



Cristiana Cervini et Marie-Pierre Jouannaud

Ouvertures et tensions liées à la conception d'un système d'évaluation en langues, numérique, multilingue et en ligne, dans une perspective communicative et actionnelle

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Cristiana Cervini et Marie-Pierre Jouannaud, « Ouvertures et tensions liées à la conception d'un système d'évaluation en langues, numérique, multilingue et en ligne, dans une perspective communicative et actionnelle », *Alsic* [En ligne], vol. 18, n° 2 | 2015, mis en ligne le 30 septembre 2015, Consulté le 15 janvier 2016. URL : <http://alsic.revues.org/2821> ; DOI : 10.4000/alsic.2821

Éditeur : Adalsic

<http://alsic.revues.org>

<http://www.revues.org>

Document accessible en ligne sur :

<http://alsic.revues.org/2821>

Document généré automatiquement le 15 janvier 2016.

CC-by-nc-nd

Cristiana Cervini et Marie-Pierre Jouannaud

Ouvertures et tensions liées à la conception d'un système d'évaluation en langues, numérique, multilingue et en ligne, dans une perspective communicative et actionnelle

1. Contextualisation

- 1 Les institutions d'enseignement supérieur font face à la contrainte d'évaluer – et de former – un grand nombre d'étudiants en langues étrangères (Carrère *et al.*, 2009), ce qu'on appelle dans la littérature anglo-saxonne "*large-scale assessment*" (Kunnan & Jang, 2009 : 611). En particulier, chaque année la passation d'un test en langues devient une étape incontournable pour évaluer leurs compétences et les orienter vers des formations ciblées sur leurs besoins. Pour répondre à cette exigence, des questions didactiques, scientifiques et organisationnelles se posent au sein des institutions universitaires.
- 2 Dans cette contribution, nous nous interrogeons sur la cohérence possible entre le *testing* autocorrectif – en tant qu'outil opérationnel de l'évaluation au sens large – et les paradigmes communicatifs qui lui servent de socle. La culture de l'institution, ainsi que la présence ou l'absence de possibilités concrètes d'investissement, peuvent déterminer le choix de créer un nouvel outil d'évaluation, ou bien d'en adopter un déjà existant.
- 3 Cette réflexion sera menée et argumentée parallèlement à la description d'un "Système d'évaluation en langues à visée formative (Self)" multilingue (Cervini *et al.*, 2013 ; Cervini, 2104), conçu au service Lansad de l'université Stendhal Grenoble 3, dans le cadre du projet Idefi (initiatives d'excellence en formations innovantes) *Innovalangues* (Masperi, 2012) financé par l'Agence nationale de la recherche. Self, administré en ligne, s'inscrit dans une vaste gamme de modèles de formation qui sont proposés au niveau institutionnel (*présentiel enrichi, en formation hybride, en autonomie guidée, en autonomie totale*), et répond à deux exigences prioritaires : d'un côté, positionner les étudiants dans des groupes-classes de langues selon les niveaux du *CECRL*, de l'autre, servir de soutien et de guide aux étudiants / candidats et à leurs tuteurs, par des retours de type diagnostique et formatif. Cette identité double constitue une source de richesse et d'innovation aussi bien que de complexité conceptuelle et opérationnelle.
- 4 Pour répondre aux buts de positionnement, des contraintes organisationnelles et procédurales s'imposent : le temps de passation doit être réduit (50 minutes dans notre cas) et des procédures de validation psychométrique sont nécessaires pour que le test puisse être jugé utile, valide et fiable (Doucet, 2001) ; en même temps, l'identité diagnostique-formatif du système (Lengyel, 2011 ; Kunnan & Jang, 2009 ; Alderson, 2006) invite à mettre en place des méthodes d'observation mixtes. À côté d'une approche quantitative dont la validation passe par des analyses psychométriques de données récoltées sur des échantillons représentatifs (application de la théorie classique des tests et du modèle de Rasch), nous nous rapprochons des paradigmes sociocognitifs, par la prise en compte de facteurs biographiques, motivationnels et attitudeux des apprenants candidats au test. La question est alors de savoir comment on peut mettre en relation tous ces facteurs.
- 5 Le but principal du diagnostic consiste à aider les étudiants à repérer leurs points forts et leurs lacunes en langues, à collecter des indices "in-formatifs", et à faciliter le démarrage d'un parcours d'apprentissage et / ou de remédiation qui se veut autonomisant.
- 6 La définition du construit des compétences dans Self (compréhension de l'oral, de l'écrit et expression écrite courte) se propose de représenter la complexité du langage à travers les approches didactiques les plus récentes. Cependant, la traduction de ce construit lors de la conception d'une banque de tâches numériques multilingues autocorrectives détermine les

tensions et les ouvertures qui sont anticipées dans le titre du papier. Nous nous proposons ci-dessous d'illustrer la manière dont nous avons interprété la perspective communicative-actionnelle dans la conception de ce système d'évaluation autocorrectif.

2. Compétence communicative et *testing*

2.1. Perspectives historiques

- 7 La capacité que nous voulons tester, ou "construit" de notre test, est la compétence communicative, c'est-à-dire une capacité à utiliser la langue en contexte, pour agir avec la langue. Cette conception peut être reliée à plusieurs approches qui ont des origines et des positionnements théoriques différents mais se rejoignent sur l'importance de la contextualisation de la langue.
- 8 En linguistique appliquée, l'approche communicative est développée depuis les années 1960 (Hymes, 1972 ; Widdowson, 1978 ; Canale, 1983 ; Savignon, 1983) en réaction aux approches structurales d'enseignement des langues étrangères ("*audio-lingual method*" dans le contexte anglo-saxon), qui avaient tendance à laisser systématiquement de côté le sens, en particulier pragmatique. Cette focalisation sur la compétence de communication a ensuite été transposée dans le contexte de l'évaluation ("*testing*"), d'abord par Bachman (1990), puis par d'autres chercheurs comme Savignon (1990) ou Celce-Murcia *et al.* (1995), qui ont tenté de clarifier et d'explicitier le construit de compétence communicative, afin de permettre une opérationnalisation mieux étayée et plus complète lors de la conception de tests de langue. Celce-Murcia *et al.* (1995) distinguent, par exemple, plusieurs sous-compétences dans leur modèle :
- compétence discursive (qui correspond à la compétence "textuelle" chez Bachman) ;
 - compétence actionnelle ("fonctionnelle" chez Bachman) ;
 - compétence linguistique ("grammaticale") ;
 - compétence socioculturelle (même terme chez Bachman) ;
 - compétence stratégique (même terme chez Bachman).
- 9 Parallèlement à ces développements anglo-saxons en linguistique appliquée, on peut citer les approches énonciatives héritées de Benveniste développées en France à partir des années 1970 (par exemple la théorie des opérations énonciatives), qui s'intéressent, non aux phrases grammaticalement bien formées (comme en linguistique structurale ou ensuite générative), mais aux énoncés qui ne prennent leur sens que dans une situation d'énonciation particulière, c'est-à-dire les énoncés pris en charge par un énonciateur à un moment donné et dans un lieu donné, et adressés à un co-énonciateur (Culioli, 1978 ; Bouscaren & Chuquet, 1987 ; Guillemin-Flescher, 1981). Il est donc important d'aider les étudiants placés dans la situation artificielle du test à recréer les situations d'énonciation correspondant aux énoncés qui leur sont proposés. Nous expliquerons plus bas comment nous avons opérationnalisé le contexte de chaque tâche d'évaluation¹.
- 10 Toujours en linguistique théorique, mais dans une tradition plus anglo-saxonne, on peut citer l'approche systémique fonctionnelle de Halliday (1985), qui rejoint les théories énonciatives sur l'importance des choix linguistiques opérés par les locuteurs, et qui insiste sur l'importance de la prise en compte des genres textuels dans la fabrication du sens des textes ("*meaning making*"). Nous prévoyons également d'essayer d'aider les candidats à repérer le genre des textes par des indices visuels tels que des indices typographiques ou de mise en page, qui sont souvent utilisés dans l'enseignement de la compréhension écrite en langue étrangère pour aider les apprenants à activer leurs schémas textuels (Cicurel, 1991).
- 11 Enfin, depuis quelques années se développe en Europe une perspective "actionnelle"², exemplifiée par la publication du *CECRL – Cadre européen commun de référence en langues* (Conseil de l'Europe, 2001). Cette perspective, qui peut être considérée comme une extension de l'approche communicative (certains chercheurs comme Bourguignon, 2006, utilisent le terme de "communic'actionnelle" pour expliciter cette parenté, même si elle est parfois considérée comme "une approche complètement nouvelle", Éduscol, nd), suppose une inscription des apprenants dans un contexte social plus vaste que la situation d'énonciation

décrite plus haut. Nous expliquerons plus bas en quoi il est difficile de considérer qu'un test de positionnement / diagnostic puisse s'ancrer aisément dans cette approche.

2.2. Applications au *testing* autocorrectif

- 12 Le *CECRL*, sur lequel nous nous fondons pour guider la conception de notre test, est l'héritier plus ou moins lointain des différentes traditions décrites dans le paragraphe précédent (Hilton, 2009). Nous utilisons le *CECRL* pour définir un apprenant prototypique comme un locuteur qui doit faire face à des situations de communication variées, et une tâche comme "toute visée actionnelle que l'acteur se représente comme devant parvenir à un résultat donné en fonction d'un problème à résoudre, d'une obligation à remplir, d'un but qu'on s'est fixé" (Conseil de l'Europe, 2001).
- 13 Le modèle de communication en langues promu dans le *CECRL*, à la suite de modèles comme ceux de Bachman et Palmer ou Celce-Murcia, mais sous une forme plus ramassée, identifie plusieurs composantes de la "compétence à communiquer langagièrement" (Conseil de l'Europe, 2001 : 17-18) :
1. compétence linguistique ;
 2. compétence sociolinguistique ;
 3. compétence pragmatique.
- 14 La compétence pragmatique se sépare en deux sous-compétences :
1. la compétence discursive, que nous prenons en compte avec des items d'expression écrite courte testant la cohérence, la cohésion ou l'organisation logique ;
 2. la compétence fonctionnelle, qui correspond à des items où les étudiants doivent identifier la fonction illocutoire d'un énoncé (Austin, 1962), par exemple reconnaître que telle question n'est pas une demande d'information mais un ordre déguisé.
- 15 La compétence sociolinguistique, quant à elle, est composée de plusieurs sous-parties :
- les marqueurs de relation sociale, les règles de politesse et les différences de registre sont représentés dans Self (par exemple par le choix d'une salutation appropriée dans un contexte donné) ;
 - les "expressions de la sagesse populaire" (Conseil de l'Europe, 2001 : 93) ne sont pas testées car elles paraissent trop culturellement spécifiques et peu adaptées à un public d'apprenants de langue étrangère ;
 - les "dialectes et accents" ; même si une variété d'accents est présentée, les étudiants ne sont pas censés reconnaître des formes très marquées socialement, régionalement, nationalement ou professionnellement (l'exemple illustrant cette sous-compétence dans le *CECRL* est le verbe "magasiner" pour "faire les courses" en québécois).
- 16 Dans le domaine du *testing*, ce travail de traduction de conceptions abstraites de la compétence langagière en tâches d'évaluation s'appelle "opérationnalisation du construit". Dans la conception de Self, nous avons suivi les pratiques habituelles des tests en excluant du construit les compétences langagières qui sont liées explicitement et directement aux connaissances encyclopédiques des candidats, et les savoirs socioculturels qui en font partie (*CECRL*, Conseil de l'Europe, 2001 : 82). En effet, il est difficile alors de savoir ce qui, des compétences linguistiques ou des connaissances culturelles / encyclopédiques, a aidé ou au contraire empêché les candidats de répondre³.
- 17 Enfin, le Cadre promeut une approche actionnelle de l'enseignement / apprentissage des langues qu'il paraît difficile de transposer dans un test tel que le nôtre. En effet, un positionnement suppose la mise en relation d'un candidat et d'un échantillonnage de situations de communication représentatives. Il n'est pas possible de positionner un candidat sur la réalisation d'une seule tâche, comme le rappelle Alderson.

Language proficiency testers these days (...) are almost obliged to examine how somebody would perform in a range of communicative settings, across a variety of tasks, tests, activities and all

the other dimensions of communicative behavior that are discussed, however briefly, in the CEFR (Alderson, 2010 : 245-246)⁴.

18 Nous avons cependant des contraintes temporelles fortes. Il serait impossible de passer cinq heures (examen UNICert, décrit dans Fischer *et al.*, 2012), ni même deux heures 30 (Cles2), pour positionner nos étudiants, pour ne citer que deux certifications s'inspirant de l'approche actionnelle. Par ailleurs, le modèle autocorrectif nous oblige à proposer essentiellement des exercices à réponse fermée, ce qui est rarement l'aboutissement d'une tâche authentique. Nous avons également quelques questions à réponse ouverte courte (QROC) dans la partie expression écrite courte, mais elles sont très coûteuses en temps (aux niveaux A1 / A2 par exemple, le temps de réponse moyen aux questions à réponse fermée est de 26 secondes, et de 49 secondes pour les QROC). De plus, les candidats sont dans une situation d'évaluation institutionnelle et non dans une situation de communication authentique (mais cela est vrai dans l'enseignement aussi bien que l'évaluation, même si l'on peut considérer avec Puren, 2006, que la "tâche d'apprentissage" a sa place dans l'approche actionnelle). Si l'on arrive à reproduire quelques caractéristiques de certaines tâches linguistiques authentiques (se relire et corriger un premier brouillon par exemple), d'autres tâches se prêteront moins à cet exercice (une discussion informelle entre amis).

2.3. Opérationnalisation de l'"authenticité"

19 Les échelles d'exemples de descripteurs du *CECRL* "montrent le continuum de la capacité en situation authentique" (Conseil de l'Europe, 2001 : 139). C'est ce concept d'authenticité, qui est crucial dans l'approche actionnelle comme dans les méthodes communicatives, que nous avons dû problématiser afin de l'opérationnaliser dans notre test.

20 L'authenticité est d'abord, d'après Jordan (1997), cité par Kilickaya (2004), une propriété des textes que l'on utilise pour la compréhension de l'oral et de l'écrit, des textes qui n'ont pas été produits pour des apprenants. Nous essayons le plus possible d'utiliser des textes authentiques, ce qui a un effet positif sur la motivation des apprenants (Kilickaya, 2004), mais aussi sur la validité faciale du test (c'est-à-dire la façon dont le test est reçu et considéré comme un "bon" test par ses utilisateurs). Nous pouvons citer ici certaines remarques de nos étudiants extraites des questionnaires d'évaluation anonymes distribués après la passation lors de l'expérimentation : "Être en situation, vrai [sic] situation = positif", "extraits de conférences super intéressants" (février 2014), "vidéos anglophones mettent en contact direct avec la langue" (avril 2014).

21 Cependant, cette conception de l'authenticité liée uniquement à la nature du document utilisé est quelque peu réductrice, dans la mesure où l'utilisation que l'on fait de ces documents est cruciale. Widdowson (1990), par exemple, réserve le terme "authentique" à l'utilisation d'un document en classe similaire à celle pour laquelle il a été conçu. Il utilise un autre terme ("*genuine*", un proche synonyme), pour des documents authentiques utilisés de façon un peu artificielle, comme par exemple des articles de journaux dont on découpe des paragraphes pour que les élèves les remettent en ordre. Bien que beaucoup des documents que nous utilisons soient authentiques, les tâches d'évaluation qui les accompagnent le sont rarement, bien que certaines s'en rapprochent. Par exemple, au niveau A1, le texte peut être un enregistrement d'un message diffusé dans un aéroport, et la question est "Vous rentrez à Paris. À quelle porte faut-il vous rendre pour prendre votre avion ?"

22 Enfin, le concept d'authenticité a été redéfini par Bachman (1990) comme une propriété de la relation entre le candidat et la tâche, c'est-à-dire que les connaissances, les stratégies et les opérations cognitives sont semblables à celles qui sont mobilisées dans une tâche de la vie réelle. Pour vérifier que c'est le cas, il faudrait mettre en place des protocoles où les étudiants verbalisent ce qu'ils font quand ils répondent à un des items (*Think Aloud Protocols*), ce que nous n'avons pas fait pour l'instant, même si certains chercheurs commencent à se pencher sur cette question (Zoghalmi, 2014).

23 En conclusion, on peut dire que dans une situation de test, la vraie tâche (sociale) que sont en train d'accomplir les étudiants, c'est l'évaluation elle-même. Nous pouvons donc au mieux

"faire semblant" même si nous essayons de mettre les étudiants "en situation" avec une définition du contexte, et une mise en relation entre le type de texte et la question qui est posée.

3. La situation d'énonciation dans les tâches d'évaluation : la contribution du genre textuel et du contexte

- 24 Le terme "tâche d'évaluation", à la place d'"exercice" ou "item", est adopté ici pour indiquer "ce qu'un candidat doit faire pour accomplir une partie du test et qui suppose plus de complexité qu'une réponse à un seul item discret" (ALTE, 2011 : 82). L'adoption volontaire de cette terminologie, très éloquente, est également un choix conceptuel qui nous conduit vers une approche du test plus holistique et inclusive, qui met en relief la valeur des compétences intégrées telles que les compétences communicatives, actionnelles, discursives.
- 25 Dans la perspective communicative-actionnelle,
- nous passons du paradigme de la connaissance, de la simplification à celui de la compétence, de la complexité, qui ne veut pas disjoindre l'objet et le sujet, la réflexion et l'action, l'apprenant et l'usager mais les conjointre pour que la finalité de l'apprentissage d'une langue ne soit pas la bonne note mais l'utilisation autonome de la langue (Bourguignon, 2006 : 63).
- 26 Comme argumenté à plusieurs reprises et selon différentes perspectives ci-dessus, les tâches d'évaluation dans un test en langues autocorrectif proposent une situation d'énonciation qui est faussée par rapport aux échanges communicatifs dans la réalité. Cependant, une réflexion autour des types et des genres textuels dans leurs contextes d'usage reste prioritaire pour déterminer le degré d'authenticité situationnelle – "se référant à l'exactitude avec laquelle les tâches et les items représentent des activités langagières telles qu'on les trouve dans la vie quotidienne" (ALTE, 2011 : 14) – et interactionnelle – "se référant au caractère naturel dans l'interaction que mène le candidat en accomplissant une tâche et les processus mentaux qui entrent en jeu" (ALTE, 2011 : 14) – de la tâche d'évaluation. En particulier, nous nous interrogeons sur la relation qui se crée entre texte et question(s), ce qui détermine et influence la posture du candidat / apprenant mis en situation par la tâche d'évaluation.
- 27 En fait, dans une perspective cognitive et didactique, la difficulté d'une tâche d'évaluation ne dépend pas uniquement des caractéristiques des inputs langagiers, au sens large (Est-ce qu'il s'agit d'un texte écrit ou oral ? Est-il monogéré ou polygéré ? Y a-t-il un ou plusieurs interlocuteurs ? Quelles sont les opérations sollicitées par les questions ?). Elle dépend également des caractéristiques préalables des candidats en termes de connaissances et compétences (encyclopédiques, culturelles, disciplinaires) et des facteurs individuels internes et externes, parmi lesquels nous citerons les orientations motivationnelles et les styles cognitifs préférentiels. Les caractéristiques individuelles des apprenants, pour reprendre la terminologie de Dörnyei (2005), peuvent influencer de manière non négligeable les résultats aux tests et favoriser ou défavoriser certains candidats (Shohamy, 2001 ; Porcelli, 2005).
- 28 Vue l'impossibilité de s'abstraire de la représentation d'un apprenant prototypique auquel le test serait adressé (Holtzer, 1995), s'interroger sur le modèle d'apprenant et sur le modèle de langue est une étape incontournable en phase de conception et de projet. Les textes et les contextes proposés dans les tâches d'évaluation contribuent à déterminer l'identité communicative, sociale et linguistique du test entier, dimension qui devient de plus en plus importante si le test est institutionnel et inséré dans un dispositif d'apprentissage et de remédiation plus vaste (macro-dispositif) (Rivens Mompean, 2013).

3.1. Autour du texte et du discours

- 29 La communication écrite ou orale s'exprime le plus souvent par des textes hybrides distingués par une hétérogénéité compositionnelle et discursive élevée et incontournable. Cependant, la conception et la rédaction d'un test d'évaluation implique, préalablement et nécessairement, un choix averti et ponctuel de plusieurs textes, appartenant à des types, genres et domaines différents.
- 30 Chaque texte est la solution spécifique d'une tâche langagière complexe, solution dans laquelle le mode d'ancrage n'entre qu'à titre de composante sous dépendance directe de la stratégie discursive engagée (avec son but et son destinataire) (Bronckart *et al.*, 1985 : 134).

- 31 Les critères qui déterminent le choix d'un texte, et donc les raisons qui nous permettent d'affirmer que le texte choisi est intéressant, utile et cohérent avec les buts du test et avec son construit, relèvent de plusieurs sphères : didactique, langagière, socio-culturelle.
- 32 Dans le cadre de Self, et plus généralement de la conception d'un test selon une approche communicative, la sphère que nous appelons "didactique" concerne entre autres la classification des ressources selon les niveaux du *CECRL*.
- 33 Le *CECRL*, aussi bien que les référentiels⁵ adossés au *Cadre*, suggèrent pour chaque niveau et pour chaque compétence des genres textuels à exploiter. Comme explicité dans certains référentiels de langues (Hawkins *et al.*, 2012 ; Benucci, 2007 ; Spinelli & Parizzi, 2010), certains genres textuels sont plus aisément proposables à des étudiants ayant un certain niveau de compétence. Cette "association", qui n'est pas univoque, se base sur les traits saillants qui sont plus souvent présents dans les textes, en termes d'enchaînement entre différentes séquences discursives mais aussi sur des aspects fonctionnels et sur le degré d'inférence que le lecteur doit activer pour comprendre et interagir avec le texte. Un genre détermine, en fait, une organisation et un enchaînement particuliers de séquences discursives, et sollicite / engendre des objectifs langagiers qui s'expriment par des aspects fonctionnels et des choix morphosyntaxiques et lexicaux. Cette concrétisation relève de la sphère que nous avons ici nommée "langagière". Une carte postale simple ou une pièce d'identité sont fréquemment proposés aux étudiants ayant un niveau débutant / élémentaire de compétence en langues, tandis qu'un article de fond est plus facile à exploiter pour un groupe d'étudiants de niveau avancé.
- 34 La sphère que nous appelons ici socio-culturelle concerne la pertinence communicative, pragmatique et culturelle du sujet traité ou véhiculé par le texte (ayant son schéma interactionnel), et donc de façon indirecte, sa cohérence avec le principe d'authenticité. Pensons, par exemple, aux nouveaux genres textuels. La riche panoplie de genres actuels qui ressortissent de la communication médiatisée par ordinateur ou sur dispositif mobile, synchrone ou asynchrone (par exemple SMS, clavardage, forum, courrier électronique, blogue, etc.) doit être fortement représentée dans un corpus de textes qui prétend s'ancrer dans des approches communicatives (Pistoiesi, 2004).
- 35 La réflexion sur la notion de *type* et de *genre* aide à conscientiser les traits les plus marquants d'un texte, et donc, dans une perspective de conception et de rédaction de tâche, à choisir les textes avec des yeux plus avertis. Les différences les plus évidentes concernent l'organisation des séquences discursives, les types de discours proposés, la densité sémantique, les contextes d'usage plus probables (en termes de domaine public, personnel, professionnel et éducationnel).
- 36 Très éclairant et pertinent pour nos buts est le débat ancien, mais encore tout à fait actuel, entre les notions-concepts de texte et de discours, récemment repris par Michèle Monte et Gilles Philippe (2014). Nous pourrions aborder le sujet depuis plusieurs perspectives mais trois en particulier nous semblent intéressantes comme socle pour justifier / argumenter l'approche Self.
- 37 i) la perspective sémiotique, où le texte est défini en raison du rapport avec son lecteur (Eco, 1985 ; Guillemette & Cossette, 2006). Le texte est un tissu de signes. Il est ouvert, interprétable, mais doit être entrevu comme un tout cohérent. Il construit son lecteur modèle, et est davantage une totalité où l'auteur amène les mots puis le lecteur le sens. Le texte est en fait une "machine paresseuse qui exige du lecteur un travail coopératif acharné pour remplir les espaces de non-dit ou de déjà-dit restés en blanc" (Eco, 1985 : 29).
- 38 Cette vision met au premier plan l'interprétation en tant que dialogue, implicite, entre l'auteur – ses intentions, ses buts, etc. – et le lecteur – ses schémas interprétatifs (Keller, 2007) et ses cadres socio-culturels. Dans l'évaluation, l'approche sémiotique est parfaitement cohérente avec la nécessité de centrer la conception et la passation d'un test sur un public de candidats prédéfini.
- 39 ii) la perspective applicative d'exploitation / usage d'un texte en didactique. Dans le *CECRL*, par exemple,

est défini comme toute séquence discursive (orale et / ou écrite) inscrite dans un domaine particulier et donnant lieu, comme objet ou comme visée, comme produit ou comme processus, à activité langagière au cours de la réalisation d'une tâche (Conseil de l'Europe, 2001 : 15).

40 Cette définition nous intéresse également puisqu'elle introduit le concept de séquence discursive (au-delà du domaine et de l'activité langagière) nous aidant à aborder le nœud de la discussion, c'est-à-dire le lien préférentiel qu'il peut y avoir entre genre textuel, séquence discursive prédominante et niveau de difficulté de la tâche d'évaluation.

41 Les séquences discursives (séquence narrative, séquence descriptive, etc.), quant à elles, constituent des unités empiriques et correspondent aux segments à partir desquels les types de discours se manifestent effectivement dans la réalité des productions langagières. Comme le rappelle Adam (1992), ces séquences présentent une configuration structurelle propre qui leur donne une autonomie relative au sein du discours dont elles relèvent (Filliettaz & Grobet, 2001 : 225).

42 Et, enfin :

iii) la perspective de la linguistique textuelle et de l'analyse du discours qui semblent rapprocher les deux précédentes.

43 On peut faire l'hypothèse que le genre comme ensemble de normes intériorisées exerce des déterminations variables sur les différents éléments de la textualité : il peut contraindre parfois fortement la structure compositionnelle du texte mais laisser une grande marge de liberté sur le plan sémantique ou énonciatif, d'autres fois contraindre l'ancrage énonciatif mais peser à peine sur la composition du texte ou sur la représentation sémantique, et ainsi de suite (Monte *et al.*, 2014 : 11).

44 Toutes ces perspectives contribuent de manière constructive à enrichir les modalités de traitement et de prise en compte du texte dans Self. Avant d'observer des exemples tirés de Self en italien et en anglais, nous nous proposons d'observer que le rôle du contexte joue dans la tâche d'évaluation.

3.2. Le contexte dans Self : entre fiction et réalité

45 Dans le *CECRL*,

le contexte renvoie à la multitude des événements et des paramètres de la situation (physiques et autres), propres à la personne mais aussi extérieurs à elle, dans laquelle s'inscrivent les actes de communication. (...) Les besoins et les désirs de communiquer naissent d'une situation donnée et, la forme comme le contenu de la situation, répondent à cette situation (Conseil de l'Europe, 2001 : 41).

46 Les contextes situationnels d'utilisation de la langue sont organisés dans le *CECRL* en catégories descriptives, chacune comportant une distinction de domaine, entre public, privé, professionnel, éducationnel. Par exemple, un "lieu" (chez soi, dans sa famille, à l'école, au bureau, etc.), des "institutions" (la famille, les hôpitaux, les entreprises, les autorités, etc.), des "événements" (fêtes, mariages, réunion, etc.), des "textes" (les journaux, les annonces, les textes d'affaire, les documents authentiques, etc.), etc.

47 C'est dans cette perspective qu'une contextualisation est nécessaire dans une tâche d'évaluation. L'idée est, en fait, celle de contextualiser la situation d'énonciation. Le contexte est donc explicité en tant qu'information courte, fournie en supplément de l'objet de la question, que l'auditeur / lecteur / rédacteur est censé avoir lorsqu'il est en situation d'écoute / lecture / rédaction authentique. Cette information aide à compenser le déficit informatif qui est le propre de la situation dans laquelle le sujet testé est placé (Masperi, 2014) et elle se relie à cette exigence naturelle explicitée dans le *CECRL* en servant de "mise en situation"⁶.

48 Les tâches d'évaluation Self proposent trois typologies de "contextes".

- *Type d'échange* : pour indiquer le rapport social et sociolinguistique entre les participants (mère / fils ; directeur / collaborateur ; entre amis, etc.), et donc donner des indications sur le registre (niveau formel / informel) et déterminer s'il s'agit d'un texte monologal ou dialogal (entre deux ou plusieurs personnes). Quelques exemples : "Conversation en famille", "Dialogue entre amis", "Entretien avec le directeur".

- *Lieu* : pour indiquer le *lieu physique* dans lequel se déroule l'action (lieu public, privé, éducationnel, professionnel). Dans cette catégorie, nous incluons aussi des *lieux fictifs*, par exemple le canal de transmission du texte. C'est le cas, par exemple, du contexte "Au téléphone". Nous aurions également pu classifier le contexte "Au téléphone" parmi la catégorie des genres textuels (si on avait voulu mettre la focale sur les routines conversationnelles et sur les échanges qui ont lieu lors d'une interaction au téléphone).
- *Genre textuel* : le contexte, dans ce cas, est indiqué d'après la classification des genres textuels (fait divers, annonce, interview, enquête, sondage, publicité, présentation, etc.).

49 Des contraintes de temps et d'espace obligent le rédacteur à réduire ou à manipuler certains textes, qui seraient trop longs ou denses d'inputs langagiers. Le cas qui se produit fréquemment concerne l'extrapolation d'une section spécifique d'un texte authentique. Prenons l'exemple d'une information radio. Le journaliste radio annonce, au début de l'émission, la liste des nouvelles et des informations qu'il abordera tout au long de la transmission, en anticipant très rapidement certains contenus. Dans notre cas de figure, où une nouvelle spécifique est reproposée en contexte didactisé, le candidat / étudiant entendra directement le texte oral qui – selon la nomenclature attribuée aux composantes des tâches Self – s'appellera "objet de la question". L'"objet de la question" est l'objet sur lequel porte(nt) la ou les questions. C'est l'objet principal sur lequel porte l'effort de compréhension écrite ou orale. Autrement dit encore, c'est l'action sur laquelle on porte l'attention.

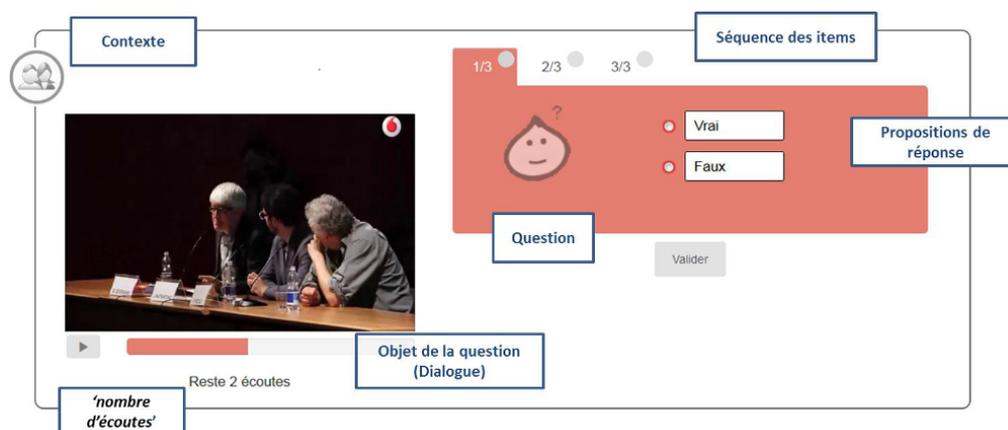
50 Évidemment en contexte de test en ligne, tout élément extérieur qui contribue à la contextualisation de la situation d'énonciation est absent. Par exemple, il n'est pas évident d'avoir à disposition des coordonnées spatio-temporelles claires ; par contre, en contexte authentique, l'écouteur connaît le canal de transmission du message (la radio mais également l'émission dont elle fait partie), le moment de la journée (matin, après-midi, soir, etc.), et peut-être parfois d'autres traits / éléments qui concernent les modalités expressives et les opinions d'un journaliste ou l'entourage culturel et social de la transmission. De plus, plusieurs nouvelles pourraient aussi être enchaînées entre elles de manière significative, en enrichissant le contexte et son utilité pour la compréhension d'un message complexe.

51 Toutes ces considérations nous ont conduites à mettre à disposition des contextes enrichis ou hybrides, où le type d'échange s'intègre à la spécification du lieu, physique ou fictif, ou bien où le genre textuel est spécifié à côté du type d'échange (par exemple : "Courriel entre père et fils" ; "Publicité informative à la gare") : les contextes enrichis sont proposés pour ne pas pénaliser la compréhension d'un texte oral ou écrit, qui est forcément proposé hors de son contexte naturel et en absence d'un titre ou d'une mise en situation. Normalement, la quantité de détails dans les "contextes enrichis" est plus élevée que dans le cas des autres types de "contexte" décrits. Les noms propres des protagonistes de l'échange peuvent être mentionnés explicitement dans le contexte surtout si les noms n'apparaissent pas dans le texte source de la tâche. C'est le cas par exemple de "Dialogo tra Paola et Martina" (dialogue entre Paola et Martina) à la place de (dialogue entre amies). Là, on décide de répéter les noms des protagonistes de l'échange pour permettre de les reconnaître sans souci lors du dialogue.

3.3 Exemplification des termes-concepts dans une tâche de compréhension de l'oral

52 Prenons le cas d'une tâche d'évaluation proposée pour évaluer la compréhension de l'oral au niveau B2, et voyons comment les éléments – texte / objet de la question, contexte, discours, rôle du candidat – décrits ci-dessus, s'enchaînent en déterminant aussi le niveau de difficulté de la tâche.

Figure 1 - Tâche intitulée "Connexions" proposée en compréhension de l'oral dans la version de Self-italien.



Comme

représenté dans l'image ci-dessus⁷ (imprimé d'écran d'une tâche de compréhension de l'oral qui s'intitule "Connexions"), la tâche est composée de plusieurs éléments, tous oralisés. En haut à gauche est placé l'icône du contexte ; en dessous, l'objet de la question qui, dans ce cas, est une vidéo ; sur la partie droite de l'écran nous trouvons les questions et les propositions de réponse. Le contexte, les questions et les propositions de réponse peuvent être réécoutés un nombre illimité de fois ; par contre, l'objet de la question jusqu'à deux, voire trois, fois selon la longueur du texte et le nombre de questions posées.

53 La vidéo proposée est tirée d'un reportage plus long intitulé "*Sconnessi o iperconnessi*" et tourné à l'occasion du Salon international du livre de Turin en Italie. Le protagoniste est un journaliste italien célèbre et expert de questions de société, Beppe Severgnini, qui décrit son point de vue sur l'usage d'Internet dans la vie quotidienne, en exposant le risque qui dérive de l'excès de connexion (*hyperconnexion*) vis-à-vis de l'absence de connexion (*déconnexion*).

54 Nous avons proposé ici un contexte enrichi, qui présente le genre textuel et mentionne le protagoniste en anticipant le contenu : *intervista a un giornalista sull'uso di internet* (interview d'un journaliste sur l'usage d'Internet). La vidéo elle-même apporte d'autres indices visuels qui éclairent la situation.

55 Il s'agit d'un texte argumentatif, qui est intéressant surtout pour la présence d'expressions lexicales imagées, qui proposent des comparaisons figurées et métaphoriques. Le rapport entre les questions et le texte-objet de la question, dans ce cas de type vrai / faux, met le candidat / apprenant dans une position d'auditeur / spectateur qui est sollicité pour activer des opérations cognitives variées : compréhension globale du texte, recherche et compréhension d'un ou plusieurs détails, inférence.

4. La centration sur l'apprenant dans un contexte d'évaluation autocorrective ? Évaluer pour positionner, diagnostiquer, former

56 La centralité de l'apprenant est l'un des concepts-clés de l'approche communicative, concept qui trouve son fondement psychologique dans "l'apprentissage en tant qu'activité personnelle du sujet apprenant" (Puren, 1995). Si les facteurs que nous venons d'illustrer dans les paragraphes précédents sont déjà révélateurs de l'attention que l'approche Self consacre à l'apprenant, d'autres éléments identitaires contribuent à en faire un système cohérent avec les approches communicatives : Self est un micro-dispositif pour l'évaluation des compétences en langues en communication intégré dans un environnement plus vaste, le macro-dispositif ENPA, "Environnement numérique personnalisé pour l'apprentissage des langues" (cf. 4.1.). Self, en tant que système diagnostique, constitue une interface entre l'évaluation et l'apprentissage et vise à mettre à disposition des informations contrôlées qui alimentent le profil de l'apprenant (cf. 4.2.). À côté des analyses psychométriques plus classiques, une attention particulière est apportée à l'autoévaluation et à la perception de soi en tant qu'apprenant de langue. Nous mentionnerons ici en particulier l'observation de la

corrélation entre la difficulté perçue par les étudiants vis-à-vis de la tâche d'évaluation et la difficulté relevée par les indices de difficulté statistique selon les analyses de la théorie classique des tests (cf. 4.3.).

4.1. Le micro-dispositif Self dans le macro-dispositif ENPA

57 Parmi les faiblesses les plus communes des tests de positionnement ou de diagnostic en langues, à disposition dans les institutions universitaires ou sur le marché, nous remarquons en particulier i) une carence dans le feedback personnalisé post-test, en tant que retour formatif qui puisse tirer profit de la valeur des erreurs comme indices d'apprentissage et mettre en exergue les points de force, et ii) une certaine autoréférentialité du test qui, dans la plupart des cas, est isolé et n'entre pas dans un rapport dialogique avec des parcours de formation et / ou de remédiation.

58 Cet état de choses est en contradiction avec deux concepts-clé de l'apprentissage holistique et contextualisé, c'est-à-dire celui du dispositif de formation (Linard, 1998 ; Peraya, 1999 ; Jacquinet-Delaunay & Monnoyer, 1999), qui est censé être autonomisant, hybride et centré sur l'apprenant, et celui d'apprentissage intégré (Charlier *et al.*, 2006). "Un apprentissage intégré renvoie à la prise en compte, à l'intégration de toutes les dimensions d'un objet de connaissance" (Charlier *et al.*, 2006 : 472). Le lien nécessaire entre évaluation et apprentissage est également souligné par Kunnan et Jang (2009 : 621).

Shohamy (1992) proposed a collaborative diagnostic feedback model in which tests provide useful information for teaching and learning advancement. (...) Alderson (2006) concluded that in order to diagnose language development, we will need to have a much better understanding of foreign language ability. If we can then incorporate such understandings into assessment and make them useful to learners through the provision of meaningful and useful feedback and follow-up, then diagnosis will truly have become the interface between learning and assessment.

59 Charlier *et al.* (2006) poursuivent leur argumentation autour du concept de dispositif de formation et d'apprentissage hybride en rappelant que Valdès (1995) décrit

l'espace hybride de formation comme étant centré sur l'apprenant et articulant : des parcours négociés, un rythme individualisé, des lieux multiples, des ressources décentralisées et accessibles à distance, des situations pédagogiques adaptées, des média diversifiés et adaptés, une pédagogie individualisée. Dans ce cadre, l'intégration de la présence et de la distance et l'usage des technologies permet de rencontrer davantage ses objectifs de "dispositif centré apprenant".

60 C'est dans cette perspective que Self se propose de poursuivre une démarche de conception et de projet fondée sur une intégration entre dispositif d'évaluation et dispositif d'apprentissage et de remédiation.

61 Comme décrit dans Masperi & Quintin (2014b), l'ENPA est le socle fédérateur de l'écosystème numérique d'apprentissage *Innovalangues*.

Un "écosystème numérique" d'apprentissage est envisagé comme un ensemble structuré, complexe et ouvert d'éléments physiques (techniques et humains) interdépendants, réunis en un seul lieu. Du fait de cette interdépendance, les propriétés globales du système ne peuvent se réduire à la somme des propriétés de chacun de ses composants. Transposé sur le versant de la formation en langues, le terme "écosystème" souligne précisément, à notre sens, tant la notion de variété et d'interdépendance des éléments qui participent à tout processus formatif et acquisitionnel, que la vocation à évoluer dont ce système doit faire preuve, afin de s'adapter aux contraintes et aux conditions auxquelles il est exposé (Masperi & Quintin, 2014b : 73).

62 Les résultats au test, soient-ils généraux et concernant le niveau global atteint ou bien organisés par compétence (compréhension de l'oral, compréhension de l'écrit, expression écrite courte) sont une des informations qui, comme dans un puzzle, contribuent à nourrir le profil de l'apprenant. Le profil ébauché après la passation du test sera approfondi, ou, au contraire, démenti et remodelé, par d'autres résultats qui s'accumulent au fur et à mesure et de façon itérative tout au long du parcours d'apprentissage. De plus, les résultats au test pourront également guider l'étudiant vers des parcours de remédiation personnalisée.

4.2. L'apport de Self dans l'arborescence du diagnostic

63 Le sujet du diagnostic dans l'évaluation des compétences langagières est très peu traité et débattu dans le domaine du *testing*, d'une part, à cause de la complexité didactique et scientifique du thème, et, de l'autre, à cause de la difficulté à investir des fonds de recherche et développement à long terme. Le processus de diagnostic en langues, en tant qu'interface (pont, lien) entre apprentissage et évaluation est en fait itératif, interactif et continu (Lengyel, 2011).

*Diagnostic tests seek to identify those areas in which a student needs further help. These tests can be fairly general, and show, for example, whether a student needs particular help with one of the four main language skills; or they can be more specific, seeking perhaps to identify weaknesses in a student's use of grammar*⁹ (Alderson, Clapham & Wall, 1995 : 12).

64 La littérature spécialisée dans le domaine semble plutôt proposer de manière réitérée certains concepts théoriques importants – tels que la remédiation, les points de force et de faiblesse dans la compétence langagière, la mise à disposition d'un feedback détaillé et formatif, l'ancrage à un syllabus interne à l'apprenant (*inbuilt syllabus*), l'auto-régulation et la personnalisation – sans mettre à disposition des solutions opérationnelles éclairantes.

65 Self vise à mettre à disposition de son public élargi, constitué bien sûr des étudiants mais aussi des enseignants et des tuteurs, des informations motivées sur les connaissances acquises et à acquérir. La contrainte cognitive et linguistique liée au processus d'inférence à faire entre les réponses collectées et l'élaboration d'un retour in-formatif pertinent engendre des questions préalables, soit sur le plan qualitatif (quelles informations saillantes collecter et comment les interpréter ?), soit sur le plan quantitatif, (combien de données observer / collecter avant de pouvoir faire une généralisation sur les points de force et de faiblesse d'un étudiant ?).

66 Ceci dit, le choix des informations et des données provenant du système Self qui pourraient alimenter le profil de l'apprenant nous conduit, encore une fois, à prendre position sur les facteurs qui influencent la difficulté d'une tâche d'évaluation. En premier lieu, nous allons considérer les facteurs concernant la perception de la compétence en autoévaluation, en termes de compétence globale / générale en langue (savoir-faire, savoir-être, savoir apprendre), puis détaillés dans les habilités productives et réceptives. En deuxième lieu, d'autres données à considérer sont issues directement du traçage des résultats après la passation du test : la / les moments de passation du test (coordonnées spatio-temporelles), le temps employé (pour le test entier et pour chaque compétence), les résultats obtenus dans les sous-compétences et selon la focalisation langagière (morphosyntaxe, lexicale, communication) de chaque question. Enfin, d'autres caractéristiques intéressantes proviennent plus directement de la tâche d'évaluation et de ses composantes, à savoir la nature du texte (objet de la question) et des questions posées de manière directe (par exemple, questionnaire à choix multiple) ou indirecte (appariements, textes lacunaires, etc.) et les opérations cognitives sollicitées (compréhension globale, détaillée, inférence, etc.).

4.3. Le rapport entre la difficulté perçue et les indices de difficulté

67 Une autre manière de se centrer sur l'apprenant est de prendre en compte ses perceptions et de favoriser une activité réflexive sur son apprentissage. À ce propos, lors de l'organisation des tests pilotes nécessaires au processus de validation de Self (ALTE, 2011 : 25-26), il a été demandé aux étudiants d'explicitier la difficulté perçue de chaque tâche sur une échelle allant de 1 à 5 (1 = très facile, 2 = facile, 3 = normal, 4 = difficile, 5 = très difficile). Nous avons ensuite calculé la corrélation entre la difficulté moyenne de chaque tâche évaluée subjectivement par les étudiants et l'indice objectif de facilité (taux de réussite à la tâche). Étant donné que ces deux variables utilisent des échelles différentes, nous avons utilisé le Rho de Spearman (en anglais, *rank order correlation*) qui calcule la corrélation, non pas directement entre les valeurs de deux variables, mais entre les rangs de ces valeurs (Bachman, 2004). L'analyse a été limitée cette fois à un échantillon de 16 tâches de compréhension de l'oral et 14 de compréhension de l'écrit testées sur un public de 59 étudiants. Les résultats sur les deux compétences ont été séparés. Les résultats sont assez encourageants, puisque l'indice de corrélation est de $r = 0,5$ pour la compréhension orale, et $r = 0,64$ pour la compréhension écrite. Ce ne sont que des résultats préliminaires qu'il faudrait confirmer en étendant l'analyse aux autres pilotes et à

toutes les compétences évaluées, mais nous pensons que c'est un premier indice montrant que les étudiants, même sans formation préalable, sont capables d'avoir une activité réflexive utile et dont les résultats pourront donc être pris en compte.

5. Conclusions

- 68 D'après Doucet (2001 : 10), "un test utile est un test qui permet de faire des prédictions raisonnables tout en limitant les sources d'erreur dans la prise de mesure et les biais dans les interprétations des résultats". Si nous voulons que le résultat du test Self fournisse des informations utilisables sur la capacité à communiquer des étudiants, il faut s'assurer que le construit de compétence communicative est bien opérationnalisé. D'après Messick (1989 : 34), le danger principal qui guette les tests de langue est la sous-représentation du construit, c'est-à-dire qu'il ne couvre pas toutes les facettes de la capacité qui est censée être mesurée ("*the test is too narrow and fails to include important dimensions or facets of the construct*"). Dans cet article, nous avons essayé de décrire les démarches adoptées pour garantir la représentativité et la cohérence du construit de Self positionnement et diagnostic, et pour compenser les faiblesses de l'évaluation autocorrective (temps limité, nombre réduit de tâches au sens actionnel, évaluation très limitée des capacités de production par rapport à la compréhension, décontextualisation, etc.) dans une perspective communicative. Ces procédés prennent la forme d'une contextualisation des tâches d'évaluation par une attention poussée prêtée à la définition du contexte / situation, d'une prise en compte des genres textuels et d'une centration sur l'apprenant par l'intégration de l'autoévaluation et de la réflexivité sur l'apprentissage.
- 69 Self étant en cours de développement et d'expérimentation, les analyses présentées sont une exploration préliminaire qui sera suivie d'analyses plus complètes sur la validité et l'utilité du test. Nous sommes engagées dans une démarche de rédaction d'un argument de validité, dont la définition du construit et la description de son opérationnalisation ne sont que les deux premières étapes. Un argument de validité "consiste en une série de propositions qui décrivent les raisons selon lesquelles les interprétations conseillées des résultats au test sont valides, et apportent la preuve pratique et la théorie qui les étayent" (ALTE, 2011). D'autres preuves devront donc être rassemblées pour montrer que la mesure de la performance des candidats est cohérente et stable, c'est-à-dire que le test est fiable, que le processus de mise en relation des scores avec le CECR est bien pensé (définition des scores de césure pour le passage d'un niveau à un autre), ou que les résultats à Self sont corrélés à ceux obtenus à d'autres tests de construit similaire.

Références

Bibliographie

- Adam, J.-M. (1992). *Les textes : types et prototypes*. Paris : Armand Colin.
- Alderson, C. (2006). *Diagnosing Foreign Language Proficiency: The Interface between Learning and Assessment*. London, New York : Continuum Edition.
- Alderson, C. (2010). "Language testing-informed SLA? SLA-informed language testing? ". In Bartning, I., Maisa, M. & Vedder, I. (dir.). *Communicative development and linguistic development: intersections between SLA and language testing research*. Modena, Reggio Emilia : Eurosla Monographs Series 1.
- Alderson, C., Clapham, C. & Wall, D. (1995). *Language Test Construction and Validation*. Cambridge : Cambridge University Press.
- ALTE, (2011). *Manuel pour l'élaboration et la passation de tests et d'examens de langue*. Division des Politiques linguistiques. Strasbourg : Conseil de l'Europe, DG II – Service de l'éducation.
- Austin, J. L. (1962). *How to do things with words*. Oxford : Oxford University Press. Traduction par Lanne, G. (1970). *Quand dire, c'est faire*. Paris : Le Seuil.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford : Oxford University Press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge : Cambridge University Press.

- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford : Oxford University Press.
- Bachman, L. F. & Palmer, A. S. (2010). *Language assessment in practice*. Oxford : Oxford University Press.
- Benucci, A. (2007). *Sillabo di italiano per stranieri*. Perugia : Guerra Edizioni.
- Bourguignon, C. (2006). "De l'approche communicative à l'approche communic'actionnelle : une rupture épistémologique en didactique des langues-cultures". *Synergie Europe*, no° 1. pp. 58-73.
- Bourguignon, C., Delahaye, P. & Vicher, A. (2005). "L'évaluation de la compétence en langue : un objectif commun pour des publics différents". *Études de linguistique appliquée (ÉLA)*, 2005/4, n° 140. pp. 459-473.
- Bouscaren, J. & Chuquet, J. (1987). *Grammaire et textes anglais – Guide pour l'analyse linguistique*. Paris : Ophrys.
- Bronckart, J.-P., Bain, D. E., Schneuwly, B., Davaud, C. & Pasquier, A. (1985). *Le fonctionnement des discours*. Neuchâtel : Delachaux et Niestlé.
- Canale, M. (1983). "From communicative competence to communicative language pedagogy". In Richards, J. C. & Schmidt, R. W. (dir.). *Language and communication*. New York : Longman.
- Carrère, C., Corcoran, A., Masperi, M., Payre-Ficout, C. (2009). *État des lieux de l'utilisation des tests de positionnement en anglais dans l'enseignement supérieur*. Enquête Grenoble Universités 2009. En ligne : http://lansad.u-grenoble3.fr/version-francaise/projets-et-bilans/bilans-et-rapports/rapports-d-enquete-166300.kjsp?RH=U3LANFR_BILRAP
- Celce-Murcia, M., Dörnyei, Z. & Thurrell, S. (1995). "Communicative competence: A pedagogically motivated model with content specifications". *Issues in Applied Linguistics*, vol. 6, n° 2. pp. 5-35.
- Cervini, C. (2014). "La valutazione multilingue nel contesto dei dispositivi formativi: il sistema 'SELF' per il posizionamento e la diagnosi delle competenze linguistiche". *LEND – Lingua e Nuova Didattica*. Periodico di Linguistica Applicata e Glottodidattica. n.° 1, février 2014. pp. 16-26.
- Cervini, C., Masperi, M., Jouannaud, M.-P. & Scanu, F. (2013). "Defining, modeling and piloting SELF, a new formative assessment test for foreign languages". In Colpaert, J., Simons, M., Aerts, A. & Oberhofer, M. (dir.). *Conference Proceedings, "Language Testing in Europe: time for a new framework", 29-31th May 2013*. Antwerp : University of Antwerp. pp. 16-26.
- Charlier, B., Deschryver, N. & Peraya, D. (2006). "Apprendre en présence et à distance". *Distances et savoirs*, vol. 4. pp. 469-496. Disponible en ligne : www.cairn.info/revue-distances-et-savoirs-2006-4-page-469.htm.
- Cicurel, F. (1991). *Lectures Interactives*. Paris : Hachette.
- Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues – Apprendre, enseigner, évaluer*. Paris : Didier.
- Culioli, A. (1978). "Valeurs modales et opérations énonciatives". *Le Français moderne*, vol. 46, n° 4, repris dans Culioli, A. (1990). *Pour une linguistique de l'énonciation, tome 1*. Paris : Ophrys. pp. 135-156.
- Dörnyei, Z. (2005). *The psychology of the language learner. Individual differences in second language acquisition*. New York & London : Routledge.
- Doucet, P. (2001), "Pour un test utile". *A5p*. n° 34. pp. 13-33. Disponible en ligne : <http://asp.revues.org/1696>. DOI : 10.4000/asp.1696
- Eco, U. (1985). *Lector in fabula*. Paris : Grasset.
- Filliettaz, L. & Grobet, A. (2001). "L'hétérogénéité compositionnelle du discours : quelques remarques préliminaires". *Cahiers de linguistique française*, n° 21. pp. 213-260.
- Fischer, J., Chouissa, C., Dugovicova, S & Virkkunen-Fullenwider, A. (2012). *Évaluer par les tâches les langues à fins spécifiques à l'université – Un guide*. Centre européen pour les langues vivantes. <http://www.ecml.at/tabid/277/PublicationID/68/Default.aspx>
- Galisson, R. & Coste, D. (dir.). (1976). *Dictionnaire de didactique des langues*. Paris : Hachette.
- Guillemette, L. & Cossette, J. (2006). "La coopération textuelle". In Hébert, L. (dir.). *Signo*. Rimouski (Québec), disponible en ligne : <http://www.signosemio.com/eco/cooperation-textuelle.asp>
- Guillemin-Flescher, J. (1981). *Syntaxe comparée du français et de l'anglais*. Gap : Ophrys.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London : Arnold.

- Hawkins J., Filipovic, L. & Green, A. (dir.). (2012). *English Profile Studies. 1 & 2*. Cambridge : Cambridge University Press.
- Hilton, H. (2009). *Systèmes émergents : acquisition, traitement et didactique des langues*. Habilitation à diriger des recherches. Université Lumière Lyon II.
- Holtzer, G. (1995). *Autonomie et didactique des langues – Le Conseil de l'Europe et les langues étrangères (1970-1990)*. Besançon : Presses universitaires de Franche-Comté.
- Hymes, D. (1972). "On communicative competence". In Pride, J. B. & Holmes, J. (dir.). *Sociolinguistics: Selected readings*. Harmondsworth : Penguin.
- Jacquino-Delaunay, G. & Monnoyer, L. (dir.) (1999). *Le dispositif – Entre usage et concept – Hermès*, n° 25. Paris : CNRS éditions.
- Jordan, R. R. (1997). *English for Academic Purposes: A Guide and Resource for Teachers*. Cambridge : Cambridge University Press.
- Keller, R. (2007). "L'analyse de discours du point de vue de la sociologie de la connaissance – Une perspective nouvelle pour les méthodes qualitatives". *Recherches qualitatives – Hors série – numéro 3. Actes du colloque "Bilan et perspectives de la recherche qualitative"*. pp. 287-306.
- Kilickaya, F. (2004). "Authentic materials and cultural content in EFL classrooms". *The Internet TESL Journal*, vol. 10, n° 7. <http://iteslj.org/Techniques/Kilickaya-AuthenticMaterial.html>
- Kunnan, A. J. & Jang, E. E. (2009). "Diagnostic Feedback in Language Assessment". In Long, M. H. & Doughty, C. J. (dir.). *The Handbook of Language Teaching*. Oxford : Wiley Blackwell.
- Lengyel, D. (2011). "La valutazione diagnostica delle competenze linguistiche in un contesto multilingue: un processo continuo che favorisce l'insegnamento e l'apprendimento individualizzato". *Italiano LinguaDue*, vol. 3, n° 1. <http://riviste.unimi.it/index.php/promoitals/article/view/1239>
- Linard, M. (1998). "L'écran de TIC, 'dispositif' d'interaction et d'apprentissage : la conception des interfaces à la lumière des théories de l'action". Colloque "Dispositifs et Médiation des Savoirs", Louvain-La-Neuve, avril 1998. Disponible en ligne : <https://edutice.archives-ouvertes.fr/edutice-00000913>
- Masperi, M. (2012). *Projet IDEFI Innovalangues*. MESR. ANR. Investissements d'avenir. Disponible en ligne : <http://lansad.u-grenoble3.fr/version-francaise/projets-et-bilans/projets/projet-innovalangues-166345.kjsp>
- Masperi, M. & Quintin, J.-J. (2014a). "L'innovation selon *Innovalangues*". *LEND – Lingua e Nuova Didattica. Periodico di Linguistica Applicata e Glottodidattica*. n° 1.
- Masperi, M. & Quintin, J.-J. (2014b). "Enseigner à l'université en France, à l'ère du numérique – L'apport de dispositifs innovants dans la formation en langues". In Cervini, C. & Valdivieso, A. (dir.). *Dispositivi formativi e modalità ibride per l'apprendimento linguistico*. Bologna : Edizioni Clueb, Collana Contesti Linguistici.
- Messick, S. (1989). "Validity". In Linn, R. L. (dir.). *Educational measurement*. New York : Macmillan.
- Monte, M., Philippe, G. (dir.). (2014). *Genres & Textes – Déterminations, évolutions, confrontations*. Lyon : Presses universitaires de Lyon.
- Peraya, D. (1999). "Médiation et médiatisation : le campus virtuel". *Hermès*, n° 25. pp. 153-167.
- Pistolesi, E. (2004). *Il parlar spedito. L'italiano di chat, e-mail e SMS*. Padova : Esedra.
- Porcelli, G. (2005). "Verifiche comode e verifiche valide". In Jafrancesco E. (dir.). *La valutazione delle competenze linguistico-comunicative in italiano L2*. Atti del XIV Convegno Nazionale ILSA. Roma : Edilingua Edizioni. pp. 13-29.
- Puren, C. (1995). "La problématique de la centration sur l'apprenant en contexte scolaire". *Études de linguistique appliquée (ÉLA)*, n° 100. pp. 129-149.
- Puren, C. (2006). "De l'approche communicative à la perspective actionnelle". *Le français dans le monde*, n° 347. pp. 37-40.
- Rivens Mompean, A. (2013). *Le centre de ressources en langues : vers la modélisation du dispositif d'apprentissage*. Villeneuve d'Ascq : Septentrion.
- Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice*. Reading : Addison-Wesley.
- Savignon, S. J. (1990). "Communicative language teaching: Definitions and directions". In Alatis, J. E. (dir.). *Georgetown University Roundtable on Language and Linguistics 1990*. Washington, DC : Georgetown University Press.

Shohamy, E. G. (1992). "Beyond Proficiency Testing: A Diagnostic Feedback Testing Model for Assessing Foreign Language Learning". *The Modern Language Journal*, vol. 76, n° 4. pp. 513-521. DOI : 10.1111/j.1540-4781. Disponible en ligne : <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-4781.1992.tb05402.x/epdf>

Shohamy, E. G. (2001). *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. London : Longman.

Spinelli, B. & Parizzi, F. (2010). *Profilo della lingua italiana. Livelli di riferimento del QCER A1, A2, B1, B2*. Milano : La Nuova Italia.

Valdès, D. (1995). *Vers de nouvelles formes de formations – Les formations hybrides*. Paris : Mémoire de DESS, université Paris 2.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford : Oxford University Press.

Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford : Oxford University Press.

Zoghalmi, N. (2014). "Think-aloud reports: Toward an understanding of the cognitive processes of spoken English (L2) comprehension". Communication au colloque international des étudiants chercheurs en didactique des langues et en linguistique (Cedil), 24 juin 2014.

Site

Éduscol (nd). Site du Ministère de l'Éducation Nationale à destination des professionnels de l'éducation. <http://eduscol.education.fr/cid45678/cadre-europeen-commun-de-referance-cecrl.htm>

Notes

1 Ce que nous appelons "contexte" est appelé d'ailleurs "situation" dans ces théories, le terme "contexte" étant réservé au co-texte proprement dit.

2 Cf. Fischer *et al.* (2012 : 17) pour une discussion intéressante des différents termes utilisés en Europe pour cette orientation didactique, et de l'hésitation entre "action" (orienté apprenant) et "tâche" (orienté enseignant) comme terme de base.

3 Par exemple, des étudiants participant aux expérimentations préliminaires nous ont demandé si l'irlandais était la même langue que l'anglais, et d'autres si une tonne correspondait à 100 ou à 1000 kilos (connaissances nécessaires pour répondre à certaines questions du test).

4 "De nos jours, les concepteurs de tests de compétence en langue sont presque forcés de tester le comportement des candidats dans une vaste gamme de situations de communication, de les confronter à une variété de tâches, de tests et d'activités, couvrant toutes les dimensions du comportement communicatif qui sont présentées, parfois brièvement, dans le CECR."

5 "Un référentiel répertorie, décrit les diverses compétences à maîtriser dans différents contextes et / ou à différents niveaux. Les référentiels de compétences sont des descriptifs d'actes, de performances observables qui sous-entendent un certain type de connaissances. Un référentiel de compétences se présente comme une typologie, une classification, un inventaire de compétences nécessaires à la réalisation d'activités. Ces compétences sont en général déclinées en capacités et en connaissances, indispensables ou nécessaires, pour réaliser avec succès une tâche, en l'occurrence langagière, dans une situation donnée ou dans différentes situations (ce qui demandera un transfert de compétence) : situation d'apprentissage, scolaire ou en entreprise, situation de travail ou situation de simulation. Les compétences sont définies au moyen de descripteurs signifiants qui permettront d'établir des jalons, des niveaux de compétence, des 'standards' auxquels on pourra se référer pour situer un individu par rapport aux compétences demandées, pour situer un apprenant dans son apprentissage (ou pour réussir dans un métier), mais aussi pour évaluer et valider les compétences acquises dans diverses situations" (Bourguignon *et al.*, 2005 : 462).

6 Dans le *Dictionnaire de Didactique des Langues* (Galisson & Coste, 1976), le contexte indique aussi l'entourage linguistique et non-linguistique dans lequel une unité est produite ou reçue (cf. Bloomfield, Miller, Prieto, etc.). "Sous le terme 'entourage', comme il n'y a pas de terme générique français rendant compte à la fois de la notion de contexte (voisinage linguistique) et de la notion de situation (voisinage extra-linguistique), nous proposons de créditer le terme entourage d'une acception nouvelle désignant l'ensemble des facteurs linguistiques (le contexte) et extralinguistiques (la situation) qui constituent le procès d'énonciation. (...) Pour permettre l'interprétation sémantique, il faut donc faire appel à la

notion d'entourage, c'est-à-dire connecter un discours (contexte) et une séquence d'images (situation fictive élocutant), de manière à les rendre complémentaires, donc interdépendants".

7 L'interface ici représentée date de janvier 2015. Il s'agit de l'interface proposée à l'occasion des tests pilotes mais non de l'interface définitive pour laquelle l'intervention d'un ergonomiste sera nécessaire en vue aussi de la cohérence entre les différents dispositifs produits par *Innovalangues*.

8 "Shohamy (1992) a proposé un modèle de retour diagnostique collaboratif dans lequel les tests fournissent des informations utiles pour améliorer l'enseignement et l'apprentissage. (...) Alderson (2006) conclut qu'avant de pouvoir faire un diagnostic du développement langagier, nous aurons besoin de mieux connaître la capacité en langue étrangère. Quand nous pourrions intégrer cette connaissance à l'évaluation, et la rendre utile aux apprenants en leur proposant des retours et suivis utiles et cohérents, alors le diagnostic aura vraiment pris sa place à l'interface entre l'apprentissage et l'évaluation".

9 "Les tests diagnostiques cherchent à identifier les domaines où un étudiant a encore besoin d'aide. Ces tests peuvent être assez généraux et montrer, par exemple, si un étudiant a besoin d'une aide ciblée sur une des quatre compétences langagières ; ils peuvent aussi être spécifiques et chercher à identifier, par exemple, les faiblesses grammaticales de l'étudiant."

Pour citer cet article

Référence

Cristiana Cervini et Marie-Pierre Jouannaud, « Ouvertures et tensions liées à la conception d'un système d'évaluation en langues, numérique, multilingue et en ligne, dans une perspective communicative et actionnelle », *Alsic* [En ligne], vol. 18, n° 2 | 2015, mis en ligne le 30 septembre 2015, Consulté le 15 janvier 2016. URL : <http://alsic.revues.org/2821> ; DOI : 10.4000/alsic.2821

À propos des auteurs

Cristiana Cervini

Cristiana Cervini est docteur en "Politique, éducation et formation linguistique-culturelle". Ingénieur de recherche dans le cadre du Projet ANR *Innovalangues* ("Idefi – Initiatives d'excellence en formations innovantes"), elle est responsable du lot Self – Système d'évaluation en langues à visée formative – et coordonne l'équipe d'italien impliquée dans le même projet. Elle est également chargée de l'enseignement en didactique des langues étrangères auprès du département Lilec (Langues, littératures et cultures modernes) de l'université de Bologne.

Affiliation : université de Bologne (Département Lilec) ; université Grenoble Alpes.

Courriel : cristiana.cervini@unibo.it

Toile : <http://www.unibo.it/docenti/cristiana.cervini>

Adresse : Via Cartoleria, 5 – 40124 Bologna, Italie; 1180 avenue centrale, 38400 Saint-Martin-d'Hères, France.

Les paragraphes 1., 3. (et sous-paragraphes), 4., 4.1 et 4.2 ont été rédigés par Cristiana Cervini.

Marie-Pierre Jouannaud

Marie-Pierre Jouannaud est PRAG d'anglais à l'université Stendhal à Grenoble. Dans le projet *Innovalangues*, elle coordonne l'équipe d'anglais de Self.

Affiliation : Université Grenoble Alpes.

Courriel : marie-pierre.jouannaud@u-grenoble3.fr

Adresse : 1180 avenue centrale, 38400 Saint-Martin-d'Hères, France.

Les paragraphes 2. (et sous-paragraphes), 4.3. et 5. ont été rédigés par Marie-Pierre Jouannaud.

Droits d'auteur

CC-by-nc-nd

Résumés

La volonté de s'inscrire dans une approche communicative, voire actionnelle, peut sembler incompatible avec les contraintes d'une évaluation à large échelle et ses exigences de rapidité et

d'autocorrection. De même, il est difficile de traduire le concept de tâche dans son sens primaire didactique ("toute visée actionnelle que l'acteur se représente comme devant parvenir à un résultat donné en fonction d'un problème à résoudre, d'une obligation à remplir, d'un but qu'on s'est fixé" ; *CECRL*, 2001 : 16) dans un contexte d'évaluation numérique en ligne. Cela ouvre plusieurs questions concernant l'opérationnalisation du construit du test. L'enjeu principal concerne la traduction du principe d'authenticité, soit situationnelle, soit interactionnelle, dans un test visant trois compétences : compréhension de l'oral, compréhension de l'écrit, expression écrite courte (Bachman & Palmer, 1996 ; 2010). Dans cette contribution nous essaierons d'expliquer comment cette tension a été gérée dans la démarche de rédaction des tâches d'évaluation sur les trois compétences, conduite dans le cadre d'un travail de création d'un "Système d'évaluation en langues à visée formative (Self)" multilingue et administré en ligne. Self s'inscrit dans le cadre du projet *Innovalangues* porté par le service Lansad de l'université Stendhal Grenoble 3 (Masperi, 2012).

Designing an online foreign language assessment tool within a communicative and action-based framework: opportunities and challenges

In this paper, we grapple with the problem of what it means for a placement test to be communicative or even "action-oriented" (*CEFR*, 2001). Communicative tasks ("any purposeful action considered by an individual as necessary in order to achieve a given result, an obligation to fulfil or an objective to be achieved", *ibid.*: 10) are especially difficult to translate into short, self-correcting assessment tasks. We explain how this difficulty was dealt with when we operationalized the construct of a multilingual online formative assessment tool (Self, "Système d'évaluation en langues à visée formative"), using among others the concepts of situational and interactional authenticity (Bachman & Palmer, 1996; 2010). Self is being developed as part of the *Innovalangues* project at Stendhal University, Grenoble, France (Masperi, 2012).

Entrées d'index

Mots-clés : évaluation, numérique

Keywords : evaluation, ICT

Thématiques : évaluation

Rubriques : Pratique et recherche

Notes de la rédaction

Date de réception de la première version de l'article : 15 janvier 2015 ; date d'acceptation : 14 avril 2015 ; date de réception de la version définitive : 2 juin 2015 ; date de mise en ligne : septembre 2015.