

Certifier la production écrite au secondaire et... rien d'autre !

Raphaël Pasquini et Sonya Florey

Résumé

Depuis l'introduction du plan d'études romand (PER) en 2011, de nombreux enseignants relèvent des difficultés à certifier la production écrite, tous genres confondus. Un des enjeux pour eux consiste à articuler l'évaluation du fonctionnement de la langue aux apprentissages spécifiques à la production écrite et à formuler des critères d'évaluation centrés sur les contenus propres à l'objet didactique. Pour aborder cette problématique, notre article présente une étude de cas analysant à l'aide d'un modèle théorique une démarche d'évaluation sommative notée au secondaire, dans le canton de Vaud. Cette analyse s'inscrit dans un questionnement professionnel permettant d'aboutir à de nouvelles réflexions, pour que les enseignants certifient la production écrite, et rien d'autre.

Mots-clés

Production écrite, évaluation sommative, alignement curriculaire élargi, parodie de conte.

⇒ *Titel, Lead und Schlüsselwörter auf Deutsch am Schluss des Artikels*

⇒ *Titolo, riassunto e parole chiave in italiano alla fine dell'articolo*

Auteur·e

Raphaël Pasquini, Haute École pédagogique du canton de Vaud, Avenue de cour 33, 1014 Lausanne, raphael.pasquini@hepl.ch

Sonya Florey, Haute École pédagogique du canton de Vaud, Avenue de cour 33, 1014 Lausanne, sonya.florey@hepl.ch

Certifier la production écrite au secondaire et... rien d'autre !

Raphaël Pasquini et Sonya Florey

1. Problématique

Des travaux de recherche et d'expérience en formation continue montrent que, depuis l'introduction de Plan d'Études Romand (PER) en 2011, bon nombre d'enseignants rencontrent des problèmes à certifier les apprentissages de leurs élèves en production écrite. Plus précisément, leurs problèmes sont de trois ordres.

Le premier est relatif à la difficulté d'exploiter le PER comme un référentiel d'évaluation alors qu'il est un référentiel d'enseignement et d'apprentissage (Carulla et al., 2013). En effet, lorsque les enseignants s'y réfèrent – ce qui n'est pas toujours le cas – ils peinent souvent à en comprendre la logique pour organiser leurs séquences, les énoncés étant décrits comme jargonneux et les contenus parfois incomplets (Pasquini, 2018). La discipline du français constitue d'ailleurs un cas emblématique, pourvoyeur de brouillages potentiels pour les enseignants : les éléments de la progression des apprentissages appartiennent à six épistémologies différentes ; des compétences ou des notions disciplinaires sont juxtaposées comme si elles constituaient des apprentissages de même niveau de complexité ; enfin, certains éléments de la progression des apprentissages apparaissent sous une forme complémentaire dans deux objectifs d'apprentissages distincts, ce qui demande aux enseignants une connaissance fine des objets disciplinaires, mais également de l'organisation interne du PER afin de construire leur séquence (Balsiger et al., 2011; Florey & Depallens, 2020). Ainsi, une traduction voire une reformulation du curriculum est régulièrement requise des enseignants, ce qui est perçu comme un surplus de travail. Dès lors, ce sont d'autres sources considérées comme plus pratiques mais parfois éloignées des contenus du PER qui jouent le rôle de référentiel (listes d'objectifs, manuels scolaires, sites Internet, ressources diverses provenant de collègues). Cet écueil est délicat du point de vue de l'enseignement, et d'autant plus problématique lorsqu'il s'agit de certifier : la pratique observée revient à « copier-coller » des objectifs du PER sur l'épreuve construite, de manière à se conformer aux prescriptions légales. Il existe alors un décalage entre les objectifs donnés comme évalués et ce qu'évalue réellement l'épreuve de production écrite au travers des situations proposées aux élèves¹.

Le deuxième problème concerne la didactisation cohérente des divers objets disciplinaires qui sont décrits dans le PER pour travailler la production écrite, jusqu'à la certification. Ce problème est particulièrement aigu lorsqu'il s'agit de certifier des objets disciplinaires relatifs au fonctionnement de la langue au sein même d'une situation de production écrite. Pourtant, cette perspective s'inscrit dans les visées du PER et est encouragée par les approches didactiques : elle peut donc se concevoir du point de vue de l'enseignement et de l'apprentissage, mais elle met en évidence des pratiques évaluatives sommatives² discutables. En effet, de nombreux élèves se retrouvent en échec car ils sont sanctionnés sur des objets grammaticaux, syntaxiques ou lexicaux tout en faisant preuve d'acquis significatifs en production écrite. La situation contraire existe aussi : par exemple, des élèves réussissent davantage grâce à leurs compétences en orthographe qu'à leur capacité à construire un schéma narratif cohérent. Dans tous les cas, de telles pratiques biaisent la cohérence des épreuves et favorisent une « pollution des notes » (Guskey & Link, 2019 ; Pope, Greene, Johnson & Mitchell, 2009) car le résultat final ne permet pas de donner des informations pertinentes sur ce que l'élève maîtrise ou non en termes de compétence à produire un texte.

Le troisième problème identifié réfère à la conception du texte qui émerge des différentes manières de certifier. Dans l'idéal, la grille d'évaluation doit permettre à l'élève de produire un texte traversé par des dimensions de contenu ou de sens, ainsi que par des dimensions formelles, le fonctionnement de la langue, par exemple. A ce titre, la définition de Halliday et Hasan (1976) précise qu'« un texte ne doit pas du tout être vu comme une unité grammaticale, mais comme une unité d'une autre espèce : une *unité sémantique* ».

¹ Cette pratique se retrouve toutefois dans toutes les disciplines.

² Nous avons choisi ce terme car les travaux anglo-saxons dont nous nous inspirons parlent de *summative assessment*, c'est-à-dire une démarche relevant d'un bilan portant sur les apprentissages essentiels (*to sum up*). Dans cette acception, l'évaluation sommative se définit principalement au regard de sa temporalité et de sa fonction première de certification : « L'évaluation sommative est une évaluation qui est effectuée après l'apprentissage pour certifier ce qui a été appris » (Brookhart, 2017, p. 8, trad.libre).

Son unité est une unité de sens en contexte, une texture qui exprime le fait que, formant un tout [as a whole], il est lié à l'environnement dans lequel il se trouve placé" (p. 293, trad. libre de J.-M. Adam, 2008, p. 10). Or, selon la manière dont les enseignants construisent leurs grilles d'évaluation, les textes produits par les élèves peuvent tantôt correspondre à la définition de Halliday et Hasan, tantôt à une définition qui considère le texte comme une addition de multiples éléments, circonscrits et segmentés. C'est notamment le cas lorsque des grilles sont articulées autour de critères quantifiés, par exemple "présence de trois adjectifs de couleur", pensant qu'on peut ainsi évaluer la maîtrise de l'élève sur ce point spécifique de grammaire.

Pragmatiquement et face à ces trois problèmes, les questions des enseignants touchent plusieurs niveaux : comment construire des situations de production écrite qui font sens, qui n'orientent pas trop les élèves sur une performance attendue tout en visant des apprentissages ciblés ? À quelles conditions formuler des critères d'évaluation qui permettent d'être en accord avec les attentes relatives à la qualité d'un texte ? Comment conjuguer l'évaluation des caractéristiques propres de la production écrite et celles des aspects plus techniques de la langue ? Ou encore, comment noter les productions des élèves de manière à être « justes » ?

Dès lors, à quelles conditions une situation d'évaluation sommative sur un tel objet didactique peut-elle être cohérente, des objectifs évalués au sein du curriculum jusqu'à la notation ? Comment gérer le dilemme consistant à évaluer simultanément deux objets aussi articulés (le fonctionnement de la langue et la production écrite) ? Quels apports théoriques permettent aux pratiques évaluatives et de notation de contourner ces écueils tout en gagnant en cohérence ?

Cet article s'empare de ces questions au travers d'une étude de cas. Pour cela, nous commençons par dresser une revue de la littérature sur l'évaluation en français et notamment en production écrite. Puis nous présentons un cadre théorique éprouvé en formation et en recherche pour penser la cohérence des pratiques en évaluation sommative. Viennent ensuite la présentation et l'analyse d'une épreuve de production écrite au secondaire qui soumet des pistes de réflexion constructives en rapport avec notre questionnement, et enfin, une discussion donnant à voir les aspects positifs et critiques d'une telle démarche.

2. État de la recherche

Une abondante littérature montre que les disciplines impactent les pratiques d'évaluation sommative (par ex., Cross & Frary, 1999 ; Duncan & Noonan, 2007 ; Mottier Lopez & Laveault, 2008 ; Prøitz, 2013 ; Resh, 2009 ; Simard, Dufays, Dolz & Garcia-Debanc, 2010 ; Zhang & Burry-Stock, 2003). Il en va de même pour les pratiques de notation, notamment en langues (Brimi, 2011 ; Hillocks, 2006). Comme le disent Braxmeyer, Guillaume et Levy (2005) : « il y aurait bien des comportements évaluatifs disciplinaires spécifiques en évaluation » (p. 52). Les pratiques d'évaluation devraient donc se diversifier en fonction, notamment, « des choix didactiques et des objectifs qui en découlent » (Mottier Lopez & Laveault, 2008, p. 13), en tenant compte des spécificités disciplinaires, mais aussi des caractéristiques des objets didactiques évalués dans une même discipline. En d'autres termes, la certification de la production écrite en français gagnerait à être spécifique et donc différente, par exemple, de celle de la structure de la langue ou de la compréhension écrite.

Force est de constater qu'à ce jour, les travaux francophones en didactique traitant de ces questions sont encore assez rares. Millon-Fauré (2013) va dans ce sens : « Bien qu'il s'agisse d'un rituel quasi-quotidien de la classe, les évaluations constituent encore des dispositifs peu étudiés en didactique » (p. 149). Le rôle de la didactique dans l'évaluation sommative semble pourtant de plus en plus partagé au sein de la communauté scientifique. Dans un texte intitulé « L'évaluation du travail des élèves » – considéré comme une référence par de nombreux auteurs consultés – Simard, Dufays, Dolz et Garcia-Debanc (2010) abordent la question de l'évaluation sommative en français et discutent les conditions de son efficacité. L'évaluation sommative est ici bien mise en relation avec les processus d'enseignement et d'apprentissage. Toutefois, les grilles de critères que les auteurs proposent visent à correspondre davantage aux effets attendus dans la production qu'à la certification des apprentissages réellement menés. Gagnon, Monnier et Dolz (2014), à ce sujet, affirment :

Les grilles permettent de stabiliser les critères, les indicateurs précis à utiliser pour apprécier le degré d'adéquation à la tâche d'écriture proposée, ainsi que la pondération à accorder aux différents critères

sélectionnés. Elles constituent ainsi une aide non seulement pour fixer le seuil de réussite, mais également pour communiquer le résultat à l'élève (p. 2).

Dans cette perspective, l'évaluation sommative devrait porter sur des savoirs déclaratifs convoqués pour être « utilisés » en même temps que des savoir-faire, dans des tâches non fragmentées, complexes, porteuses de sens et, si possible, « dotées d'une dimension épistémologique et/ou de plaisir » (Simard et al., 2010, p 19). Cette conception de l'évaluation sommative gagnerait à s'inscrire selon nous dans un souci de conceptualisation, tant du point de vue des paradigmes dans lesquels elle s'inscrit (De Ketele, 2012 ; Figari, 2012, 2013), que des approches théoriques dont découlent les outils exploités. Cette hypothèse se confirme avec l'étude des travaux de Tobola Couchepin, Sanchez Abchi et Dolz (2016) sur les pratiques d'évaluation sommative de la production écrite au primaire : la finesse de leur analyse didactique souffre d'un manque de conceptualisation des grilles critériées utilisées comme outils évaluatifs.

Ces approches didactiques font par ailleurs l'impasse sur un élément clé de toute épreuve sommative : la construction de la note, qui est toujours présente dans le questionnement des enseignants (Pasquini, 2017). Dans les sources consultées, les auteurs proposent effectivement des modèles de grilles de critères qui tiennent compte d'une forme de pondération, mais la manière de construire la note en rapport avec les critères n'est pas développée. Dans la même ligne, en 2016, la collection « Recherches en didactique du français » porte sur l'évaluation. On y revendique une approche critique que certains auteurs thématisent au niveau de l'évaluation sommative. Par exemple, Mottier Lopez, Serry et Sales Cordeiro analysent la production écrite du fait divers au primaire. Si leurs propos sont ici très rigoureux jusqu'à la pondération, rien n'est dit, encore une fois, sur les processus d'élaboration de la note. Enfin, la même année, un ouvrage important sur le jugement professionnel³ rend compte de recherches dont certaines ont pour ancrage l'évaluation sommative. On peut y lire un texte très abouti de Goasdoué, Vantourout et Bedoin (2016) qui décrit les composantes du jugement professionnel en jeu dans la correction de dissertations en français. Les auteurs affirment que « la note n'est pas la somme des jugements produits lors de la correction. Des appréciations semblables peuvent conduire à des notes divergentes et inversement des appréciations divergentes peuvent se solder par des notes voisines » (p. 75). On souhaiterait toutefois en savoir davantage sur les outils auxquels ils se réfèrent, ou sur une éventuelle conceptualisation des pratiques de notation.

Au vu de ces espaces à investir, et pour comprendre la cohérence des démarches d'évaluation sommative en général et particulièrement en français et en production écrite, il devient important de nous doter de modèles théoriques permettant à la fois de donner une importance aux objets didactiques et à leurs caractéristiques, mais aussi aux processus de pondération et de notation, omniprésents en certification. Nous avons opté pour le modèle de l'alignement curriculaire, et plus spécifiquement sa conceptualisation élargie (Pasquini, 2018, 2019b).

3. Cadre théorique

La littérature qui porte sur l'alignement curriculaire considère que toute pratique d'évaluation doit être en forte cohérence avec les objectifs prescrits, l'enseignement dispensé, les tâches d'apprentissage proposées aux élèves et les activités effectivement réalisées par ces derniers (par ex., Anderson, 2002 ; Biggs, 1999, 2003 ; Gauthier, Mellouki, Bissonnette & Richard, 2005 ; Martone & Sireci, 2009). Anderson (2002) définit l'alignement curriculaire comme : « le niveau de cohérence entre les objectifs du curriculum et l'évaluation, entre les objectifs du curriculum et les activités pédagogiques, et enfin entre l'évaluation et les activités pédagogiques » (p. 257, trad. libre). La validité du contenu, la façon dont il est enseigné, ainsi que les opportunités d'apprentissage de ce contenu offertes aux élèves au travers de tâches d'apprentissage font également partie du modèle. L'auteur le schématise sous forme de triangle dont chaque sommet en est une composante :

³ « Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages », coordonné par Lucie Mottier Lopez et Walther Tessaro, publié chez Peter Lang.

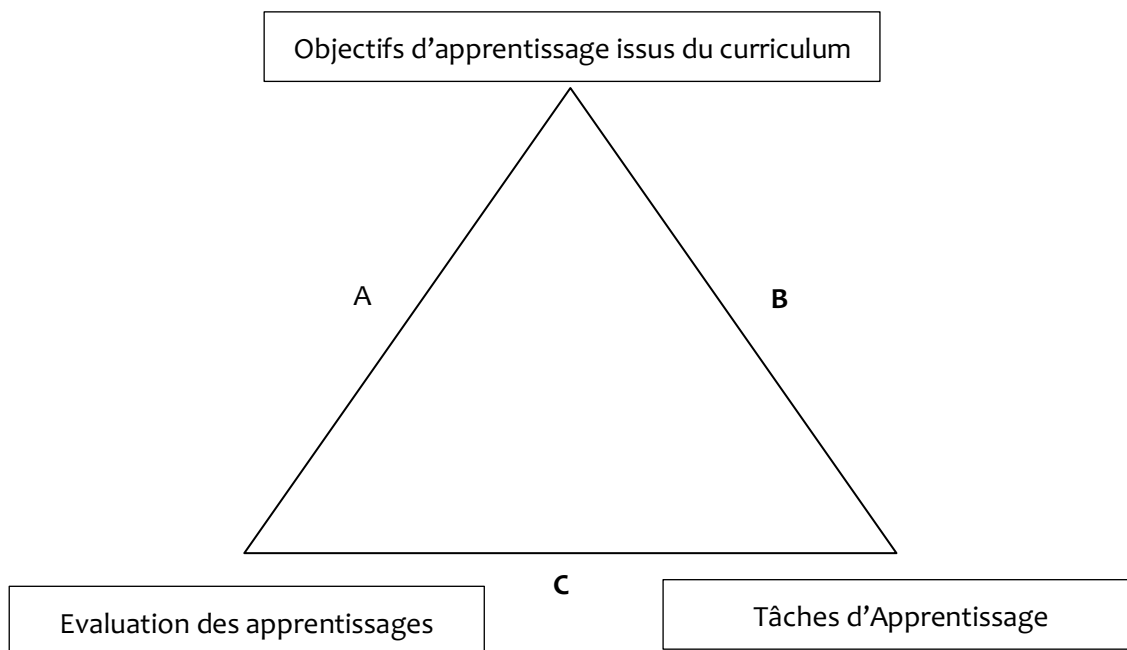


Figure 1 : L'alignement curriculaire selon Anderson (2002, p. 256)

Dans cette perspective, l'élaboration d'épreuves évaluatives sommatives est conçue comme un processus complexe, dynamique, contextuel et centré sur les apprentissages. La notion d'*alignement* relève d'une conceptualisation de la cohérence à l'aune d'une taxonomie d'habiletés cognitives appariées à des contenus, ici, celle d'Anderson et Krathwohl (2001)⁴. La dimension *curriculaire* renvoie au fait que cette cohérence est pensée depuis un curriculum.

Relativement à notre problématique, ce modèle invite à formuler deux questions permettant d'objectiver la cohérence de toute démarche d'évaluation sommative : dans la relation A, « dans quelle mesure le test évalue-t-il les objectifs importants du curriculum ? » (p. 255, trad. libre) et dans la relation C, « ce qui a été enseigné est-il testé ? » (p. 255, trad. libre). Ce questionnement est un premier pas mais il reste incomplet pour analyser finement les processus de pondération – la manière dont le poids est alloué aux apprentissages évalués au travers de critères et/ou de points – et les pratiques de notation en référence à ces mêmes apprentissages, deux composantes caractéristiques d'épreuves sommatives (Anderson, 2002 ; Biggs, 2003 ; Marzano, 2002 ; Brookhart, 2005).

Il devient dès lors nécessaire d'élargir conceptuellement le modèle d'Anderson pour mieux comprendre la cohérence en évaluation sommative, car les pratiques de notation et les pratiques d'évaluation sommative sont fortement articulées (Harlen, 2005 ; McMillan & Nash, 2000 ; Pasquini, 2018 ; Randall & Engelhard, 2010 ; Walvoord & Johnson Anderson, 2009). Par exemple, ces travaux documentent que la cohérence ne peut se penser ni se conceptualiser sans y intégrer le design des tâches évaluatives (Wiggins & McTight, 2005), la formulation des critères (Jönsson, 2014) et la construction d'une note cette fois référée à l'apprentissage pour être constructive et à haute valeur informative (Hadjji, 2016, 2017).

⁴ Dans cette perspective, deux éléments clés émergent : la complexité de l'habileté cognitive en jeu, et conjointement, les caractéristiques de l'objet disciplinaire sur lesquelles elle porte (McGrath & Noble, 2008). La première dimension renvoie à l'activité mentale de l'élève au travers des six niveaux d'habiletés cognitives de la taxonomie (se souvenir, comprendre, appliquer, analyser, évaluer ou créer). La seconde se rapporte au contenu sur lequel porte l'apprentissage en question : la phrase, le schéma quinaire, etc.. La complexité de l'objectif dépend donc de ces deux dimensions (par exemple : rédiger une synthèse de roman), qui devraient, idéalement, permettre d'établir les relations de cohérence entre chaque composante du modèle.

Ainsi, trois sources ont participé à établir ce modèle élargi :

- Les phénomènes relevés dans les pratiques d’enseignants de français du secondaire (Pasquini, 2018) ;
- Les apports d’une littérature ciblée et épistémologiquement cohérente sur la pondération et la notation ;
- Deux caractéristiques du modèle de départ d’Anderson (2002) : une perspective dynamique pour penser les relations entre ses différentes composantes « objectifs du curriculum – tâches d’apprentissage – évaluation » ; l’inscription première du modèle dans les processus d’enseignement-apprentissage.

Concrètement, nous avons porté notre attention sur plusieurs principes :

- Les critères de l’épreuve décrivent les objectifs évalués (ou annoncés comme tels) ;
- Les objectifs sont évalués dans les tâches de l’épreuve ;
- Les critères portent sur des dimensions de l’apprentissage présents dans les tâches de l’épreuve, et sont aussi en lien avec les seuils des échelles de notation. À ce niveau, il s’agit de voir dans quelle mesure les contenus et habiletés cognitifs décrivant l’objectif d’apprentissage évalué peuvent caractériser les seuils de réussite dans une échelle de notation, elle-même construite en référence aux tâches évaluatives et/ou aux critères d’évaluation qui devraient, idéalement, les rendre explicites.

L’alignement curriculaire élargi (Pasquini, 2018, 2019b) permet donc la compréhension des phénomènes de cohérence ou d’incohérence – ou de désalignements curriculaire (Airasian & Miranda, 2002) – perceptibles dans les épreuves évaluatives sommatives notées, dans une perspective systémique et contextuelle. Il se schématise de la manière suivante :

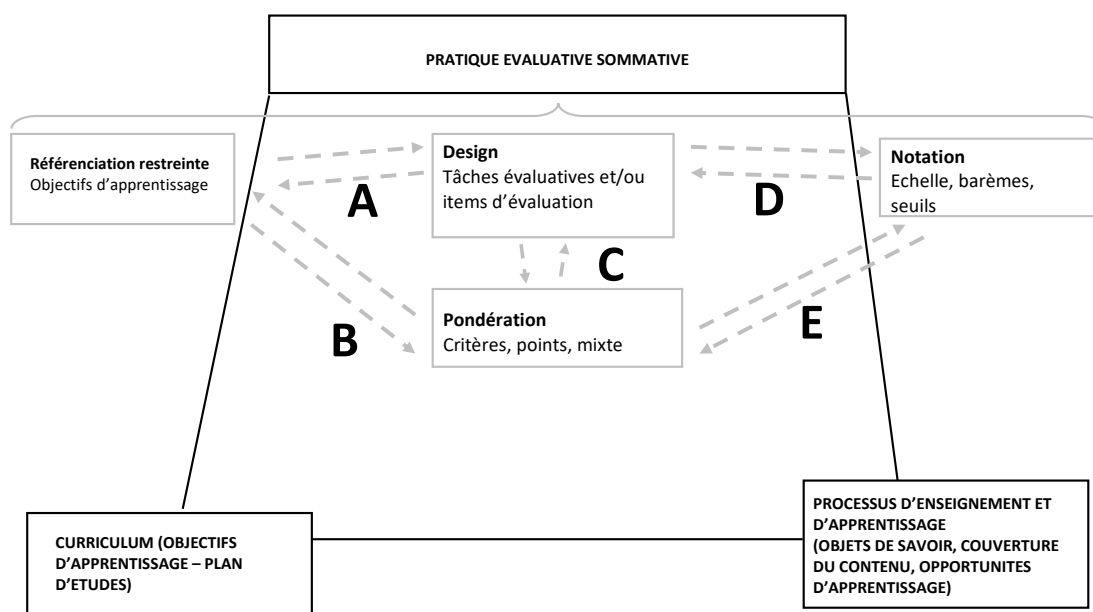


Figure 2 : L’alignement curriculaire élargi (Pasquini, 2018, 2019b)

Tout en s’inscrivant au cœur du modèle d’Anderson (2002), ce modèle élargi déploie les quatre composantes de l’évaluation sommative (accolade, en clair) en les mettant en liens réciproques (double traits clairs, pointillés). Reprenons-les successivement :

- La référenciation restreinte, inspirée de Vial (2012), décrit ce qui fait foi chez les enseignants quand ces derniers opèrent des choix au niveau des objectifs d’apprentissages à évaluer d’un point de vue sommatif. Cette idée de choix est fondamentale. Elle suppose qu’une situation d’évaluation ne porte que sur une partie des apprentissages menés en classe, et non sur leur totalité. Cet élément est particulièrement important pour comprendre comment il est possible de gérer l’évaluation du fonctionnement de la langue en production écrite. En effet, le risque de tout évaluer dans une seule

épreuve sommative est de perdre de la qualité dans les informations recueillies, ces dernières étant trop disparates.

- Le « *design* » (Wiggins & McTighe, 2005) explicite les processus de construction des tâches évaluatives au niveau des contenus et des habiletés cognitives, notamment du point de vue des consignes. Pour l'enseignant, cela supposera entre autres de prendre conscience du niveau élevé de complexité en jeu dans la production écrite.
- La pondération renvoie au poids alloué aux apprentissages et aux outils exploités pour ce faire (des critères, des points, ou des critères appariés à des points comme modalité mixte, très courante en production écrite) (Andrade, 2005 ; Jönsson, 2014). Cette étape est cruciale si l'on souhaite garder une perspective qualitative de l'évaluation.
- Enfin, la notation cible les démarches de construction de la note chiffrée en lien avec l'apprentissage à l'aide d'outils comme les échelles ou les barèmes (Brookhart, 2017 ; Marzano & Heflebower, 2011).

Ce sont donc maintenant non plus trois, mais cinq relations entre les différentes composantes qui doivent être cohérentes d'un point de vue taxonomique et de contenu, possiblement traduisibles au travers de questions non exhaustives comme :

- Les objectifs formulés dans l'épreuve sont-ils évalués au travers de la tâche de production ? La tâche de production rend-elle compte des objectifs évalués ? (relation A) ;
- Les objectifs formulés dans l'épreuve se retrouvent-ils dans les critères ? Les critères décrivent-ils les objectifs évalués ? (relation B) ;
- La tâche donne-t-elle lieu à un poids alloué à l'apprentissage cohérent ? Le poids des apprentissages évalués correspond-il à ce que la tâche évalue en termes de complexité de contenu et d'habileté cognitive ? (relation C) ;
- Les contenus et habiletés présents dans la tâche se retrouvent-ils dans les descripteurs des échelles ? Les échelons des échelles donnent-ils à voir les contenus et habiletés issus de la tâche ? (relation D) ;
- Relativement à la notation, les critères permettent-ils l'établissement des échelons, surtout le seuil de suffisance ? Les échelons et leurs progressions sont-ils en rapport avec les critères ? (relation E).

Dans cette nouvelle acception compréhensive, la cohérence curriculaire de toute démarche d'évaluation sommative se lit au travers de multiples triangulations entre plusieurs composantes du modèle, jusqu'à la notation. Aussi, à l'instar du modèle de départ, plusieurs entrées peuvent orienter le questionnement. Un enseignant peut donc se sentir plus à l'aise de questionner la cohérence depuis les objectifs évalués et leurs liens avec la tâche de production et/ou les critères. Par exemple, relativement à un texte narratif, ce travail l'amènera à décrire l'objectif « utiliser un schéma narratif » en critères et en consignes de production, en lien avec l'enseignement qu'il aura dispensé et aux apprentissages réalisés par les élèves. Un autre collègue partira quant à lui des contenus et consignes de la tâche évaluative, et les mettra en lien avec les objectifs et critères de production.

L'alignement curriculaire élargi désigne donc, en évaluation sommative, le fort lien de cohérence systémique existant entre les objectifs évalués (référenciation restreinte), les tâches évaluatives (*design*), les points et/ou les critères attribués (pondération) et l'échelle (notation) (Pasquini, 2019b, pp. 83-84).

Rappelons que le modèle poursuit une visée compréhensive et non prescriptive : le but est bien de permettre aux enseignants une nouvelle intelligibilité de la cohérence de leurs épreuves de production écrite, et non de leur prescrire une manière de procéder pour obtenir de « bons » modèles d'épreuves, interchangeables et reproductibles. Pour mieux comprendre la complexité de cette approche, arrêtons-nous à présent en détail sur un exemple de pratique.

4. Étude de cas

La démarche présentée a été élaborée par Line⁵, une enseignante du secondaire, lors d'une formation continue à l'évaluation dispensée à un groupe de collègues volontaires de diverses disciplines d'un même établissement⁶. Cette épreuve de production écrite sur la parodie du conte – appelée Travail Significatif dans le canton de Vaud (TS) – se destine à des élèves de 9^e année du secondaire (moyenne d'âge 13 ans) de voie Prégymnasiale (la plus exigeante du système vaudois sur les deux existantes).

4.1 La parodie, caractéristiques du genre

La parodie de conte n'est pas explicitement nommée dans les genres conseillés par le PER au cycle 3, mais permet, au même titre que d'autres genres, de travailler des compétences liées au texte narratif. Arrêtons-nous un instant sur la définition de la parodie et sur la pertinence à travailler ce genre relativement au conte. Pour Sangsue (2007), le sens de la parodie est à rechercher dans son étymologie : « para- » signifiant à la fois « à côté » et « contre », la parodie est un genre littéraire impliquant la réécriture ludique d'un texte, combinant proximité et distance. Autrement dit, pour qu'une parodie fonctionne, le lecteur doit être capable de déceler le texte-source, objet de la parodie, et simultanément d'identifier les évolutions du texte-source (Sangsue, 2007). Reuter (1984) définit quant à lui la parodie comme « la production d'un texte (B) où l'on doit reconnaître facilement un texte ou un genre (A) transformé à des fins humoristiques » (p. 82). Des différentes caractéristiques de la parodie nous retiendrons les suivantes, qui semblent particulièrement adaptées à un travail en contexte scolaire : la transposition spatio-temporelle, la dégradation et l'inversion (Reuter, 1984 ; Petitjean, 1984). A la place du lieu et du temps indéterminés du conte-source, la parodie situe l'histoire dans un espace-temps situé. Dans le cas d'une actualisation du récit, le comique est porté par un décalage entre des éléments jugés désuets et le contexte contemporain dans lequel ces derniers sont transposés. La trivialisation qui s'ensuit apparaît comme une forme de dégradation de l'histoire. L'inversion, enfin, peut porter sur des associations de mots, des contrepèteries, mais également sur les personnages, les situations et les valeurs⁷.

Si l'on considère à présent la parodie dans la classe de français, Connan-Pintado (2009) énonce cinq compétences inhérentes au travail sur la parodie de conte à développer chez les élèves, justifiant ainsi l'intérêt à en faire un genre scolaire :

- linguistique, car les élèves doivent porter une attention soutenue aux formes du texte-source et du texte second ;
- encyclopédique, lorsque les élèves réalisent une transposition spatio-temporelle ;
- logique, lorsqu'ils relient entre eux ces éléments issus du texte-source et du texte second ;
- rhétorique, s'agissant de s'interroger sur le nouveau genre et sur la réécriture ;
- idéologique, concernant les nouvelles valeurs qui s'incarnent dans le texte second.

Dans cette perspective, l'enseignement de Line a porté sur les éléments caractérisant une parodie, et sur l'étude compréhensive des deux versions du Petit Chaperon Rouge (celle des frères Grimm et de Perrault).

4.2 Certifier la parodie

En ce qui concerne la situation de certification à proprement parler, les élèves reçoivent deux documents : une feuille de consignes (figure 3) et une grille de critères (figure 4).

5 Prénom d'emprunt.

6 Elle a également été présentée par l'enseignante à la journée d'étude « Certifier en soutenant les apprentissages ? Enjeux et débats » organisée par Collectif Romand d'Etudes des Pratiques Evaluatives (CREPE) à la HEP Vaud en avril 2018.

7 Par exemple, un Chaperon Rouge qui berne un loup devenu végétarien, comme dans l'album de de Lestrade et Bravo « Le plus gentil loup du monde », publié en 2005 aux éditions La joie de lire.

| | | |
|--------------|--|----------|
| Français 9VP | TS Expression écrite – La parodie de conte | Prénom : |
|--------------|--|----------|

Pour ce TS, tu devras écrire un conte parodique du *Petit Chaperon Rouge*, que tu connais déjà. Tu as le droit à tout moment de consulter les deux versions de ce conte (celle de Perrault et celle des frères Grimm). **Ton texte fera minimum 200 mots.**

- Choisis **une** des deux versions comme texte de base pour ta parodie.
- Liste ensuite les modifications que tu veux apporter à l'histoire du Petit Chaperon Rouge : les étapes principales du schéma narratif doivent être respectées pour que l'on puisse reconnaître le conte source, mais tout le reste peut être modifié.
- Dans le premier paragraphe, soigne la description de ton personnage principal : quels sont ses caractéristiques ? dans quel univers, quel décor évolue-t-il ?...
- Le système de temps utilisé dans ton texte sera **le système du passé** : passé simple / imparfait / plus-que-parfait pour marquer l'antériorité. D'autres temps sont acceptés dans les dialogues! Tu peux utiliser le tableau collé dans ton cahier de structuration pour t'aider.
- Structure ton texte à l'aide de paragraphes et de mots organisateurs (marqueurs temporels et connecteurs).
- Si tu insères des dialogues dans ton texte, fais-le selon les règles.
- Utilise un registre courant (ou soutenu) mais pas familier, même dans les dialogues.
- Soigne l'orthographe, la ponctuation, la syntaxe.
- Trouve un nouveau titre en fonction de ton histoire : là aussi, on doit pouvoir faire le lien avec *Le Petit Chaperon Rouge*.
- Commence par faire un brouillon avec tes idées, puis rédige la version finale.

Matériel autorisé :

- Dictionnaire
- Bescherelle conjugaison + Bescherelle Collège
- Les deux versions du *Petit Chaperon Rouge*
- Le cahier de structuration

Figure 3 : Consignes distribuées aux élèves pour la rédaction de la parodie de conte

Line décrit un rapport « distancé » au curriculum : elle en connaît bien le contenu et il est une référence pour organiser son enseignement. Mais lorsqu'elle construit ses épreuves sommatives avec sa collègue de discipline, elle ne se sent pas obligée de s'y référer explicitement. C'est une des raisons pour lesquelles elle privilégie une communication des apprentissages poursuivis aux élèves durant les leçons, et qu'il n'y a pas, littéralement, de formulation d'objectifs issus du PER sur la feuille de consignes. Toutefois, les élèves ont reçu la grille d'évaluation avant le jour de l'épreuve, de manière à se familiariser avec son contenu et sa logique, et à faire des liens avec l'enseignement dispensé sous le guidage de l'enseignante.

L'analyse de la démarche montre que les objectifs se retrouvent en lien avec les consignes, et surtout avec les critères. L'objectif principal est d'« écrire un texte parodique adapté aux situations d'énonciation ». Les objectifs plus fins correspondant qui se déclinent dans les critères sont les suivants : respecter un schéma narratif, décrire des personnages, structurer un texte, organiser des dialogues, respecter un registre de langue, utiliser une orthographe et une syntaxe au service de la compréhension, ou encore utiliser le système du passé. Les objectifs correspondent aux caractéristiques du genre, impliquant un travail de didactisation, même si l'enseignante ne semble pas avoir insisté sur la dimension ludique dans le texte parodique – autant d'éléments présents formellement dans le PER.

Line organise dans le même temps sa grille d'évaluation (figure 4) avec six critères obligatoires décrivant ce qu'elle appelle « le fond », « la forme », la « langue » et « le système de temps » au sein du texte, et quatre critères de perfectionnement. Nous observons ici qu'elle adopte une logique où les seconds sont des versions complexifiées des premiers.

| Evaluation – TS Expression écrite : parodie de conte | | |
|--|---|---|
| | Critères obligatoires Tous les critères doivent être cochés pour avoir 4 | Critères d'approfondissement Pour chaque critère coché, obtention d'un demi-point supplémentaire |
| 4 | Le schéma narratif du texte original est respecté (ses étapes principales figurent dans la parodie) et l'on perçoit le lien avec le conte source. Le personnage principal ainsi que son univers sont décrits. | Les liens entre le conte source et la parodie sont subtils et ne se limitent pas à des caractéristiques évidentes. Les personnages sont intéressants : leur description est riche et détaillée ; les actions qu'ils accomplissent font avancer le récit. |
| 3 | Le texte est structuré par des paragraphes et par des mots organisateurs (marqueurs temporels et connecteurs). Lorsqu'il y a des dialogues, ils sont insérés correctement. Le registre de langue est courant ou soutenu. | Le texte est divisé de manière pertinente. Les mots organisateurs utilisés sont variés et toujours adéquats. Les dialogues apportent une dimension intéressante au texte et font avancer le récit. |
| 2 | L'orthographe et la syntaxe sont généralement correctes, elles n'empêchent pas la compréhension du texte. | |
| 1 | Le système de temps du passé est utilisé correctement : p. simple / imparfait / p-q-parfait pour marquer l'antériorité, D'autres temps sont acceptés dans les dialogues. | Le système de temps du passé est utilisé sans erreurs dans l'ensemble du récit. |

Fig 4 : Grille d'évaluation critériée donnée aux élèves

Avec cette manière de procéder, il n'y a pas véritablement d'échelle de notation, puisque la validation de tous les critères obligatoires donne une note de 4, et que les notes et demi-notes supérieures s'obtiennent chaque fois qu'un critère de perfectionnement est validé. Pour l'insuffisance, le raisonnement de Line suit la même logique : à chaque critère obligatoire non validé correspond une « descente » d'une demi-note sur l'échelle.

4.3 Analyse

Nous observons ici une épreuve peu « formelle », dans le sens où les objectifs du curriculum n'y sont qu'indirectement mentionnés. Cette pratique reste tout de même innovante, dans la mesure où, comme nous le montrons plus loin, elle se différencie des pratiques courantes. La cohérence d'alignement curriculaire élargi est optimale : il existe en effet un fort rapport de cohérence du point de vue des habiletés cognitives et du contenu entre les apprentissages évalués, les consignes de la tâche de production, les critères et l'échelle de notation. C'est bien l'évaluation de la qualité de la production qui est recherchée ici, d'où l'abandon des points. A ce niveau, nous notons qu'une des caractéristiques fondamentales de la parodie – écrire un texte qui se distingue du texte-source tout en permettant d'identifier ce dernier – est placée dans les critères obligatoires, ce qui renforce la cohérence évaluative.

Fait intéressant allant dans ce sens, Line propose aux élèves, en début d'épreuve, de les guider dans la mise en place du récit, de manière à ce qu'ils ne puissent pas partir sur une fausse voie. En leur distribuant un tableau comparatif (figure 5), elle se donne le droit d'interagir avec eux sur leurs choix et de les valider après argumentation afin qu'ils puissent rédiger leur parodie du Petit Chaperon Rouge. Cette décision émane des choix qu'elle a posés du point de vue des contenus à évaluer. Ces outils peuvent mener à la crainte de « donner les réponses » aux élèves. Or, de notre point de vue, cette crainte est infondée. Comme le souligne Line, l'enjeu, ici, est d'orienter les élèves sur les caractéristiques de l'objet et sur la manière de rendre une parodie opératoire. Autrement dit, de s'assurer que les élèves connaissent les caractéristiques de la parodie de conte, mais également leur mise en œuvre dans le texte, à l'aide d'outils tangibles. Cette manière de faire rend possible l'explicitation des éléments de contenu dans la situation d'évaluation, sans mettre en danger la pertinence de la certification : ce n'est pas parce qu'un élève a réalisé avec succès la comparaison d'éléments de la liste qu'il maîtrisera *de facto* l'ensemble des processus pour écrire un texte.

| Éléments du conte source | Éléments modifiés dans mon conte |
|---------------------------------------|---|
| Cendrillon a 2 méchantes belles-sœurs | Cendrillon a deux méchants beaux-frères |
| | |
| | |

Fig 5 : Tableau fourni aux élèves en début de certification pour poser les éléments parodiques

Cette manière d'évaluer permet à Line de fixer les exigences, de clarifier les attentes, d'explicitier les spécificités des contenus et après correction de réaliser des feedbacks ciblés sur les apprentissages réussis et non maîtrisés, autorisés par la haute valeur informative de la note. La figure 6 exemplifie bien ces possibles : la production de l'élève, qui met en évidence des acquis en fonctionnement de la langue, montre toutefois des insuffisances importantes en production écrite. En effet, l'échec se justifie ici car il porte sur la maîtrise insuffisante du schéma narratif et sur l'incomplétude de la description du personnage. Line, en termes de feedback, écrit : « ton personnage principal n'est pas décrit, tu donnes son nom sans préciser rien d'autre, tu ne fais pas le lien entre son caractère ou son physique et son surnom. D'autre part, il manque toute une partie où le dragon arrive chez la princesse et se fait passer pour Capo Rosso ; puis plus loin le fameux dialogue (« comme tu as de grands yeux »...) à modifier selon ton histoire. Ton texte est plus un résumé qu'une parodie, c'est dommage ! » (figure 7).

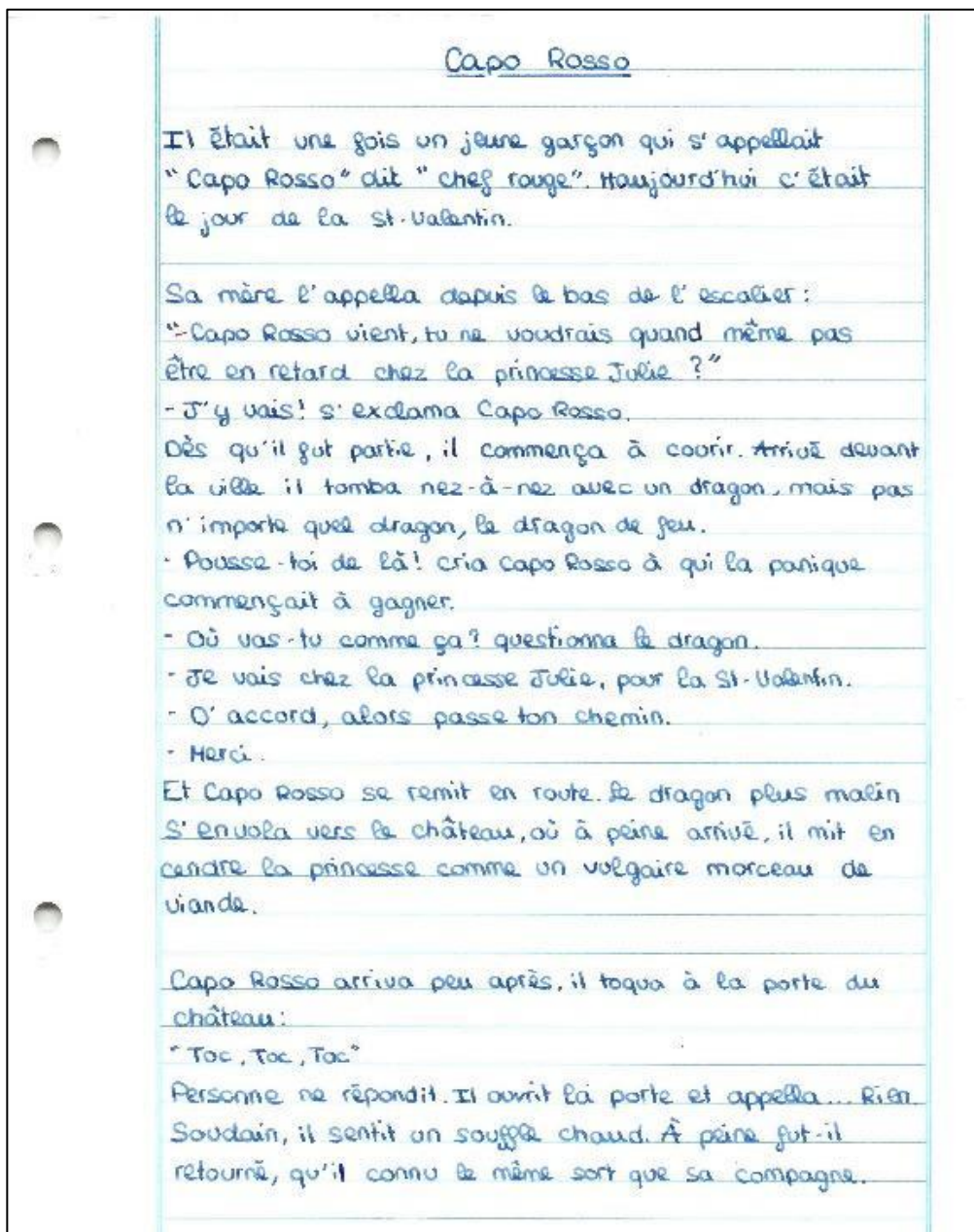


Fig 6 : production écrite d'une élève jugée insuffisante

| | | |
|--------------|--|----------|
| Français 2VP | TS Expression écrite – La parodie de conte | Prénom : |
|--------------|--|----------|

Pour ce TS, tu devras écrire un conte parodique du Petit Chaperon Rouge que tu connais déjà. Tu as le droit à tout moment de consulter les deux versions de ce conte (celle de Perrault et celle des frères Grimm). Ton texte fera minimum 200 mots.

- Choisis une des deux versions comme texte de base pour ta parodie.
- Il te faut ensuite les modifications que tu veux apporter à l'histoire du Petit Chaperon Rouge : les étapes principales du schéma narratif doivent être respectées pour que l'on puisse reconnaître le conte source, mais tout le reste peut être modifié.
- Dans le premier paragraphe, soigne la description de ton personnage principal : quels sont ses caractéristiques ? dans quel univers, quel décor évolue-t-il ?...
- Le système de temps utilisé dans ton texte sera le système du passé : passé simple / imparfait / plus-que-parfait pour marquer l'antériorité. D'autres temps sont acceptés dans les dialogues ! Tu peux utiliser le tableau collé dans ton cahier de structurel'im pour t'aider.
- Structure ton texte à l'aide de paragraphes et de mots organisateurs (marqueurs temporels et connecteurs).
- Si tu insères des dialogues dans ton texte, fais-le selon les règles.
- Utilise un registre courant (ou soutenu) mais pas familier, même dans les dialogues.
- Soigne l'orthographe, la ponctuation, la syntaxe.
- Trouve un nouveau titre en fonction de ton histoire : si aussi, on doit pouvoir faire le lien avec le Petit Chaperon Rouge.
- Commence par faire un brouillon avec tes idées, puis rédige la version finale.

Matériel autorisé :

- Dictionnaire
- Bescherelle conjugaison + Bescherelle Collège
- Les deux versions du Petit Chaperon Rouge
- Le cahier de structuration

3

Ton texte est globalement bien rédigé en ce qui concerne la langue. Cependant il présente quelques problèmes : ton personnage principal n'est pas décrit, tu donnes son nom sans préciser rien d'autre, tu ne fais pas le lien entre son caractère et son surnom. D'autre part, il manque toute une partie où le dragon arrive chez la princesse et se fait passer pour Capot Rosso, puis plus loin le fameux dialogue ("comme tu as de grands yeux...") à modifier selon ton histoire. Ton texte est plus un résumé qu'une parodie, c'est dommage ! Evaluation – TS Expression écrite : parodie de conte

| | Critères collégiaux Tous les critères doivent être notés pour avoir 4 | Critères d'approfondissement Pour chaque critère noté, attribution d'un demi-point supplémentaire |
|------------------|---|---|
| Fond | <input type="checkbox"/> Le schéma narratif du texte original est respecté (ses étapes principales figurent dans la parodie) et l'on peut faire le lien avec le conte source. <input type="checkbox"/> Le personnage principal ainsi que son univers sont décrits. | <input type="checkbox"/> Les liens entre le conte source et la parodie sont subtils et ne se limitent pas à des caractéristiques évidentes. <input type="checkbox"/> Les personnages sont intéressants : leur description est riche et détaillée ; les actions qu'ils accomplissent font avancer le récit. |
| Forme | <input checked="" type="checkbox"/> Le texte est structuré par des paragraphes et par des mots organisateurs (marqueurs temporels et connecteurs). Lorsqu'il y a des dialogues, ils sont insérés correctement. | <input type="checkbox"/> Le texte est dit avec pertinence. Les mots organisateurs utilisés sont variés et toujours adéquats. Les dialogues apportent une dimension intéressante au texte et font avancer le récit. |
| Langue | <input checked="" type="checkbox"/> Le registre de langue est courant ou soutenu. <input checked="" type="checkbox"/> L'orthographe et la syntaxe sont généralement correctes, elles n'empêchent pas la compréhension du texte. | |
| Système de temps | <input checked="" type="checkbox"/> Le système de temps du passé est utilisé correctement : p. simple / imparfait / plus-que-parfait pour marquer l'antériorité, D'autres temps sont acceptés dans les dialogues. | <input type="checkbox"/> Le système de temps du passé est utilisé sans erreur dans l'ensemble du récit. |

Fig 7 : feedback rendu à l'élève par Line.

4.4 Synthèse des atouts et limites de la pratique

Les liens de cohérence entre les différentes composantes de cette épreuve sont les suivants :

- Les critères décrivent les objectifs. Ils sont qualitatifs et servent à l'élaboration des différentes notes, faisant ainsi l'impasse sur l'utilisation des points – relations B et E du modèle de l'alignement curriculaire élargi ;
- La compensation induite par le cumul de points disparaît, laissant place à un jugement professionnel centré sur les apprentissages clés à maîtriser (critères de base) et ceux considérés comme plus difficiles (critères de perfectionnement). Les choix opérés sont donc contextuels, dépendent de l'importance accordée aux différents contenus par Line dans sa connaissance de l'objet disciplinaire, mais aussi dans son enseignement et en référence aux apprentissages effectifs des élèves – relations A, B, C et E ;

- Le fonctionnement de la langue est évalué de manière continue durant la séquence ou au travers de prises d'informations différées et ciblées et non plus de manière consécutive au sein même de l'épreuve. Ainsi, le contenu de cette dernière peut se focaliser sur un seul objet didactique, évitant ainsi une pollution de note et donc une invalidité du résultat – A, B, C, D et E.

Les éléments convergeant vers une cohérence d'alignement curriculaire élargi sont ici nombreux. La reformulation des objectifs pour l'évaluation (référenciation restreinte), les contenus de la tâche de production et ses consignes (design), la formulation des critères (pondération) et l'échelle de notation entretiennent une cohérence du point de vue des contenus et des habiletés cognitives. La validation des critères, après explicitation aux élèves, est compréhensible, avec l'idée du minima à atteindre (critères obligatoires). Tous les seuils des notes sont posés en amont de la correction au travers d'un raisonnement visant à choisir les critères obligatoires à valider, puis ceux de perfectionnement. Le fait que la notation se base uniquement sur des éléments qualitatifs permet à la note d'être constructive, car à haute valeur informative, puisque des feedbacks centrés sur les apprentissages sont possibles. Ainsi, le recours à une telle grille de critères permet de signifier clairement à l'élève les raisons de son échec. Les commentaires centrés sur les contenus et les apprentissages sont alors possibles comme nous l'avons vu, encourageant à juger les performances plutôt que la personne. En d'autres termes, cette évaluation permet à Line de distinguer les textes qui sont des parodies de ceux qui ne le sont pas.

D'autres éléments méritent toutefois d'être discutés. Par exemple, Line exploite des critères qui touchent souvent plusieurs dimensions de l'apprentissage, ce qui pose l'épineuse question des conditions de leur validité. Dès lors, c'est la formulation des critères qui peut questionner : que signifient des liens subtils, des personnages intéressants ? En guise de réponse, Line explique que la grille est distribuée à l'avance et discutée en permanence avec les élèves, que les critères sont explicités de manière à lever les éventuelles ambiguïtés, soulignant ainsi les liens entre évaluation, enseignement et apprentissage. Le guidage à l'aide du tableau de production prend également forme dans des interactions permanentes avec les élèves autour de la compréhension des critères, ce qui participe à augmenter encore la cohérence de l'alignement curriculaire élargi dans cette situation.

5. Discussion

La démarche de Line contraste avec les pratiques évaluatives sommatives traditionnelles de production écrite : le plus souvent, les grilles d'évaluation observées en formation et en recherche ressemblent à celle discutée dans la figure 8 (Pasquini, 2018). Dès lors, de manière à rester proches des pratiques effectives, nous optons pour une discussion par comparaison qui permet de dégager les spécificités de la situation de Line et étayer la problématique de la certification de l'écrit. Pour cela, nous présentons succinctement une épreuve courante et en décrivons les biais.

Cette épreuve, également destinée à des élèves de 9^e année, mais de voie Générale, demande de produire deux textes liés, mais distincts : le premier texte est la retranscription d'un message de réclamation laissé sur un répondeur, utilisant le registre familier, tandis que le second est la lettre de réponse à ce message, rédigée en langage soutenu. L'enseignant a construit deux grilles, une par texte produit, qui participent de la même évaluation. Le total des points est obtenu par l'addition des points de chacune des grilles.

Date : _____ Prénom : _____

—
d'énonciation...

...en organisant ses idées, en personnalisant son message et en précisant sa pensée

...en analysant la situation

...en respectant les règles de l'expression écrite (temporalité, orthographe, syntaxe, reprises,...) et en utilisant un vocabulaire adapté à la situation

...en enrichissant le contenu du texte (idée, vocabulaire, grammaire, ...)

...en

...en améliorant sa production (relecture, réécriture, réorganisation, enrichissement,...)

...en utilisant la dimension créative de la langue et sa propre créativité

Objectifs d'apprentissage

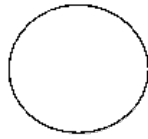
L'élève sera capable

d'écrire un texte, une lettre, un dialogue comportant toutes les caractéristiques des

de relire et corriger l'orthographe

d'utiliser une ponctuation adéquate

d'améliorer son texte en évitant les répétitions et en enrichissant son vocabulaire



/33pts

A. Tu es allé manger dans un restaurant dans lequel tout ne s'est pas passé comme tu le souhaitais. Parmi les problèmes proposés, choisis-en 1 et imagine que tu laisses un message de réclamation sur le répondeur du restaurant. Tu te présentes, expliques le problème, donnes au restaurateur une chance de se racheter et le salues. Ecris ce message en langage familier en respectant les consignes.

- *Personnel pas sympathique*
- *Mobilier pas confortable*
- *Nourriture mauvaise*

| | | |
|--|-----------------------------------|--|
| 1. Ton texte comporte au moins 80 mots (mets une X après le 80 ^{ème} mot) | < 70 mots : 0pt | |
| | 70-80 mots : 0.5pt | |
| | >80 mots : 1pt | |
| 2. Il contient ta présentation, la description du problème rencontré et des salutations à la fin | Présentation et salutations : 1pt | |
| | Description : 1pt | |
| 3. Tu fais une demande de réparation sous forme de question | Si juste : 1pt | |
| 4. Il y a au moins deux phrases à la forme négative | Si 1 phrase : 0.5pt | |
| | Si 2 phrases : 1pt | |
| 5. Syntaxe adaptée au registre de langue | 0fte : 2pts | |
| | -1/2 pt par faute | |
| | 0-3 mots : 0pt | |
| 6. Choix du vocabulaire adapté au registre de langue | 4-6 mots : 1pt | |
| | 7-9 mots : 1.5pt | |
| | >10 mots : 2pts | |
| 7. Touche d'humour | Si bien : 1pt | |

FRA 9VG Expression écrite Date : _____ Prénom : _____

B. Ecris ensuite l'e-mail de réponse du restaurateur en langage soutenu. Commence la lettre de manière officielle, indique que tu as bien reçu le message, formule des excuses et propose un moyen de réparer les torts. Termine par les salutations d'usage.

| | | |
|--|---|--|
| 1. Ton texte comporte au moins 80 mots (mets une X après le 80 ^{ème} mot) | < 70 mots : 0pt 70-80 mots : 0.5pt >80 mots : 1pt | |
| 2. Il contient des excuses | Si juste : 1pt | |
| 3. Il contient une solution de réparation | Si juste : 1pt | |
| 4. Il commence de manière officielle et se termine par des salutations d'usage | Si juste : 1pt | |
| 5. Syntaxe adaptée au registre de langue | Offte : 2pts -1/2 pt par faute | |
| 6. Choix du vocabulaire adapté au registre de langue | 0-3 mots : 0pt 4-6 mots : 1pt 7-9 mots : 1.5pt >10 mots : 2pts | |
| A. Accords | Offte : 5pts -1/2 pt par faute | |
| O. Orthographe lexicale | Offte : 5pts | |
| P. Ponctuation | -1/2 pt par faute Offte : 3pts | |
| R. Répétitions | -1/2 pt par faute Pas plus de 2fois le même GN : 2pts | |
| Présentation | -1/2 pt par faute Si bonne présentation, pas de ratures : 1pt | |

Figure 8 : Évaluation traditionnelle d'une production écrite (9^è année, Voie Générale)
 Nous observons dans cet exemple de nombreux désalignements curriculaires.

- Les deux grilles, à différents égards, ne ciblent pas d'éléments spécifiques au genre. La première grille concerne un texte que l'énonciateur aurait laissé sur un répondeur. Pourtant, le texte en question n'est pas évalué dans le cadre d'une production orale, mais bien celui d'une production écrite. Or, la recherche en didactique montre que l'oral est régi par des spécificités syntaxiques qui lui sont propres et qui ne sont justement pas la transposition à l'oral des règles écrites (Béguelin, 1996 ; Blanche-Benveniste, 2010). Il s'agit là d'un désalignement curriculaire important, source de brouillage pour les élèves. La seconde grille, si elle pointe des caractéristiques du genre (Dolz & Schneuwly, 1998) – deux éléments de contenu (le texte contient des excuses / le texte contient une solution de réparation), ainsi qu'un élément de planification (il commence de manière officielle et se termine par des salutations d'usage) – accorde un poids mineur (3 points sur 21) à ces éléments pourtant centraux du point de vue didactique.
- Il est difficile de percevoir la cohérence entre les objectifs énoncés issus du curriculum et déclinés par l'enseignant, les consignes et les critères (par exemple, où les élèves analysent-ils une situation d'énonciation ?) – relation A.
- De nombreux critères ne décrivent pas les objectifs en termes de contenus ni d'habiletés cognitives, mais reposent sur le respect de consignes (le texte comporte au moins 80 mots) ou de procédures (le texte contient des excuses) – relation B.
- L'attribution des points ne rend pas compte d'une pondération relative à la complexité des apprentissages présents dans les critères. Ainsi, par exemple, la touche d'humour (!) vaut autant que la description du problème – relation B.
- Le fait de recourir à des points engendre un phénomène de compensation qui nuit à la qualité et à la cohérence de l'épreuve (Balan & Jönsson, 2018). Ici, peu importe où l'élève capitalise de la valeur, pourvu qu'il y en ait assez. Avec un tel système, non seulement les critères disparaissent derrière les points, mais un élève peut réussir avec des lacunes importantes ou pire, échouer alors qu'il maîtrise des éléments clés de production écrite – relation E. Ce phénomène est d'autant plus problématique que 17 points sur 33 sont attribués au fonctionnement de la langue.
- Les critères relatifs au fonctionnement de la langue comptant pour moitié dans le total de points, la note reflète potentiellement autant d'acquisitions sur ces objets qu'en production écrite – relations B et E.
- Même sans disposer de l'échelle de notation, nous pouvons observer qu'avec de tels écarts dans la pondération des différents "critères", la note ne pourra pas rendre compte de la maîtrise avérée d'apprentissages significatifs en matière de production écrite.
- De manière plus générale, ce type de grilles interroge la conception du texte qui est sous-jacente. En effet, des grilles qui séquencent la production écrite en de micro-unités tendent à faire penser qu'un texte n'est qu'une addition de parties. Or, Halliday et Hasan (1976), comme nous l'avons vu, suggèrent qu'une cohérence globale doit également traverser un texte pour qu'on puisse nommer ce dernier ainsi.

Tous ces biais ont été abondamment documenté par la recherche en évaluation (par ex., Biggs, 2003 ; Guskey, 2013 ; Jönsson & Balan, 2018 ; Pasquini, 2019b). Ils n'en restent pas moins très présents dans les pratiques.

6. Conclusion

Au travers de cette étude de cas, nous avons tenté de montrer que toute démarche d'évaluation sommative de la production écrite relève d'une pratique complexe, délicate à enfermer dans des procédures statiques ou des manières de faire standardisées. En effet, l'exemple de Line souligne qu'un travail préalable sur l'objet didactique, les apprentissages à mener et l'enseignement permettant leur appropriation, idéalement en équipe de file et en référence au contexte, est indispensable (Pasquini, 2019a). En cela, nous avons mis en évidence le caractère dynamique de la formulation des critères et de la construction de la note. Orienter ses pratiques dans une telle perspective nécessite toutefois de prendre quelques précautions avec les élèves, habitués à des formes de certifications traditionnelles (Mottier Lopez, 2013 ; Ohlsen, 2007) où les points, les totaux de points et les seuils de notation standardisés sont rois. Ces pratiques, dans une certaine mesure, sont rassurantes. Mais elles contribuent aussi à éloigner les élèves d'une logique d'apprentissage, leur souci majeur étant de réussir. Pour cela, ils développent des compétences de calculateurs, d'expertise

en marchandages de points et de négociation diverses, sans comprendre pourquoi ils réussissent ou échouent (Feldman, 2018).

A l'instar de Line, engager des démarches sommatives qualitatives demande donc d'explicitier les pratiques, d'en montrer les spécificités et de souligner leur valeur en termes de soutien à l'apprentissage. C'est notamment pour cela qu'une communication continue et précoce avec les élèves sur les critères est encouragée par la recherche et la pratique, de manière à clarifier le contrat d'évaluation (Jönsson & Balan, 2018).

Du point de vue de la formation et de la recherche, le recours à des modèles théoriques mettant au centre de leur conceptualisation les caractéristiques des objets disciplinaires et les liens de cohérence entre les diverses composantes de l'évaluation sommative devient incontournable. Ainsi en est-il de l'alignement curriculaire élargi, dont les enseignants soulignent, aussi, la pertinence. Ce modèle permet également de balayer une réserve qui pourrait être formulée par certains enseignants, craignant de « trop » guider les élèves et donc de niveler leurs compétences par le bas. Nous pensons avoir montré qu'au contraire, évaluer en se concentrant sur les objets disciplinaires permet d'attirer l'attention de l'enseignant et de l'élève sur les caractéristiques centrales d'un genre particulier, sans se laisser distraire par des critères structurels tels que le fonctionnement de la langue, importants certes, mais non spécifiques car présents dans tous les genres de textes. Au final, on peut postuler que la qualité des productions écrites augmente, car un enseignement ciblé sur les apprentissages-clés du genre a été rendu possible (Pasquini, 2018).

Notons toutefois que penser ses pratiques à l'aune d'un tel modèle est exigeant. Cela suppose tout d'abord de solides connaissances des objets disciplinaires. C'est en effet à cette seule condition qu'une appropriation du curriculum est possible, permettant ainsi de l'exploiter de manière critique pour enseigner, faire apprendre et enfin certifier (Marcel, 2011). Une attention particulière doit également être portée sur le design des tâches, et plus spécifiquement sur les consignes. À ce niveau, il est primordial que ces dernières dirigent les élèves, qu'elles soient davantage orientées sur l'« apprendre » (rédige un titre en rapport avec le thème de l'histoire) que sur le « faire » (mets un titre). Mais peut-être surtout, l'enseignant doit posséder des connaissances en évaluation sommative, et spécifiquement celles liées aux critères, aux échelles et aux barèmes (Moss, 2013). Ce n'est qu'à ce prix qu'il lui sera possible de se distancer d'idées reçues, de façons de faire ou de directives, d'assumer un jugement professionnel centré sur les apprentissages (Mottier Lopez & Tessaro, 2016). Ainsi, il sera en mesure de faire des choix au service d'une certification cohérente et porteuse de sens.

Bibliographie

- Adam, J.-M. (2008). *La linguistique textuelle : introduction à l'analyse textuelle des discours*. Armand Colin.
- Airasian, W., & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory into Practice*, 41(4), 249-254.
- Anderson, L. W. (2002). Curricular alignment : a re-examination. *Theory into Practice*, 41(4), 255-260.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing : a revision of Bloom's taxonomy of educational objectives*. Longman.
- Andrade, H. G. (2005). Teaching with rubrics : the good, the bad, and the ugly. *College Teaching*, 53(1), 27-31.
- Balan, A., & Jönsson, A. (2018). Increased explicitness of assessment criteria : effects on student motivation and performance. *Frontiers in Education*, 3(81), 56-68.
- Balsiger, C., Béatrix, D., Cordonier, N., Panchout-Dubois, M., & Ticon, J. (2011). Le Plan d'études romand de 2010 : un curriculum d'un nouveau type et les effets de son introduction. In J.-L. Dumortier, J. Van Beveren & D. Vrydaghs (Eds.), *Curriculum et progression en français* (pp. 19-50). Presses Universitaires de Namur.
- Béguelin, M.-J. (1996). Le rapport écrit-oral : tendances dissimultrices, tendances assimilatrices. *Cahier de linguistique française*, 20, 229-253.
- Biggs, J. (1999). What the student does : teaching for enhanced learning. *Higher Education Research and Development*, 18(1), 57-75.
- Biggs, J. (2003, 21 mars). Aligning teaching for constructive learning. *Advance HE*. <https://www.advance-he.ac.uk/knowledge-hub/aligning-teaching-constructing-learning>
- Blanche-Benveniste, C. (2010). *Le français : usages de la langue parlée*. Peeters.
- Braxmeyer, N., Guillaume, J.-C., & Levy, J.-F. (2005). Les pratiques d'évaluation des enseignants au collège. Les dossiers du Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Direction de l'évaluation et de la prospective, 160, 5-117.
- Brimi, H. M. (2011). Reliability of grading high school work in english. *Practical Assessment, Research and Evaluation*, 16(17). <https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1287&context=pape>
- Brookhart, S. M. (2005). Assessment theory for college classrooms. *New Directions for Teaching and Learning*, 100, 5-14.
- Brookhart, S. M. (2017). How to use grading to improve learning. Association for Supervision and Curriculum Development (ASCD).
- Carulla, C., Corti, D., de Pietro, J.-F., Kassam, S., Roth, M., SÁnchez Abchi, V., & Singh, L. (2013). Quoi et comment évaluer en référence au Plan d'études romand (PER) : premières pistes. (Rapport scientifique intermédiaire). Institut de recherche et de documentation pédagogique (IRDp).
- Connan-Pintado, C. (2009). Lire des contes détournés à l'école à partir des Contes de Perrault, de la GS au CM2. Hatier pédagogie.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading : endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72.
- De Ketele, J.-M. (2012). A la recherche de paradigmes unificateurs ou de paradigmes intégrateurs? In L. Mottier Lopez & G. Figari (Eds.), *Modélisations de l'évaluation en éducation* (pp. 195-210). De Boeck.
- De Lestrade, A. & Bravo, C. (2005). Le plus gentil loup du monde. La joie de lire.
- Dolz, J., & Schneuwly, B. (1998). *Pour un enseignement de l'oral : initiation aux genres formels à l'école*. ESF éditeur.
- Duncan, R. C., & Noonan, B. (2007). Factors affecting teachers'grading and assessment practices. *The Alberta Journal of Educational Research*, 53(1), 1-21.
- Feldman, J. (2018). The end of points. *Educational Leadership*, 75(5), 36-40.
- Figari, G. (2012). Configurations pour une épistémologie de l'évaluation. In L. Mottier Lopez & G. Figari (Eds.), *Modélisations de l'évaluation en éducation* (pp. 211-224). De Boeck.
- Figari, G. (2013). L'évaluation entre « technicité » et « théorisation ». *Mesure et évaluation en éducation*, 36(3), 5-24.
- Florey, S., & Depallens, V. (2020). L'accès à la littérature, un nouveau défi pour l'enseignement au primaire. In A.-C. Blanc & V. Capt (Eds.), *La tête et le texte : formation initiale des enseignants primaires en didactique de la lecture et de l'écriture* (pp. 65-81). Peter Lang.
- Gagnon, R., Monnier, A., & Dolz, J. (2014). L'évaluation dans la formation à l'enseignement de la production écrite : l'exemple suisse romand. *Pratiques*, 161/162. <https://doi.org/10.4000/pratiques.2120>
- Gauthier, C., Mellouki, M., Bissonnette, S., & Richard, M. (2005). Écoles efficaces et réussite scolaire des élèves à risque : un état de la recherche. (Rapport de recherche du Centre de recherche interuniversitaire sur la formation et la profession enseignante). CRIFPE, Université Laval.

- Goasdoué, R., Vantourout, M., & Bedoin, D. (2016). La correction de dissertations : nouveau regard sur la construction des jugements des évaluateurs. In L. Mottier Lopez & W. Tessaro (Eds.), *Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages* (pp. 73-93). Peter Lang.
- Guskey, T. R. (2013, septembre). The case against percentage grades. Educational, School, and Counseling Psychology Faculty Publications, 22. Uknowledge. https://uknowledge.uky.edu/edp_facpub/22
- Guskey, T. R., & Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education : Principles, Policy & Practice*, 26(3), 303-320.
- Hadjj, C. (2016). Le système éducatif peut-il passer de l'évaluation normative à une évaluation constructive? In B. Bocquet (Ed.), *La fièvre de l'évaluation quels symptômes, quels traitements?* (pp. 113-135). Presses Universitaires du Septentrion.
- Hadjj, C. (2017). Savoir mettre en œuvre une évaluation constructive. In A. Bentolila (Ed.), *L'essentiel de la pédagogie* (pp. 229-250). Nathan.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Harlen, W. (2005). Teacher's summative practices and assessment for learning : tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Hillocks, G. (2006). Two decades of research on teaching writing in the secondary schools in the US. *L1-Educational Studies in Language and Literature*, 6(2), 29-51.
- Jönsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840-852.
- Jönsson, A., & Balan, A. (2018). Analytic or holistic : a study of agreement between different grading models. *Practical Assessment, Research and Evaluation*, 23(12). <https://doi.org/10.7275/z3gm-fp34>
- Marcel, J.-F. (2011). Des interrelations entre les processus curriculaires et les processus évaluatifs. *Les Dossiers des sciences de l'éducation*, 25, 7-12.
- Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, 79(4), 1332-1361.
- Marzano, R. J. (2002). A comparison of selected methods of scoring classroom assessments. *Applied Measurement in Education*, 15(3), 249-267.
- Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Effective Grading Practices*, 69(3), 34-39.
- McGrath H. et Noble, T. (2008). *Huit façons d'enseigner, d'apprendre et d'évaluer*. Chenelière Education.
- McMillan, J. H., & Nash, S. (2000). Teacher classroom assessment and grading practices decision making. (Report ERIC Document Reproduction Service No. ED447195). ERIC. <https://files.eric.ed.gov/fulltext/ED447195.pdf>
- Millon-Fauré, K. (2013). Processus de négociation didactique et mesure du niveau des élèves : des fonctions concurrentes de l'évaluation. *Carrefours de l'éducation*, 36(2), 149-166.
- Moss, C. (2013). Research on classroom summative assessment. In J. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 235-256). SAGE Publications.
- Mottier Lopez, L. (2013, 20 juin). Pour une culture partagée de l'évaluation : du retour à des pratiques d'évaluation traditionnelles vers la construction d'un nouveau rapport entre la recherche, le terrain et le politique. *Revue Éducation Canada*. Le Réseau EdCan. <https://www.edcan.ca/articles/pour-une-culture-partagee-de-levaluation/?lang=fr>
- Mottier Lopez, L., & Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5-34.
- Mottier Lopez, L., Serry, S., & Sales Cordeiro, G. (2016). Évaluation certificative de productions écrites d'élèves : référence au modèle didactique du genre textuel dans l'épreuve et lors de la correction de faits divers. In É. Falardeau, J. Dolz, J.-L. Dumortier & P. Lefrançois (Eds.), *L'évaluation en classe de français, outil didactique et politique* (pp. 27-61). Presses Universitaires de Namur.
- Mottier Lopez, L., & Tessaro, W. (Eds.). (2016). *Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages*. Peter Lang.
- Ohlsen, M. T. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education*, 36 (1), 4-14.
- Pasquini, R. (2017). Évaluation : quatrième question vive de la pratique. *Éducateur*, 7, 34-35.
- Pasquini, R. (2018). Le modèle théorique de l'alignement curriculaire élargi pour étudier des pratiques évaluatives sommatives d'enseignants de mathématiques et de français du secondaire : enjeux conceptuels et pragmatiques. (Thèse de doctorat inédite). Université de Genève. <https://archive-ouverte.unige.ch/unige:106442>
- Pasquini, R. (2019a). Construire des évaluations sommatives communes : une mauvaise « bonne idée »? *Éducateur*, 6, 4-5.

- Pasquini, R. (2019b). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63-92.
- Petitjean, A. (1984). Pastiche et parodie : enjeux théoriques et pédagogiques. *Pratiques : linguistique, littérature, didactique*, 42, 3-33.
- Pope, N., Green, S. K., Johnson, R. L., & Michelle, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25, 778-782.
- Prøitz, T. S. (2013). Variations in grading practice : subjects matter. *Education Inquiry*, 4(3), 555-575.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380.
- Resh, N. (2009). Justice in grades allocation : teachers' perspective. *Social Psychology of Education : An International Journal*, 12(3), 315-325.
- Reuter, (Y). (1984). Contes et parodies : un exemple d'introduction de l'histoire littéraire dans le premier cycle. *Pratiques : linguistique, littérature, didactique*, 42, 79-88.
- Sangsue, D. (2007). La relation parodique. José Corti.
- Simard, C., Dufays, J.-L., Dolz, J., & Garcia-Debanc, C. (2010). L'évaluation du travail des élèves. In C. Simard, J.-L. Dufays, J. Dolz & C. Garcia-Debanc (Dir.), *Didactique du français langue première* (pp. 139-164). De Boeck.
- Tobola Couchepin, C., SÁnchez Abchi, V., & Dolz, J. (2016). Pratiques d'évaluation de la production écrite à l'école primaire : outils et régulations. In É. Falardeau, J. Dolz, J.-L. Dumortier & P. Lefrançois (Eds.), *L'évaluation en classe de français, outil didactique et politique* (pp. 63-84). Presses Universitaires de Namur.
- Vial, M. (2012). Se repérer dans les modèles de l'évaluation. De Boeck.
- Walvoord, B. E., & Johnson Anderson, V. (2009). *Effective grading : a tool for learning and assessment*. Jossey Bass.
- Wigging, G., & Mc Tighe, J. M. (2005). *Understanding by design*. Association for Supervision and Curriculum Development (ASCD).
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self perceived assessment skills. *Applied Measurement in Education*, 16(4), 323-342.

Auteur·e

Raphaël Pasquini, Professeur HEP associé en évaluation certificative des apprentissages, en évaluation formative et dans les processus d'orientation scolaire, Haute École pédagogique du canton de Vaud.

Sonya Florey, Professeur HEP ordinaire en didactique du français (didactique de la littérature de l'extrême-contemporain, littérature de jeunesse, analyse des traces du néolibéralisme dans le contexte scolaire), Haute École pédagogique du canton de Vaud.

Cet article a été publié dans le numéro 3/2020 de forumlecture.ch

Schriftliche Arbeiten auf der Sekundarstufe benoten und ... nichts Anderes!

Raphaël Pasquini und Sonya Florey

Abstract

Seit der Einführung des Lehrplans für die französischsprachige Schweiz (plan d'études romand, PER) 2011 weisen viele Lehrpersonen darauf hin, dass es allgemein schwierig ist, schriftliche Texte zu beurteilen. Eine der Herausforderungen besteht für sie darin, die Beurteilung der sprachlichen Qualität mit den spezifischen Lerninhalten der schriftlichen Arbeit in Bezug zu setzen und Beurteilungskriterien festzulegen, die sich auf den eigentlichen Lerninhalt beziehen. In diesem Beitrag wird eine Fallstudie vorgestellt, die mit Hilfe eines theoretischen Modells ein umfassendes, benotendes Beurteilungsverfahren auf der Sekundarstufe im Kanton Waadt analysiert. Die Analyse ist Teil einer professionellen Fragestellung, die zu neuen Überlegungen führen kann, damit die Lehrpersonen die schriftlichen Arbeiten benoten, und nichts Anderes.

Schlüsselwörter

Schriftliche Arbeit, allgemeine Beurteilung, erweiterte Lehrplanausrichtung, Märchenparodie.

Dieser Beitrag wurde in der Nummer 3/2020 von leseforum.ch veröffentlicht.

Certificare la scrittura nella scuola secondaria e... nient'altro!

Raphaël Pasquini e Sonya Florey

Riassunto

Dall'introduzione del plan d'études romand (PER), nel 2011, molti insegnanti hanno riscontrato difficoltà nel certificare la produzione scritta di tutti i generi. Una delle sfide, per loro, consiste nel collegare la valutazione del funzionamento della lingua all'apprendimento specifico della produzione scritta e nel formulare criteri di valutazione incentrati sul contenuto specifico dell'oggetto didattico. Per affrontare questa problematica, il nostro articolo presenta un caso di studio che utilizza un modello teorico per analizzare un approccio di valutazione sommativa osservato nella scuola secondaria nel Canton Vaud. Questa analisi si inserisce in un'interrogazione professionale che porterà a sua volta a nuove riflessioni, in modo che gli insegnanti possano certificare la produzione scritta, e nient'altro.

Parole chiave

Produzione scritta, valutazione sommativa, allineamento esteso del curriculum, parodia della narrazione.

Questo articolo è stato pubblicato nel numero 3/2020 di forumlettura.ch