnelle	42	43,3%	25	40,32%	29	34,52%
Modérateur de fréquence et d'intensité	20	20,62%	10	16,13%	18	21,43%
Contre balancement	11	18,56%	10	16,13%	8	9,52%
Déresponsabilisation	6	6,18%	8	12,90%	11	13,09%
Amplificateur	3	3,09%	0	0%	12	14,29%
Reconnaissance de la difficulté	8	8,25%	0	0%	1	1,19%
Comparaison	0	0%	4	6,45%	2	2,28%

Synthèse de la catégorie « Nature des commentaires »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Informatifs	280	84,48%	239	85,67%	135	74,99%
Suggestifs	30	9,09%	18	6,45%	16	8,89%
Prescriptifs	20	6,06%	22	7,88%	29	16,11%

Synthèse de la catégorie « Explication du commentaire »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Extrait de la présentation	28	47,46%	10	26,32%	23	54,76%
Expérience subjective	6	10,17%	11	28,95%	12	28,57%
Comparaison	3	5,08%	4	10,53%	4	9,52%
Appui sur le jeu	8	13,56%	2	5,26%	1	2,38%
Réaction interlocuteur	5	8,47%	4	10,53%	1	2,38%
Exemple fictif	8	8,47%	6	15,79%	0	0
Environnement	1	1,69%	1	2,63%	1	2,38%

Synthèse de la catégorie « Utilisation de notes pour formuler les commentaires »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Réponse grille	495	73,71%	43	34,68%	0	0%
Grille comme checklist	135	20,18%	76	61,29%	0	0%
Organisation spatiale	23	3,44%	0	0%	0	0%
Signification d'un item	13	1,94%	2	1,61%	0	0%
Référence à un brouillon	0	0%	0	0%	13	100%

Synthèse de la catégorie « Réactions de l'orateur »

Sous-catégories	Grille imposée		Grille libre		Sans grille	
	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage	Nombre d'unités de sens	Pourcentage
Accord	88	44,22%	6	14,28%	17	32,69%
Désaccord	33	16,58%	11	26,19%	22	42,31%
Évaluation formatrice	27	13,57%	3	7,14%	4	7,69%
Interrogation	28	14,07%	8	19,05%	2	3,85%
Contenu émotionnel	3	1,51%	5	11,9%	4	7,69%
Remerciement	3	1,15%	8	19,05%	1	1,92%
Humour	4	2,01%	1	2,38%	0	0%
Proposition d'amélioration	13	6,53%	0	3,85%	2	0%