

Analyses des classes et des transitions latentes : Des outils pour documenter le parcours de développement des compétences

Léon Harvey ,^a

^a Département des sciences de l'éducation, Université du Québec à Rimouski

Abstract ■ Our proposal is that the process of validating the construct of a competency inferred from an observation grid raises a multivariate and multi-dimensional latent classification problem. The analysis models a given competency based on the identification of discrete classes using latent classification (LCA) and latent transition (LTA) models. Then, the approach is illustrated in the context of the growth of the morphosyntax competency of learners of French as a second language. Advantages and limits of the analysis are outlined.

Résumé ■ Notre proposition est que le processus de validation du construit d'une compétence inférée à partir d'une grille d'observation soulève un problème multivarié et multidimensionnel de classification latente. L'analyse modélise une compétence à partir de l'identification de stades discrets et de transitions d'un stade à un autre par le biais de modèles de classes latentes (LCA) et de transitions latentes (LTA). L'approche est par la suite illustrée dans le contexte du développement de la compétence morphosyntaxique chez des apprenants de français langue seconde. Les limites et les avantages de l'analyse sont discutés.

Keywords ■ Latent class, Latent transition, Competency, Modelling, Morphosyntax // classe latente, transition latente, compétence, modélisation, morphosyntaxe

 Leon.Harvey@uqar.ca

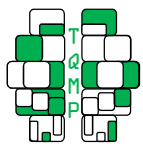
Introduction

Depuis quelques années, la notion de compétence est au centre de la conception de plusieurs programmes de formation à tous les ordres d'enseignement. Des progrès sur la définition de la notion de compétence sont réalisés en lien avec la modélisation de ce concept dans des contextes éducatifs (Klieme, Hartig, & Rauch, 2008). Dans les écrits francophones, Ayotte-Beaudet (2013) analyse une vingtaine de définitions dont émergent différentes dimensions; l'aspect central est la mobilisation par un individu d'un ensemble de ressources internes et externes afin de réaliser des tâches authentiques et signifiantes issues de familles de situations. Le présent article s'intéresse en particulier aux aspects de ce concept qui sont liés à sa mesure. En particulier, l'une des propriétés les plus intéressantes est qu'une compétence est non directement observable et est, par conséquent, considérée comme un construit latent. Cependant, sa structure peut être inférée à partir d'indicateurs qui, eux, sont observables. De cette propriété découle l'une des plus grandes difficultés de la notion de compétence: définir des outils qui permettent d'inférer, à partir des indicateurs sélectionnés, l'état dans lequel se trouve cette compétence chez un apprenant donné dans une famille

de situations. L'utilisation de modèles psychométriques est ici essentielle pour valider ces inférences. « *Psychometric models are the link between theoretical construct and the results of empirical assessments* » (Klieme et al., 2008, p. 12).

Or, les compétences peuvent être mesurées dans différents contextes éducatifs. Elles peuvent être mesurées lors de tests nationaux à grande échelle (p. ex., PISA), lors d'évaluation de programmes ou d'institutions spécifiques, en recherche et lors de la qualification d'individus par un ou plusieurs enseignants en contexte de classe ou en situation authentique (Klieme et al., 2008). Dans certaines de ces situations, des échelles standards peuvent être utilisées; les modèles psychométriques tels que la théorie de réponses aux items (Blais, Laurier, & Rousseau, 2009) ou l'analyse factorielle peuvent alors être utilisés pour valider le construit en modélisant la compétence d'un apprenant. C'est le cas, notamment, lors de la passation de tests à grande échelle.

Par contre, dans d'autres cas, l'observation en situation authentique ou en classe fait appel à des grilles ou à des outils non standardisés. Le plus souvent, il s'agit d'échelles descriptives qui définissent de



manière narrative les standards nécessaires pour atteindre différents niveaux d'une compétence (Bélaïr, 1999 citée par Campanale & Raïche, 2008). Dans ces situations, l'addition des indicateurs pour former un résultat global pour chaque individu n'a généralement pas de sens. Les indicateurs sont souvent de nature très différente et sont parfois même très peu corrélés. Il faut alors conserver intacte l'information des indicateurs afin de porter un jugement diagnostique sur la configuration particulière (ou le profil de réponses) qui émerge de l'ensemble (Bloom, 1942¹; Boscardin, 2012). L'utilisation de grille d'observation répond à cette prérogative. De telles grilles de compétence sont d'ailleurs largement répandues dans le monde de l'éducation.

Au départ, l'analyse théorique de la compétence et son explicitation sont essentielles; elles permettent de dégager des niveaux de compétence qui se distinguent les uns des autres en fonction des indicateurs sélectionnés, produisant ainsi une grille d'observation descriptive (Blais et al., 2009; Côté, 2014; Scallon, 2004). En soi, une démarche théorique de cette nature est inhérente au développement de tout instrument de mesure. Par contre, comment déterminer si la grille d'observation théorique est valide et fiable?

La fiabilité réfère à la précision et à la consistance interne d'un instrument, c'est-à-dire la proportion de variance attribuable au score vrai par rapport au total de la variance (Brown, 2006; Brunner, Nagy, & Wilhelm, 2012; Raykov, 2004). Elle permet de déterminer jusqu'à quel point un score est affecté par la présence de sources d'erreur. Certaines solutions pour attester de la fiabilité d'une grille sont connues. La fiabilité d'une grille peut être obtenue, par exemple, par une étude de l'accord interjuges (Hallgren, 2012); la grille est alors mise à l'essai par différents juges qui l'expérimentent auprès d'un même échantillon d'individus. Par ailleurs, la classification obtenue lors d'un premier test devant être relativement stable dans le temps, la fiabilité d'une grille de classification peut être obtenue par la procédure de test-retest. Cependant, ces preuves liées à la fiabilité d'une grille ne permettent pas d'apprécier pleinement l'adéquation entre cette grille et les observations qu'elle génère. En soi, une grille est un modèle idéal de développement de la compétence qu'il faut infirmer ou confirmer à partir d'observations en situations authentiques.

Ainsi, certaines questions en lien avec la validation

¹ Merci à L. Laurencelle (2014) pour cette référence.

du construit mesuré par ces grilles restent intuitives. Il faut préciser que la validité est un concept unitaire. C'est un jugement qui est porté sur l'ensemble des preuves accumulées à propos d'un instrument et qui détermine si les interprétations réalisées à partir de celui-ci sont conformes à l'intention visée (American Educational Research Association, 1999; Laveault & Grégoire, 2002). Le concept de validité de construit (Zaltman, Pinson, & Angelmar, 1973) réfère au degré auquel une opérationnalisation (un instrument) permet de mesurer le concept qu'elle est supposée représenter. La littérature (voir Laveault & Grégoire, 2002) utilise aussi l'expression validité de concept comme synonyme et considère que les procédures associées donnent un sens aux scores obtenus au test en lien avec le modèle théorique que l'on souhaite évaluer. Linn et Miller (2005) spécifient qu'il s'agit de déterminer le degré auquel la performance à un instrument peut être interprétée en termes d'un ou de plusieurs construits. Ce degré est établi à partir de l'adéquation entre les structures théorique et empirique.

Par conséquent, quelle est l'adéquation entre le construit théorique d'une grille d'observation systématique et les résultats issus de son application dans des situations authentiques? L'utilisation de la grille dans ces contextes permet-elle de confirmer ou d'infirmer le modèle théorique (le construit) de la compétence? Cette question est importante, car d'autres questions y sont reliées. Les indicateurs qui composent la grille sont-ils discriminants? Le nombre de niveaux de la grille est-il optimal? La grille permet-elle de porter des jugements valides sur les apprenants observés?

En particulier, la distribution des indicateurs est rarement communiquée et, de ce fait, un véritable processus de modélisation de la compétence n'est alors pas possible. En fait, de nouveaux modèles psychométriques sont nécessaires (Klieme et al., 2008) pour valider et pour modéliser le processus de développement d'une compétence au-delà de la grille théorique qui est produite. Ce processus de validation permet notamment de vérifier si le nombre de niveaux de la grille est optimal et si les indicateurs sélectionnés sont utiles.

Devant la complexité du concept de compétence, ces modèles doivent être parcimonieux, posséder une base théorique forte et permettre de réaliser les inférences quant à la maîtrise des compétences en situations authentiques. De plus, ces modèles doivent être multidimensionnels et permettre la mesure de

Tableau 1. Grille de compétence fictive à quatre niveaux et trois indicateurs

Niveau	Indicateur 1	Indicateur 2	Indicateur 3
A			✓
B		✓	✓
C	✓	✓	✓
D	✓	✓	✓

plusieurs variables latentes. Ces variables représentent alors des niveaux de développement de la compétence qui sont qualitativement différents les uns des autres. Les modèles doivent également être multivariés et prendre en considération les inférences basées sur plusieurs indicateurs de performance qui sont les éléments constitutants de la grille d'observation. Finalement, ils doivent être souples quant à la nature des indicateurs considérés afin de concilier les indicateurs de nature binaire, multinominale ou continue, car, comme le suggère la littérature (Côté, 2014; Scallon, 2004), il existe différents types d'échelles descriptives.

L'objectif du présent article est de proposer deux outils qui permettent de modéliser le développement des compétences en situations authentiques à partir de grilles d'observation. Notre proposition est que la validation du construit d'une compétence observée à partir d'une grille d'observation représente un problème multivarié de classification latente (LCA) ou d'analyse des transitions latentes (LTA).

Notre démarche suggère que les modèles LCA et LTA permettent la validation du construit mesuré par des grilles utilisées dans le cadre de l'observation d'une compétence. De plus, lorsque des modèles LTA sont utilisés, ils permettent l'obtention d'itinéraires de développement de la compétence.

L'origine mathématique des deux outils considérés est similaire; il s'agit de modèles à mixtures de distribution. La différence est que les modèles par classification latente s'appliquent lorsque des données issues de protocoles transversaux sont recueillies. En contrepartie, les modèles en transitions latentes requièrent des données issues de protocoles longitudinaux et sont avantageux lorsque les individus changent de classe pendant la période d'observation. Des modèles LCA peuvent à l'occasion être utilisés avec des protocoles longitudinaux lorsque les individus ou

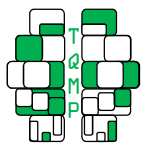
certaines classes d'individus demeurent dans la même classe durant toute la période d'observation.

Notre démarche sera d'abord illustrée à partir d'une compétence fictive. Puis, elle sera illustrée dans le contexte du développement de la compétence morphosyntaxique chez des apprenants en français langue seconde. Le choix de ce contexte est motivé par le fait qu'il existe des travaux détaillés sur cette compétence et des corpus de textes longitudinaux (Ågren, 2008) qui vont permettre d'illustrer nos propos à l'aide d'observations obtenues à partir de situations authentiques d'écriture.

La modélisation des compétences à l'aide de modèles latents

La modélisation des compétences à partir de modèles par classes ou transitions latentes est relativement nouvelle et en émergence. Les modèles à transitions latentes ont été proposés dans les années 1950 pour modéliser l'apprentissage (Miller, 1952)², tandis que Vermunt et Magidson (2002) considèrent que l'analyse par classes latentes remonte également aux années 1950 à 1970. Ce n'est que récemment, avec la prolifération des ordinateurs, que ces modèles ont pu être utilisés dans des contextes nombreux et variés. Les écrits récents rapportent différents travaux sur la modélisation de la compétence des étudiants à l'aide de modèles développementaux latents (Lissitz, 2006). Plus spécifiquement, les modèles par classes latentes sont reconnus comme étant des modèles généraux pouvant aider à porter un diagnostic sur les profils cognitifs des élèves à la suite d'une épreuve donnée (Boscardin, 2012; von Davier, DiBello, & Yamamoto, 2008; von Davier, 2008). De même, Kaplan et Sweetman (2006) ont modélisé la compétence en mathématiques à l'aide de modèles de transition latente. En santé, Harvey (2011) a utilisé un modèle LTA afin de modéliser la compétence de stagiaires en soins cliniques infirmiers à partir d'observations réalisées par la superviseure en situation authentique. De façon générale, les modèles LTA sont très utilisés en psychologie (Visser, Raijmakers, & Molenaar, 2002; Wickens, 1982) pour modéliser différents comportements. Des modèles plus complexes sont aussi développés en sciences cognitives pour modéliser des compétences comportant des séquences d'actions nécessitant une prise de décision

² Dans le présent texte, modèle par transitions latentes et modèle de Markov à états cachés sont considérés synonymes.



(Fu & Anderson, 2006; Littman, 2010; Niv, 2009).

À titre illustratif, le tableau 1 présente une grille d'observation pour une compétence fictive. La grille contient trois indicateurs et quatre niveaux de compétence. Cette grille est typique des grilles qui sont rapportées dans les écrits. À ce stade-ci, la décision d'inclure trois indicateurs et quatre niveaux est relativement arbitraire. Comme il a été mentionné précédemment, cette décision relève généralement de considérations théoriques et repose sur l'avis d'experts ou de praticiens. Le choix actuel n'est qu'illustratif.

Ainsi, cette grille suggère que le premier stade de développement de la compétence, le stade A, est caractérisé par la présence unique de l'indicateur 3. Les indicateurs 1 et 2 ne sont pas observés. Aux fins de la modélisation, une valeur nulle (0) est attribuée à ces deux indicateurs à ce niveau. Normalement, ce premier niveau correspond à la compétence d'un novice. À l'opposé, le niveau de développement D est le plus avancé de notre compétence fictive et est caractérisé par un profil où sont observés les trois indicateurs de performance. Ce stade correspond au niveau de maîtrise de la compétence observé chez des experts. Les niveaux B et C sont des stades intermédiaires. Les niveaux C et D semblent identiques; en fait, il s'agit d'un cas hypothétique où l'ensemble des indicateurs sont observés aux deux niveaux mais s'avèrent plus saillants en D qu'en C.

Sur le plan de la modélisation, les niveaux A, B, C et D représentent des classes latentes qui se distinguent par les valeurs obtenues dans chacun des indicateurs. Les appellations stades, classes ou niveaux de compétence seront considérées ici comme équivalentes. Ainsi, les individus qui sont dans la classe latente D sont qualitativement différents de ceux qui sont dans les classes A, B et ainsi de suite. Les indicateurs 1, 2 et 3 sont, dans cette grille, de nature qualitative (binaire). Lorsque des indicateurs continus sont utilisés, les indicateurs sont à leur plus bas niveau au stade A. À l'opposé, au stade D, ils sont à leur valeur maximale.

Dans cette perspective, la grille de compétence peut être modélisée à partir de mélanges de distributions. Ces modèles sont de plus en plus utilisés lorsque des distributions d'individus issus de sous-populations hétérogènes sont observées. C'est notamment le cas lorsqu'une grille de compétence est construite afin de décrire la progression d'individus possédant des niveaux de compétence différents. Ces individus proviennent ainsi de sous-populations qui diffèrent

quant à leur niveau de compétence. Ces sous-populations hétérogènes forment alors ce que l'on appelle une mixture de distribution.

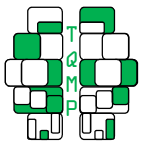
Les modèles de classification latente ne sont pas les seuls outils pour classer les individus, mais ils gagnent en popularité par rapport à l'approche classique (*cluster analysis*). Vermunt et Magidson (2002) ainsi que Boscardin (2012) discutent des avantages des modèles LCA par rapport à cette approche classique. Le premier avantage est que l'approche par classification latente est basée sur la modélisation des données. Le modèle considère que les données sont générées par une mixture de distribution de probabilités latentes. L'estimation des paramètres de distribution est généralement obtenue par la maximisation de la vraisemblance du modèle. La modélisation est très flexible et permet des mélanges simples ou complexes, et des restrictions peuvent être imposées sur les paramètres afin de vérifier la validité de certains paramètres. De plus, des indicateurs dichotomiques, multinominaux ou continus peuvent être modélisés. Par ailleurs, les indicateurs d'échelles différentes n'ont pas besoin d'être normalisés. Un autre avantage est que les modèles de classification latente sont des modèles probabilistes, où le résultat de la classification contient de l'incertitude et s'accompagne de la probabilité d'appartenir à chacune des classes. Il est alors possible d'inspecter ces probabilités; dans un bon modèle, l'incertitude associée à la classification doit être minimale et la probabilité pour un apprenant d'appartenir à sa classe doit être forte par rapport aux autres probabilités. En raison de ces avantages, il n'est pas surprenant que la classification latente soit préférée à l'approche classique.

Le développement mathématique des modèles LCA et LTA

Dans le présent article, les modèles de classification latente permettent d'identifier les caractéristiques des niveaux de développement d'une compétence. L'objectif d'un modèle à mélanges de distribution est d'explicitier les densités de probabilité des indicateurs, à partir de la somme pondérée de distributions de densité associées à un certain nombre de classes latentes. Mathématiquement,

$$P(Y_i) = \sum_j \pi_j P^j(Y_i)$$

où $[\pi_1, \pi_2, \dots, \pi_j]$ constitue un vecteur π dont les



éléments sont la probabilité d'observer la classe latente j ; et où $P^j(Y_i)$ est la probabilité que la valeur Y_i appartienne à la classe j .

Pour une variable dichotomique (0,1) ou multinominale, une matrice d'observation B permet de présenter les probabilités conditionnelles qu'une observation y_i appartienne à la classe Y_i .

$$B = \begin{bmatrix} y_1 & y_2 \\ y_3 & y_4 \end{bmatrix} = p(y/Y_i)$$

Pour une variable continue, $p(y/Y_i)$ est déterminée à partir de la densité de distribution de la variable latente. Les variables continues sont présumées être normalement distribuées à l'intérieur de chaque classe latente; en fait, le problème consiste à connaître les paramètres de cette distribution, généralement une distribution multivariée normale. Les paramètres recherchés sont la moyenne et la matrice de variance-covariance de la distribution multivariée. Généralement, des classes latentes qui diffèrent quant à la moyenne sont recherchées, mais des différences quant aux autres paramètres peuvent exister. Ainsi, la matrice de variance-covariance n'a pas à être homogène à toutes les classes; elle peut être spécifique de chaque classe.

Les modèles par transition latente (LTA) ajoutent à la modélisation les taux de transition entre les classes. Un modèle LTA précise l'évolution temporelle des classes dans le temps. Il assume qu'il existe une probabilité qu'un ou plusieurs individus changeront de classe pendant la période d'observation. Une matrice A présente les probabilités conditionnelles de passer de la classe j_1 à j_2 et vice-versa. Le vecteur π exprime alors les probabilités initiales d'appartenir à une classe ou à une autre.

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

Les modèles LCA et LTA s'appliquent lorsque le choix d'une variable latente discrète s'impose, ce qui est le cas lorsque l'objectif de la modélisation est l'identification de stades (ou classes)³.

³ Il est possible de modéliser le développement d'une compétence à partir de variables latentes continues plutôt que discrètes. Cependant, considérant que notre objectif est de se doter d'outils pour modéliser les niveaux de développement des compétences, les

Comme mentionné antérieurement, ces modèles peuvent être utilisés à partir de protocoles transversaux (LTA) ou longitudinaux (LCA et LTA). Notons qu'un modèle LCA peut être exceptionnellement utilisé avec un protocole longitudinal lorsque l'évolution des sujets d'un temps à l'autre est nulle. En fait, il peut exister, dans un même modèle longitudinal, une mixture de classes latentes et de transitions latentes (Muthén, 2008). Un tel modèle serait l'indication qu'il existe des classes d'individus stables qui n'évoluent pas et des classes d'individus dont la compétence évolue durant la période d'observation.

Le nombre de niveaux du modèle est important; il donne une indication de la capacité de la grille à classer les individus. Il est difficile de déterminer a priori le nombre de niveaux qu'une grille doit inclure. Laurencelle (2014) mentionne que plus un instrument est fiable, plus il peut classer les individus en de nombreuses classes. Moins il est fiable, moins le nombre de classes sera élevé. Ainsi, à titre comparatif, Laurencelle estime qu'un instrument dont les résultats sont normalement distribués et qui possède une bonne fiabilité (fiabilité = 0,90) peut classer les individus en neuf catégories distinctes, et ce, en considérant que la probabilité d'un individu d'appartenir à sa catégorie soit toujours de 50 %. Par contre, si la fiabilité n'est que de 0,50, le même instrument verra sa capacité discriminante passer à environ quatre catégories.

Dans le cadre de modèles LCA et LTA, l'identification du nombre de niveaux n'est pas basée sur une étude de la fiabilité. Elle est faite *a posteriori* par la génération de plusieurs modèles; des modèles simples à un ou deux niveaux sont alors comparés à des modèles plus complexes (trois niveaux et plus). Le modèle qui possède la meilleure adéquation aux données est sélectionné. Des indices d'adéquation tels que les indices AIC et BIC sont utilisés. Ces indices sont parcimonieux et tiennent compte de l'ajout de paramètres supplémentaires qui accompagnent les modèles plus complexes. Pour une définition plus exhaustive de ces indices, voir Hélie (2006).

L'utilisation des modèles LTA et LCA est sujette à des conditions d'application (von Davier et al., 2008). Ces conditions sont l'indépendance locale des observations étant donné l'appartenance à une classe, et la présomption que les classes latentes sont exhaustives et disjonctives. Il est également attendu

modèles à variables latentes discrètes sont une option naturelle.

Tableau 2 ■ Fluidité et marquage nominal du pluriel avec les morphèmes « s » et « x » (extrait d'Ågren, 2008) – étude transversale

Stade	Fluidité	Marquage du pluriel		
		Moyenne	Écart-type	Test
A	130	79 %	29 %	12,2 %
B	202	87 %	16 %	21,4 %
C	276	91 %	12 %	37,5 %
D	406	97 %	5 %	51,5 %

que les probabilités conditionnelles entre les classes (matrice B) soient distinctes (la distinction). L'indépendance locale stipule qu'étant donné la classe latente, les variables observées sont indépendantes. La présomption des classes latentes exhaustives et disjonctives garantit que chaque apprenant est un membre d'une seule et unique classe latente. Cette présomption nécessite qu'il existe suffisamment d'information à partir des indicateurs pour classifier chacun des apprenants dans une classe donnée avec un haut taux de probabilité, et ce, dans la perspective où le modèle est adéquat. Finalement, la distinction suggère que chaque classe latente diffère quant au profil de réponses qui la caractérise. Dans la perspective où ces conditions d'application ne sont pas satisfaites, une révision du modèle et, par conséquent, de la grille d'observation s'avère nécessaire.

Un exemple : la compétence morphosyntaxique

La morphosyntaxe est définie comme étant « l'ensemble des structures qui permettent de construire grammaticalement un énoncé » (Parisse, 2009, p. 7). L'étude de l'acquisition de la morphosyntaxe chez des apprenants nécessite la spécification d'un grand nombre d'indicateurs de performance. Toujours selon cet auteur, la morphosyntaxe porte aussi bien sur les formes des mots, les flexions régulières et irrégulières, les variantes irrégulières de certains noms et verbes, l'agencement des marques syntaxiques autour du nom (déterminants, etc.), du verbe (pronoms, etc.), de l'adjectif, de l'adverbe, que sur l'organisation des mots et groupes de mots dans un énoncé ou une phrase.

Ågren (2008), à l'instar de Bartning et Schlyter (2004), conçoit le développement de cette compétence par la succession de stades distincts. À chaque stade, des indicateurs de performance sont identifiés et des valeurs caractéristiques de chacun des stades sont déterminées. Ågren a notamment observé quatre des six stades identifiés par Bartning et Schlyter.

À des fins d'illustration, nous nous limiterons à modéliser simultanément trois indicateurs. Une modélisation exhaustive de la compétence nécessiterait la prise en compte d'un plus grand nombre d'indicateurs. Néanmoins, les indicateurs choisis sont justifiés par le fait qu'ils permettent de tracer un portrait global de la compétence. Ces indicateurs sont la fluidité du scripteur, sa précision globale ainsi que le marquage nominal du pluriel avec les morphèmes « s » et « x ».

Le tableau 2 extrait d'Ågren (2008) montre que la fluidité et le marquage nominal du pluriel fluctuent bien d'un stade à l'autre de la compétence morphosyntaxique. Ces statistiques proviennent d'une étude transversale menée auprès d'étudiants en français langue seconde en Suède. Un test de connaissances complémentaire permet de montrer qu'il y a évolution des connaissances en parallèle avec le développement des stades de la compétence.

Le tableau 2 constitue une grille de compétence au même titre que la grille du tableau 1. La différence est que les indicateurs sont de type continu plutôt que binaire.

Le présent article propose d'inférer les mêmes quatre stades et la même classification en utilisant des modèles LCA et LTA. La modélisation sera faite à partir d'un échantillon longitudinal restreint et en tenant compte des trois indicateurs. Le modèle le plus simple considéré est un modèle à 2 classes latentes (LCA-2). Ce modèle est comparé à un modèle à 4 classes latentes (LCA-4) afin de mieux comparer les stades obtenus à ceux de l'étude transversale rapportée par Ågren (2008). Il est également comparé à un modèle LTA à 2 niveaux (LTA-2) afin de déterminer si des transitions d'un niveau de compétence à l'autre se sont produites durant l'année scolaire. Tel qu'il a déjà été mentionné, le modèle LTA permet d'estimer les taux de transition d'un niveau à l'autre. Pour des fins de comparaisons supplémentaires, le modèle à 4 classes sera comparé avec les classifications rapportées avec l'analyse d'Ågren.

Description du corpus longitudinal

Comme mentionné auparavant, le corpus de textes provient d'Ågren (2008) et est disponible sur Internet. Le choix de ce corpus est justifié par sa disponibilité et par le fait que les niveaux de la compétence morphosyntaxique des apprenants ayant produit les textes ont déjà été déterminés à partir de l'application de la grille de Bartning et Schlyter (2004). Une

Tableau 3 ■ Description des deux classes du modèle LCA-2

Indicateur	Classe 1		Classe 2	
	Moyenne	Écart-type	Moyenne	Écart-type
Fluidité	159,33	54,44	396,92	125,62
Précision	75,88 %	6,98 %	87,68 %	3,37 %
Marquage du pluriel	74,30 %	12,90 %	89,80 %	7,50 %
	$\pi_1 = 60 \%$		$\pi_2 = 40 \%$	

Tableau 4 ■ Distribution des textes selon les probabilités d'appartenir à sa classe propre

(modèle LCA-2)					
Classe	50-70 %	70-79 %	80-89 %	90-99 %	100 %
1	1	0	1	2	23
2	1	2	0	1	14

comparaison de cette classification avec celle de l'analyse en classes latentes sera alors possible. Il s'agit d'une forme de validité relative à un critère externe qui est connu, ce qui représente un avantage à des fins d'illustration. Un lien entre ce corpus et un critère concomitant a aussi été présenté par Ågren, ce qui renforce l'ensemble des preuves disponibles pour ce corpus.

Quinze élèves ont produit quatre textes répartis durant l'année scolaire. Ces quinze élèves se situent dans l'un ou l'autre des stades de compétence (A à D). Les mises en situation pour chacun des textes sont différentes, mais les indicateurs de compétence sont comparables d'un texte à l'autre. Ces élèves ont de une à six années d'études du français. Ils proviennent donc a priori de classes différentes.

La procédure pour déterminer la valeur de chacun des indicateurs est la suivante. La *fluidité* de chacun des textes est déterminée par le nombre total de mots écrits. La *précision* est le taux de réussite, soit 1 - le pourcentage d'erreurs. Tous les types d'erreurs sont compilés, qu'ils soient grammaticaux, syntaxiques ou orthographiques. Les indicateurs *fluidité* et *précision* sont obtenus à partir de l'application du logiciel de correction grammaticale *Antidote 6.1* sur chacun des textes. La précision est d'ailleurs l'indicateur utilisé par *Antidote* pour porter un diagnostic sur la compétence du scripteur. Quant au *marquage du pluriel*, cet indicateur est obtenu à partir d'une compilation globale des données d'Ågren (2008) quant au taux de réussite

dans les différents types d'accord du pluriel. Comme la compilation effectuée à partir d'Ågren n'est possible que pour trois des quatre textes, seulement trois textes par élève sont utilisés.

La base de données est par conséquent constituée de 135 données (15 élèves x 3 textes x 3 indicateurs). Considérant la taille restreinte de ce corpus, l'analyse se limitera à des modèles simples avec peu de paramètres à estimer. En fait, les modèles considérés ont entre 13 et 27 paramètres, ce qui représente un ratio se situant entre 5 et 10 données par paramètre. Ce ratio est considéré comme minimal pour ce type de modèle, mais suffisant dans une perspective d'illustration. Évidemment, un corpus de textes plus considérable serait souhaitable dans une perspective de généralisation⁴.

Les indicateurs sont moyennement corrélés. Les coefficients de corrélation entre la *fluidité* et la *précision* et la *fluidité* et le *marquage du pluriel* sont respectivement de 0,55 et 0,61. Il est de 0,52 entre la *précision* et le *marquage du pluriel*.

Résultats

La figure 1 présente la distribution de fréquence de la *fluidité*. Cette distribution est asymétrique et ne peut être considérée comme une distribution normale. La figure 2 est le résultat de la modélisation à deux classes latentes. La première classe se distingue par des productions écrites ayant une moyenne de 159,33 mots, par rapport à 396,92 mots pour la seconde classe (tableau 3). Les productions écrites de cette première classe sont donc moins fluides que celles de la classe plus avancée. Soixante pour cent des élèves se situent dans cette première classe, alors que 40 % se situent dans la seconde. L'écart-type de la première classe est d'environ 54,44 mots, par rapport à 125,62 mots pour la seconde.

Les productions écrites de la première classe se distinguent également par une précision moindre. La précision moyenne est de 75,88 %, par rapport à 87,68 % pour la seconde classe. Finalement, les deux classes se distinguent par des taux différents quant à l'accord du pluriel, passant de 74,3 % de réussite à 89,80 %. Le pourcentage d'élèves catégorisé dans la première classe est de 60 %, par rapport à 40 % pour la

⁴ Les solutions des modèles LTA et LCA ont été produites par depmix (Visser et al., 2002) en R. L'annexe I présente le détail de la programmation.

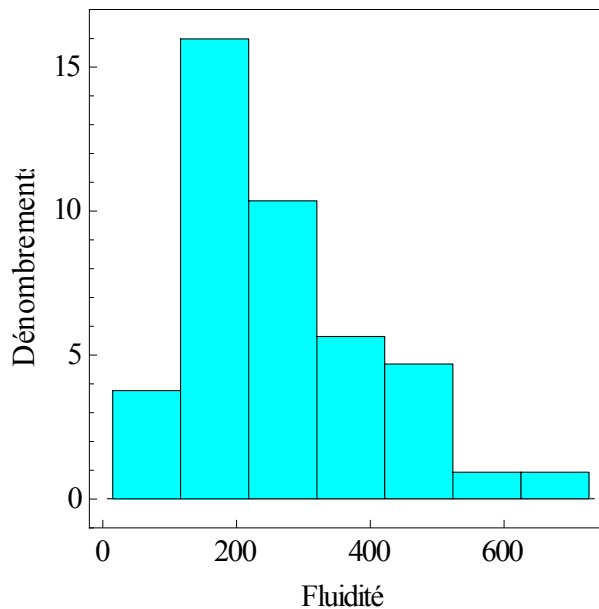


Figure 1 ■ Distribution de fréquences pour l'indicateur *Fluidité*.

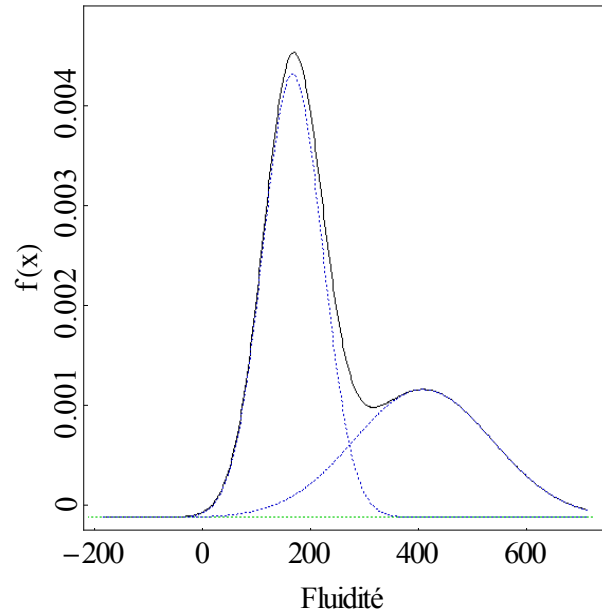


Figure 2 ■ Modélisation à deux classes (modèle lca2) pour l'indicateur *Fluidité*.

seconde classe.

Les deux classes produites sont donc très distinctes; les profils générés diffèrent quant aux valeurs des indicateurs. En ce sens, nous pouvons considérer que tous les indicateurs s'avèrent utiles dans la classification.

L'inspection des probabilités d'appartenir à une classe suggère que la classification est également disjonctive. Les probabilités garantissent que tous les apprenants sont classifiés et que chaque apprenant est un membre d'une seule et unique classe latente. Le tableau 4 présente les probabilités d'appartenir à la classe 1 ou 2 pour chacun des 45 textes; 27 textes ont été classifiés dans la classe 1, tandis que 18 textes l'ont été pour la classe 2.

Comparaison de modèles

À titre comparatif, le modèle à 2 classes latentes (LCA-2) a été comparé aux modèles alternatifs LTA-2 et LCA-4 (voir le tableau 5).

La première comparaison est faite avec le modèle LTA-2. Ce modèle permet d'évaluer si, durant l'année en cours, les élèves ont évolué et s'ils ont progressé d'un niveau de compétence à un autre. Or, le taux de progression obtenu par le modèle est de 0,07 (7 %) et ce paramètre s'avère ne pas être différent de zéro. Dans un tel cas, un modèle plus simple peut être privilégié. Ici, le modèle LCA-2 est un modèle plus simple et, effectivement, les indices d'adéquation AIC et BIC

indiquent que ce modèle est plus parcimonieux que le modèle LTA-2. Les propriétés du modèle LTA-2 ne seront donc pas examinées plus en profondeur.

L'examen du modèle LCA-4 est cependant intéressant; l'indice AIC considère que ce modèle est celui qui décrit le mieux les données, ce qui rejoint l'analyse d'Ågren (2008). Notons que l'indice BIC favorise un modèle plus parcimonieux, soit le modèle LCA-2. Il y a donc divergence entre les deux indices quant au modèle à favoriser. De par son intérêt théorique, le modèle LCA-4 sera donc détaillé ici. L'avantage de ce modèle à quatre niveaux est qu'il est plus exhaustif que le modèle à deux niveaux.

Le tableau 6 présente le résultat de la modélisation pour chacun des indicateurs en fonction des niveaux de compétence.

Le tableau 7 présente une comparaison des moyennes de chacun des stades identifiés par Ågren et celles obtenues à partir du modèle LCA-4 de la présente étude. Dans ce tableau, la première lettre du tableau identifie le stade de la personne catégorisée (Andréa => stade A, Bror => stade B, etc.).

Bien que le modèle LCA-4 soit basé sur un nombre très restreint de textes, la catégorisation qu'il réalise s'apparente à celle qui est rapportée par Ågren (2008). Elle présente cependant certaines différences. Il est possible de constater que le stade A du modèle LCA-4 fusionne les stades A et B de l'étude d'Ågren. Les élèves catégorisés totalisent 53 % de l'échantillon et

Tableau 5 ■ Indices d'adéquation des trois modèles considérés

Modèle	Vraisemblance	dl	AIC	BIC
LCA-2	-366,64	13	759,28	782,76
LTA-2	-365,13	15	760,23	787,37
LCA-4	-351,70	27	757,40	806,19

regroupent les élèves dont le prénom débute par A, B et même un C (Conny). Le stade B n'a pas émergé dans le modèle LCA-4. Par ailleurs, la fluidité du stade C est comparable dans les deux études (276 mots). Ce stade regroupe principalement les élèves dont le prénom débute par C et un D (Deborah). Quant au stade D, il regroupe dans les deux études deux apprenants communs. Finalement, le modèle LCA-4 identifie un stade supplémentaire (E) pour Daisy, qui n'est pas identifié par l'étude d'Ågren.

La grille modélisée génère une classification avec un taux élevé de certitude. Le tableau 8 présente pour chaque stade un sommaire des probabilités d'appartenir à une classe donnée pour chacun des textes classifiés. Ainsi, sur les 45 textes qui ont été classifiés, 80 % (soit 36/45) l'ont été avec une probabilité supérieure ou égale à 99 %; 11 % (5/45) l'ont été avec une probabilité qui se situe entre 0,90 et 0,989; et seulement 8,8 % l'ont été avec une probabilité d'appartenance inférieure à 90 %. Un seul texte a été classifié avec une probabilité d'appartenir à cette catégorie qui est faible (inférieure à 50 %) et qui représente une erreur de classification du modèle. Il s'agit d'un léger manque de disjonction de ce modèle. Notons que ce type d'information n'était pas disponible au départ dans l'étude d'Ågren (2008). Il s'agit donc d'une contribution importante de l'analyse par classes latentes.

Ainsi, dans l'ensemble, la probabilité qu'un texte soit mal classifié reste faible (1/45). La grille permet de poser, avec une forte certitude, un jugement précis sur la compétence des individus qui ont produit les textes.

Cette comparaison illustre qu'un modèle à classes latentes permet de faire émerger avec un haut niveau de certitude des stades de compétence relativement similaires à ceux identifiés dans l'étude d'Ågren, et ce, malgré des différences d'échantillonnage et de méthodologie.

Tableau 6 ■ Moyennes des indicateurs (fluidité, précision et marquage du pluriel) en fonction des niveaux de développement de la compétence

Niveau	Fluidité	Précision	Marquage du pluriel	π
A	154,9	0,75 %	0,74 %	53,4 %
B	275,8	0,86 %	0,85 %	26,6 %
C	413,9	0,89 %	0,90 %	13,3 %
D	538,4	0,89 %	0,97 %	6,7 %

Conclusion

La présente étude démontre qu'il est possible de modéliser le développement des compétences à partir de modèles à classes ou à transitions latentes. La modélisation peut soit survenir avant la production d'une grille composée d'indicateurs de performance et faciliter la création de cette grille, soit être utilisée *a posteriori* afin de valider une grille existante. Dans le présent cas, il est possible de conclure que le modèle à quatre stades d'Ågren (2008) convient bien à la description de la compétence. Le stade initial se compose d'un niveau faible de fluidité, d'une précision globale d'environ 75 %, taux comparable à ce qui est observé pour le marquage du pluriel (environ 74 %). Le stade le plus avancé qui a été inféré dans la présente étude est caractérisé par une fluidité accrue (plus de 500 mots), par une précision globale qui s'approche de 90 % et par une excellente maîtrise du marquage du pluriel (97 %).

L'analyse modélise simultanément les multiples indicateurs de performance. La modélisation simultanée suggère que les paramètres estimés sont les mêmes pour tous les apprenants. Les apprenants sont donc évalués selon les mêmes critères de performance et les mêmes taux de progression. Une telle approche est souhaitable pour la production d'une grille uniforme.

Lorsqu'un devis transversal est utilisé, une modélisation des classes latentes permet l'obtention des caractéristiques des classes. Lorsqu'un devis longitudinal est présent, les taux de transition entre les niveaux (classes) sont obtenus par un modèle LTA. Les écrits suggèrent que les coefficients de transition, qui témoignent de la vitesse d'apprentissage, sont généralement faibles (entre 2 % et 7 %). Dans le présent cas, ce paramètre s'est avéré non significatif et ce modèle n'a pas été retenu. Il y a eu peu d'évolution dans la compétence des apprenants durant la période observée. Cette absence apparente de progression

Tableau 7 ■ Comparaison de la fluidité de chacun des stades identifiés à partir du modèle LCA-4 avec les caractéristiques des stades de la compétence morphosyntaxique d'Ågren (2008) – π représente la proportion d'élèves dans chaque classe

Stade	Ågren	LCA-4	Élèves
A	130	155	Andréa, Amie, Anders, Bibbi, Barbara, Billy, Bror, Conny
B	202	s. o.	
C	276	276	Ceasar, Cecilia, Carolina, Deborah
D	406	413	Daniella, Dagmar
E	s. o.	538	Daisy

Tableau 8 ■ Distribution des textes selon les probabilités d'appartenir à sa classe propre (modèle LCA-4)

Stade	>0,99 %	0,90-0,98	< 0,90
A	23		1
C	7	2	3
D	3	3	
E	3		
Total	36	5	4

suggère que la compétence morphosyntaxique d'une langue seconde évolue très lentement en une année et que plusieurs années d'études sont nécessaires pour observer des changements notables dans la fluidité et la précision globale d'un texte.

L'analyse permet de déterminer que les niveaux sont bien distincts en ce qui a trait aux valeurs des indicateurs. L'inspection des probabilités d'appartenir à une classe ou à une autre permet d'apprécier la certitude quant à la classification qui est effectuée. Une classification caractérisée par de faibles probabilités d'appartenir à une classe est indicative d'un manque de disjonction et, par conséquent, d'un modèle inadéquat. En ce sens, les modèles considérés dans le présent article ont généré des classes disjonctives à une exception près. Le modèle LCA-4 est par ailleurs considéré comme plus exhaustif que les modèles alternatifs à deux niveaux.

Il est possible d'examiner plus en profondeur les conditions de développement des compétences et

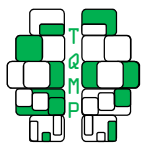
d'enrichir les modèles par l'imposition de contraintes sur les modèles et par l'ajout de covariables. Ces aspects plus avancés n'ont cependant pas été illustrés dans le cadre du présent article. Ainsi, la création ou la validation d'une grille nécessite de bien documenter les situations étudiées et d'être systématique dans l'observation des indicateurs communs aux situations. Il est souhaitable de disposer d'une banque suffisante de données pour procéder à la modélisation. Dans cette perspective, le présent effort de modélisation est réalisé à des fins d'illustration, et non de généralisation.

En ce qui a trait à la fiabilité, la classification obtenue à partir des premiers, deuxièmes et troisièmes textes est identique. En fait, à une exception près, nous n'avons observé aucune transition d'un stade à un autre. La classification est donc stable. Nous ne croyons pas qu'il soit pertinent de calculer un indice de fiabilité test-retest dans ce contexte. De plus, comme les données ont été obtenues uniquement par le logiciel de correction grammaticale *Antidote*, nous ne procéderons pas à une étude de la fiabilité interjuges. Nous savons que les données obtenues par le logiciel *Antidote* ne sont elles-mêmes pas fiables à 100 % et ne remplacent pas l'analyse minutieuse d'un correcteur linguistique expérimenté. En ce sens, la classification obtenue peut être limitée par la qualité de notre base de données.

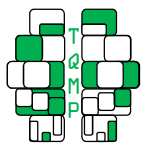
En fait, la contribution du présent article est de suggérer une méthodologie pour établir la validité de construit d'une grille. D'autres preuves complémentaires pourraient aussi être envisagées dans une perspective de validation. Ces preuves pourraient, entre autres, être liées à un critère externe de compétence ou à une épreuve prédictive ou concomitante, ou à une étude approfondie de la fiabilité.

References

- Ågren, M. (2008). À la recherche de la morphologie silencieuse : sur le développement du pluriel en français L2 écrit (Thèse de doctorat non-publiée). Université de Lunds, Suède.
- American Educational Research Association (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Ayotte-Beaudet, J.-P. (2013). The concept of competence in the French-language education literature. *Prospect*, 168, 419-427. doi : 10.1007/s11125-013-9284-8.
- Bartning, I., & Schlyter, S. (2004). Itinéraires



- acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, 14, 281-299. doi : 10.1017/S0959269504001802.
- Blais, J.-G., Laurier, M., & Rousseau, C. (2009). Deriving proficiency scales from performance indicators using the Rasch model. Dans E.V. Smith et G.E. Stone (dir.), *Criterion referenced testing: Practice analysis to score reporting using Rasch measurement* (chap. 24, pp. 528-540). Maple Grove, MN: Journal of Applied Measurement Press.
- Bloom, S.B. (1942). Test reliability for what? *The Journal of Educational Psychology*, 33, 517-526.
- Boscardin, C.K. (2012). Profiling students for remediation using latent class analysis. *Advances in Health Sciences Education*, 17, 55-63. doi : 10.1007/s10459-011-9293-4
- Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.
- Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, 80(4), 796-846. doi: 10.1111/j.1467-6494.2011.00749.x.
- Campanale, F., & Raïche, G. (2008). L'évaluation de la formation supérieure et professionnelle. *Mesure et évaluation en éducation*, 31(3), 35-59.
- Côté, F. (2014). *Construire des grilles d'évaluation descriptives au collégial: guide d'élaboration et exemples de grilles*. Québec : Presses de l'Université du Québec.
- Fu, W.-T., & Anderson, J.R. (2006). From recurrent choice to skill learning: A reinforcement-learning model. *Journal of Experimental Psychology*, 135(2), 184-206. doi : dx.doi.org/10.1037/0096-3445.135.2.184.
- Gibson, W.A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, 24, 229-252.
- Hallgren, K.A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Harvey, L. (2011). Hidden Markov models and learning in authentic situations. *Tutorials in Quantitative Methods for Psychology*, 7(2), 12-21.
- Hélie, S. (2006). An introduction to model selection: Tools and algorithms. *Tutorials in Quantitative Methods for Psychology*, 2(1), 1-10.
- Kaplan, D., & Sweetman, H. (2006). Finite mixture modeling approaches to the study of growth in academic achievement. Dans R.W. Lissitz (dir.), *Longitudinal and value added modeling of student performance* (pp. 130-169). Maple Grove, MN: Journal of Applied Measurement Press.
- Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. Dans J. Harting, E. Klieme, & D. Leutner (dir.), *Assessment of competencies in educational contexts* (pp. 3-176). Cambridge, MA: Hogrefe & Huber.
- Laurencelle, L. (2014). The discriminating capacity of a measuring instrument: Revisiting Bloom (1942)'s theory and formula. *The Quantitative Methods for Psychology*, 10, 5-12.
- Laveault, D., & Grégoire, J. (2002). *Introduction aux théories des tests en psychologie et en sciences de l'éducation* (2^e éd.). Bruxelles : De Boeck.
- Lazarsfeld, P.F., & Henry, N.W. (1968). *Latent structure analysis*. Boston: Houghton Mill.
- Linn, R.L., & Miller, M.D. (2005). *Measurement and assessment in teaching* (9^e éd.). Columbus, OH: Prentice Hall.
- Lissitz, R.W. (dir.) (2006). *Longitudinal and value added modeling of student performance*. Maple Grove, MN: Journal of Applied Measurement Press.
- Littman, M.L. (2010). A tutorial on partially observable Markov decision processes. *Journal of Mathematical Psychology*, 53(3), 119-125. doi:10.1016/j.jmp.2009.01.005.
- Miller, G.A. (1952). Finite Markov processes in psychology. *Psychometrika*, 17(2), 149-167.
- Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. Dans G.R. Hancock & K.M. Samuelsen (dir.), *Advances in latent variable mixture models* (pp. 1-24). Charlotte, NC: IAP.
- Parisse, C. (2009). La morphosyntaxe : qu'est-ce que c'est? Application au cas de la langue française. *Rééducation orthophonique*, 238, 7-20.
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3), 139-154. doi:10.1016/j.jmp.2008.12.005.
- Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, 35, 299-331.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck.
- Vermunt, J.K., & Magidson, J. (2002). Latent class cluster analysis. Dans J. Hagenars & A. McCutcheons (dir.), *Applied latent class analysis* (pp. 89-106). Cambridge: Cambridge University Press.
- Visser, I., Raijmakers, M.E.J., & Molenaar, P.C.M. (2002).



- Fitting hidden Markov models to psychological data. *Scientific Programming*, 10, 185-199.
- von Davier, M. (2008). The mixture general diagnostic model. Dans G.R. Hancock & K.M. Samuelsen (dir.), *Advances in latent variable mixture models* (pp. 255-274). Charlotte, NC: IAP.
- von Davier, M., DiBello, L., & Yamamoto, K. (2008). Reporting test outcomes using models for cognitive diagnosis. Dans J. Harting, E. Klieme & D. Leutner (dir.), *Assessment of competencies in educational contexts* (pp. 151-176). Cambridge, MA: Hogrefe & Huber.
- Wickens, D.W. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco, CA: W. H. Freeman and Company.
- Zaltman, G., Pinson, C.R.A., & Angelmar, R. (1973). *Metatheory and consumer research*. New York: Holt, Rinehart & Winston.

Annexe I: Utilisation de depmix et détails de la modélisation

Le programme *depmix* (Visser et al, 2002) est nécessaire pour la modélisation des données. Le programme *lattice* est utilisé pour générer les histogrammes afin d'obtenir une exploration préliminaire des données. Dans une première section, le fichier de données est présenté. Dans une seconde section, les détails de la programmation sont explicités. Les principales fonctions de *depmix* telles que requises dans le cadre du présent article seront brièvement présentées. Ainsi, quatre fonctions *markovdata*, *lca/dmm*, *fitdmm* et *summary* seront présentées. Dans une dernière section, le rapport produit par la fonction *summary* pour le modèle *lca2* sera commenté. Notez que cette présentation des fonctionnalités de *depmix* n'est pas exhaustive. Le lecteur intéressé devra se référer aux documents originaux produits par les concepteurs.

Fichier de données. Les données se retrouvent dans le fichier nommé *Indicateurs*. Il contient les informations des 45 textes. Les colonnes correspondent respectivement à **V1=Fluidité**, **V2=Précision**, **V3=Accord du pluriel**, **V4=élève**, **V5=temps**. La création de ce fichier se retrouve à la fin de l'annexe mais doit être exécuté en début de programmation.

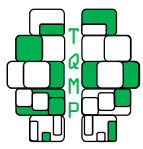
Détails de la programmation. Afin de pouvoir être traité, le fichier de données original doit être formaté par *depmix*. La fonction *markovdata* effectue cette opération. Le fichier résultant s'appelle *Indicateurs2*. Les paramètres de la fonction *markovdata* nécessitent trois types d'informations. Le premier paramètre présente les indicateurs de départ qui se trouvent dans les trois premières colonnes de la variable *Indicateurs* (*Indicateurs*[1:3]). Le second paramètre définit la nature des indicateurs. Dans le présent cas, il s'agit de trois variables continues (*itemtypes=c("cont","cont","cont")*). Le dernier paramètre est un vecteur qui spécifie le nombre d'observations par sujet (*ntimes=nt*). Le vecteur *nt* qui contient 15 éléments (un par sujet) précise qu'il y a 3 observations pour chacun des sujets.

Par la suite, il s'avère nécessaire de définir le modèle qui sera appliqué. Les fonctions *lca* ou *dmm* sont alors utilisées : *lca* sera de mise pour définir un modèle à classes latentes tandis que *dmm* définira un modèle à transitions latentes. La syntaxe de ces deux fonctions est similaire. Le premier paramètre correspond aux nombres de classes et est défini par *ncl* ou il correspond au nombre d'états latents désirés qui est alors spécifié par *nsta*. Il faut par la suite définir le type de distribution qui sera utilisé pour la modélisation. Il s'agit ici d'un vecteur qui sera associé à *itemt*. Des distributions gaussiennes («gaus») sont spécifiées pour les trois indicateurs (*itemt=c("gaus","gaus","gaus")*). Dans les exemples qui suivent trois modèles furent générés sous les noms de *modele1* (*lca2*), *modele2* (*lta2*) et *modele3* (*lca4*). Ces modèles sont traités dans l'article.

La fonction *fitdmm* applique les spécifications du modèle aux données. Le premier paramètre est le fichier de données (*Indicateurs2*) et le second paramètre précise le modèle désiré (*modele1*, *modele2* ou *modele3*). Les résultats de la modélisation se trouvent dans les listes *lca2*, *lta2* et *lca4*. La fonction *summary* permet de visualiser les informations premières.

Les probabilités d'appartenir à une classe peuvent être inspectées via un énoncé tel que *lca4\$post\$states[[1]]* pour le modèle *lca4*. La fonction *round* permet de limiter à deux décimales l'affichage de ces probabilités. La fonction *table* produit une synthèse pour chaque état/classe.

La reproduction du modèle *lca4* tel que présenté dans l'article nécessite que des valeurs initiales soient insérées dans le modèle. Cette situation s'explique par le fait qu'il existe plusieurs solutions vraisemblables pour un modèle à 4 classes latentes et qu'en l'absence de valeurs initiales, le modèle peut converger vers différentes solutions; il semble exister différents maxima locaux pour ce modèle. L'insertion de valeurs initiales permet à l'analyse de



converger vers la solution qui s'est avérée la plus intéressante et qui a été publiée dans l'article.

Pour ce faire, un vecteur nommé *st* spécifiant les valeurs initiales à considérer est créé. La première valeur de ce vecteur est fixée à 1. Par la suite la moyenne et l'écart-type de chacun des trois indicateurs dans chacune des classes sont insérés dans un vecteur appelé *obsmat*. Un vecteur contenant les proportions de textes dans chacune des classes est également créé. Ces valeurs ont été obtenues lors de la solution du modèle publié dans l'article et reproduites ici. Les vecteurs *st* et *obsmat* définissent le vecteur de valeurs initiales qui est inséré dans la fonction *lca* via `stval = st`.

Analyse du fichier de sortie. La fonction *summary* produit un ensemble d'informations générales sur la modélisation concernant la nature du modèle, la date, l'heure et l'année de la modélisation, la vraisemblance du modèle obtenu (-367.075), les indices AIC (760.15) et BIC (783.637), le nombre d'observations utilisées par l'indice BIC, le nombre de paramètres du modèle (19), le nombre de paramètres libres (13), le nombre d'états/classes (2), le nombre d'items/indicateurs considérés (3) ainsi que le type d'items (1 pour continu).

Par la suite, les paramètres sont présentés dans deux matrices. La première présente les valeurs moyennes et les écart-types pour chaque indicateur ainsi que l'erreur standard et la valeur d'un test *t* associé. Ces statistiques sont produites pour chaque classe/état du modèle. Le test *t* permet de déterminer si le paramètre obtenu est significativement différent de zéro. On constate que pour le premier item (la fluidité), la moyenne est de 396.915 mots avec un écart-type de 125.849 mots. Les valeurs des tests *t* associés sont de 12.605 et 4.969; ces paramètres sont donc différents de zéro.

La seconde matrice présente la proportion de textes classifiée soit 40,1% pour la classe 1 et 59,9% pour la classe 2. Les erreurs standards et les tests *t* sont aussi produits pour ces deux paramètres.

Programmation

```
require(depmix)
require(lattice)
```

#Inspection de la distribution des trois indicateurs avec la fonction histogram

```
histogram(Indicateurs[,1], xlab="Fluidité", type="density",
  breaks=c(0,100,200,300,400,500,600,700))
histogram(Indicateurs[,2], xlab="Précision", type="count",
  breaks=c(0,5,10,15,20,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100))
histogram(Indicateurs[,3], xlab="Accord du pluriel", type="count",
  breaks=c(0,.05,.10,.15,.20,.25,.30,.35,.40,.45,.50,.55,.60,.65,.70,.75,.80,.85,.90,.95,1))
```

#Modèle à 2 classes latentes LCA-2

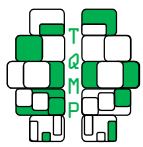
```
nt <- c(3,3,3,3,3,3,3,3,3,3,3,3,3,3,3)
Indicateurs2 <- markovdata(Indicateurs[,1:3],
  itemtypes=c("cont", "cont", "cont"), ntimes=nt)
modele1 <- lca(ncl=2, itemt=c("gaus", "gaus", "gaus"))
lca2 <- fitdmm(Indicateurs2, modele1)
summary(lca2)
```

#Modèle à transitions latentes à deux niveaux LTA-2

```
modele2 <- dmm(nsta=2, itemt=c("gaus", "gaus", "gaus"))
lta2 <- fitdmm(Indicateurs2, modele2)
summary(lta2)
```

#Modèle à 4 classes latentes LCA-4

```
obsmat <- c( 538.417,    79.157,    88.667,    1.700,    0.973,    0.024,
  413.949,    58.008,    89.333,    4.066,    0.908,    0.055,
  275.843,    72.799,    86.000,    2.646,    0.850,    0.098,
  154.870,    49.725,    74.797,    6.509,    0.736,    0.127)
```

```
prop <- c(0.067, 0.133, 0.266, 0.534)
st = c(1,obsmat,prop)
modele3 <- lca(ncl=4,itemt=c("gaus","gaus","gaus"),stval=st)
lca4 <- fitdmm(Indicateurs2, modele3)
summary(lca4)
```

#Inspection des probabilités d'appartenir à une classe pour le modèle LCA-4

```
postround <- round(lca4$post$states[[1]],digits=2)
postlca4 <- table(postround[,2],postround[,3])
```

Sortie

```
> summary(lca2)
```

```
Model: 2 class model fitted at Tue Feb 03 16:31:10 2015
Optimization information, method is optim
Iterations: 83 32
Inform: 0 (look up the respective manuals for more information.)
```

```
Loglikelihood of fitted model: -367.075
AIC: 760.15
BIC: 783.637
Number of observations (used in BIC): 45
Fitted model
Model: 2 class model
Number of parameters: 19
Free parameters: 13
Number of states: 2
Number of items: 3
Item types: 1 1 1
```

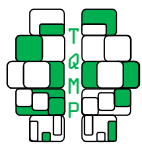
Parameter values, observation parameters

	Item1, mean	Item1, stddev	Item2, mean	Item2, stddev	Item3, mean	Item3, stddev
Class1	396.915	125.849	87.681	3.367	0.898	0.075
se	31.488	25.329	0.797	0.556	0.018	0.013
t	12.605	4.969	110.005	6.056	49.802	5.592
Class2	159.329	54.437	75.881	6.983	0.743	0.129
se	10.612	7.559	1.364	0.960	0.025	0.018
t	15.014	7.201	55.618	7.276	29.940	7.327

Parameter values, unconditional (class) probabilities

	Class1	Class2
val	0.401	0.599
se	0.127	0.127
t	3.152	4.701

(appendix continues on next page)



Fichier de données.

```
Indicateurs <- read.table(header=TRUE,
  sep=" ", dec=".", text=
" V1 V2 V3 V4 V5
158 76 0.59 1 1
 58 87 0.62 1 2
132 72 0.60 1 3
100 72 0.90 2 1
 58 79 0.65 2 2
179 72 0.65 2 3
197 71 0.59 3 1
138 75 0.61 3 2
124 63 0.64 3 3
 89 86 0.82 4 1
163 78 0.68 4 2
207 64 0.92 4 3
112 71 0.68 5 1
137 77 1.00 5 2
138 77 0.90 5 3
213 73 0.82 6 1
170 64 0.57 6 2
262 64 0.70 6 3
197 76 0.82 7 1
129 80 0.69 7 2
206 78 0.83 7 3
159 78 0.93 8 1
182 85 0.67 8 2
209 77 0.79 8 3
185 88 0.80 9 1
179 82 0.64 9 2
295 85 0.96 9 3
287 86 0.90 10 1
206 80 0.82 10 2
352 85 0.89 10 3
229 88 0.78 11 1
258 90 0.73 11 2
364 88 0.90 11 3
354 87 1.00 12 1
207 87 0.85 12 2
393 86 0.93 12 3
413 90 0.86 13 1
320 95 0.97 13 2
432 94 0.99 13 3
457 87 0.91 14 1
367 86 0.84 14 2
496 84 0.88 14 3
630 91 0.94 15 1
437 87 0.99 15 2
543 88 0.99 15 3"
)
```

Citation

Harvey, L. (2015). Analyses des classes et des transitions latentes : des outils pour documenter le parcours de développement des compétences. *The Quantitative Methods for Psychology*, 11 (2), 63-77.

Copyright © 2015 Harvey. This is an open-access article distributed under the terms of the *Creative Commons Attribution License (CC BY)*. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 07/01/15 ~ Accepted: 09/02/15