

# HIERARCHISATION DE COMPETENCES PAR L'INTERMEDIAIRE DU LOGICIEL C.H.I.C. : EXEMPLES D'ANALYSES DE DONNEES DANS DEUX DOMAINES DIFFERENTS

Stéphanie Malaise\*, Dominique Casanova\*\*

\* Université de Mons, [stephanie.malaise@umons.ac.be](mailto:stephanie.malaise@umons.ac.be)

\*\* Chambre de Commerce et d'Industrie de Paris, [dcasanova@ccip.fr](mailto:dcasanova@ccip.fr)

---

**Mots-clés :** évaluation de compétences - hiérarchisation de compétences - analyse implicative

**Résumé.** A travers deux exemples issus de contextes différents, les auteurs proposent de montrer comment l'exploitation de deux types de résultats obtenus par l'intermédiaire du logiciel C.H.I.C.<sup>1</sup> – le graphe implicatif et l'arbre de similarités – permet d'appréhender la hiérarchie existant entre différentes compétences. Le premier exemple est issu d'une recherche menée à l'Université de Mons et portant sur le diagnostic de la maîtrise de compétences issues du traitement de données, par des élèves de grade 5 et 6 (10-12 ans). Le second exemple est issu d'une recherche menée à la Chambre de Commerce et d'Industrie de Paris et portant sur l'évaluation de la maîtrise de la langue française par des adultes.

---

## 1. Introduction

La hiérarchisation de compétences peut être effectuée dans des buts différents. Au niveau didactique, distinguer les compétences qui sont les plus simples de celles qui sont le plus complexes peut aider l'enseignant à organiser les séquences d'enseignement/apprentissage. La connaissance de la structure hiérarchique des différentes compétences qui composent un curriculum est également un avantage en termes de diagnostic. En effet, identifier dans quelle mesure la maîtrise d'une compétence est nécessaire à la maîtrise d'une autre compétence peut aider l'enseignant à proposer une remédiation adéquate lorsqu'un élève éprouve des difficultés. Enfin, la hiérarchisation de compétence est particulièrement utile dans le cadre de l'élaboration de tests standardisés amenant à l'attestation d'un niveau de maîtrise précis d'une compétence.

Bien que différents modèles de l'évaluation de la maîtrise de compétences aient été élaborés (Carette, 2007 ; De Ketele & Gerard, 2005 ; Scallon, 2007), aucun ne s'avère être totalement satisfaisant et, à l'heure actuelle, « on sait peu de choses sur la façon de graduer la progression des élèves dans le développement ou la construction de compétences » (Scallon, 2007, p. 121). Dans ce cadre, le logiciel C.H.I.C., utilisant l'analyse statistique implicative (Couturier, 2001 ; Gras & Régnier, 2009 ; Gras, Kuntz & Briand, 2001), offre l'opportunité de dépasser ce constat. A travers deux exemples différents, les auteurs illustrent comment, grâce à l'utilisation du logiciel, il est possible de représenter la structure hiérarchique d'un ensemble de compétences.

## 2. Le logiciel C.H.I.C.

Le logiciel C.H.I.C. est un outil qui, grâce l'utilisation d'analyses statistiques implicatives, permet d'identifier des relations de quasi-implication du type « si a alors presque b » (Gras, 1996) existant entre des couples de variables. Il permet, en outre, de quantifier l'intensité de cette relation

---

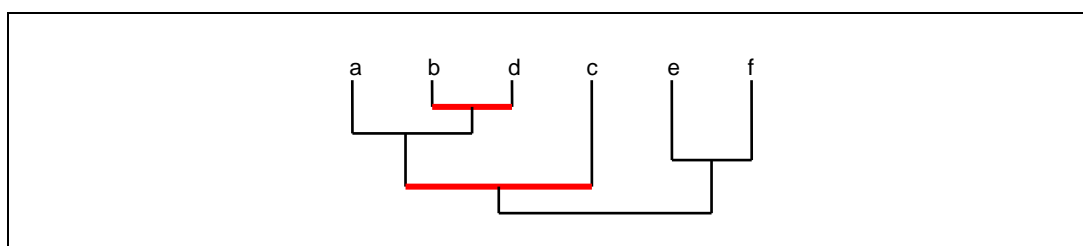
<sup>1</sup> Classification Hiérarchique Implicative et Cohésitive

lorsque « l'implication est non stricte, c'est-à-dire lorsque l'ensemble  $A$  des individus qui vérifient  $a$  n'est pas inclus dans l'ensemble  $B$  des individus qui vérifient  $b$  » (Ottaviani & Zannoni, 2001, p. 62).

L'utilisation d'une approche probabiliste, telle que le propose le logiciel C.H.I.C. apparaît comme étant plus appropriée que les analyses fondées sur des échelles de mesure « classiques » pour rendre compte d'une hiérarchie (Alexandre, 1971 ; Delhaxhe & Crahay, 1983 ; Loubet Del Baye, 2000). En effet, la mesure de la relation implicative « ne relève pas uniquement de la comparaison des seules fréquences de réussite, ni même de l'examen des coefficients de corrélation entre deux réussites » (Gras, 1996, p. 13) - ces derniers ne pouvant pas « restituer l'information recherchée à travers la relation d'implication » (*Ibid.*, p.14) - mais détermine la qualité confirmatoire d'une relation implicative  $a \Rightarrow b$  « à partir de l'in vraisemblance de l'apparition, dans les données, du nombre de cas qui l'infirmant, c'est-à-dire pour lesquels  $a$  est vérifié sans que  $b$  ne le soit » (Gras & Reigner, 2009, p.19). « Le nombre  $k$  de contre-exemples est considéré comme la réalisation d'une variable aléatoire d'un modèle de Poisson ou d'un modèle binomial, l'un et l'autre approximatés par une loi de Gauss lorsqu'elle est légitimée par les paramètres. (...) L'implication est admissible au seuil  $\alpha$  si la probabilité que cette variable aléatoire soit supérieure à  $k$  est elle-même supérieure à  $1 - \alpha$  » (Couturier, Bodin & Gras, 2010), c'est-à-dire à l'intensité d'implication. Lorsque  $n$ ,  $a$  et  $b$  sont supérieures à plusieurs dizaines d'unités, une intensité de 0,95 constitue une bonne valeur d'admissibilité.

Le logiciel C.H.I.C. fournit des représentations des résultats qu'il permet d'obtenir, notamment à travers la création de graphes implicatifs et d'arbres de similarités. Le graphe implicatif représente de manière synthétique l'ensemble des relations de quasi-implication existant entre les couples de variables étudiées (Gras, 1996). Utilisé dans le cadre de l'évaluation de compétences, il permet d'identifier des précurseurs, c'est-à-dire des compétences qui, lorsqu'elles sont maîtrisées favorisent l'acquisition de la maîtrise d'autres compétences, d'un niveau supérieur.

L'arbre de similarités est un type de représentation qui permet d'identifier les variables comportant les mêmes caractéristiques. Au premier niveau, l'arbre réunit les deux variables les plus similaires en une classe. Au niveau suivant, il réunit, selon les cas, deux autres variables ou une variable et la classe du premier niveau. Au troisième niveau, soit deux nouvelles variables sont réunies, soit une variable est associée à une classe, soit encore, les deux premières classes sont réunies (Gras & Régnier, 2009). Dans l'exemple suivant (*cf.* figure 1), empruntée à Couturier (2001),  $b$  et  $d$  sont les variables qui sont les plus semblables. Elles forment donc la première classe de l'arbre hiérarchique. La variable  $a$  est ensuite agrégée à cette première classe. Au troisième niveau, ce sont les variables  $e$  et  $f$  qui sont réunies car elles sont plus similaires que toute extension de  $(a,b,d)$ . Au quatrième niveau, la variable  $c$  est groupée à la classe  $(a,(b,d))$ . Elle est donc meilleure que toute extension de la classe  $(e,f)$ . Finalement, la classe  $(a,b,d,c)$  et la classe  $(e,f)$  sont regroupées en une classe. Les niveaux représentés par un trait gras ont « plus de signification classifiante que les autres niveaux » (Couturier, 2001, p. 35).



**Figure 1** : Exemple d'arbre des similarités (Couturier, 2001, p. 35)

### 3. Hiérarchisation de compétences relatives au traitement de données

Dans le cadre d'une recherche pluriannuelle commanditée par la Communauté française de Belgique<sup>2</sup> (Dehon, Demierbe, Derobertmeasure & Malaise, 2008 ; Malaise, Dehon & Franquet, 2009 ; Malaise & Dandoy, 2010), des chercheurs de l'Institut d'Administration Scolaire (UMONS) avaient pour objectif de déterminer, empiriquement, les liens hiérarchiques existant entre la maîtrise, par des élèves de grades 5 et 6 (10 - 12 ans), de quatre compétences issues des Socles de compétences<sup>3</sup> et relatives au traitement de données :

- Organiser selon un critère ;
- Lire un graphique, un tableau, un diagramme ;
- Interpréter un tableau de nombres, un graphique, un diagramme ;
- Représenter des données par un graphique, un diagramme.

Outre la hiérarchisation des quatre compétences, l'équipe de recherche s'est également intéressée aux liens existant entre trois niveaux distincts de chacune de ces compétences, du plus simple (niveau 1) au plus complexe (niveau 3). Les trois niveaux de complexité ont été définis à partir de la classification d'exercices issus de manuels scolaires et selon une logique proche de celle de Rey, Carette, Defrance et Kahn (2006) qui, grâce à leur modèle de l'évaluation en trois étapes, offrent un moyen de cibler l'origine des difficultés rencontrées par l'élève, lors de la résolution d'une situation complexe.

#### 3.1. Méthode

Dans le cadre de la première année de recherche, 121 élèves ont fait l'objet d'un test papier/crayon évaluant la maîtrise de six des douze combinaisons « compétence x niveau ». Des analyses ont alors permis de proposer une première hiérarchie partielle (Malaise, 2010 ; Malaise, 2011).

Lors de la deuxième année de recherche, ce sont les douze combinaisons qui ont fait l'objet d'une évaluation. Deux cahiers de test similaires - composés des mêmes exercices mais variant soit au niveau de complexité, soit au niveau de la présentation - ont été proposés à 217 élèves (N=108 pour l'un des cahiers N = 109 pour l'autre). L'objectif était de valider la hiérarchie créée et de mettre en évidence les facteurs qui influencent la complexité d'une situation.

Enfin, la troisième année de recherche a donné lieu à la validation des facteurs influençant la complexité d'un exercice et a permis de proposer une hiérarchie entre les différentes combinaisons. Au total, 619 élèves ont été interrogés à travers 4 cahiers de test similaires. Ce sont les résultats de cette troisième année de recherche qui feront l'objet d'une description dans cet article.

#### 3.2. Analyses

##### 3.2.1. Vérification des niveaux de complexité par l'intermédiaire de l'arbre des similarités

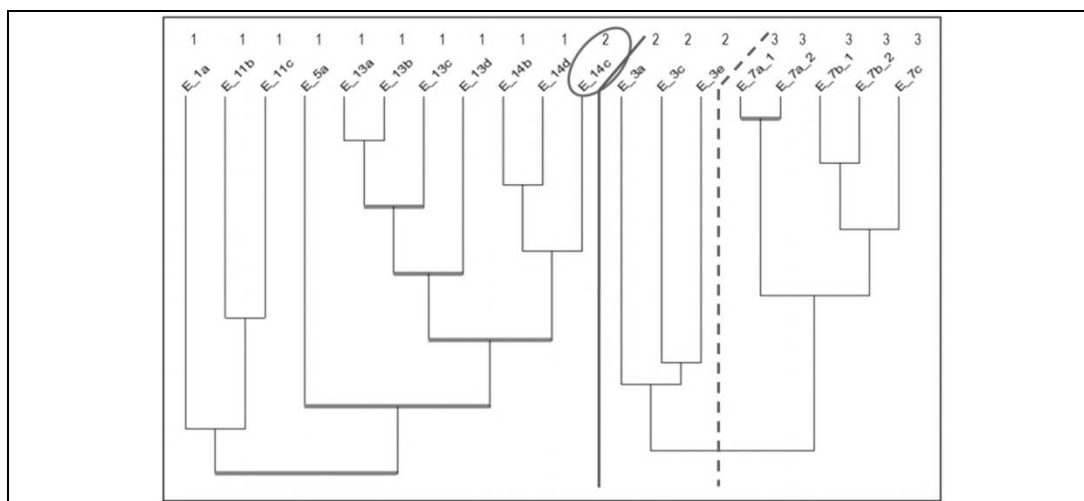
Dans le premier temps de la troisième année de recherche, les similarités entre items ont été analysées en vue de déterminer si, pour chaque item des tests, le niveau de complexité déterminé *a priori* correspondait bien au niveau de complexité réellement observé. Pour chacun des quatre Cahiers de tests et pour chaque compétence, un arbre des similarités a été construit.

---

<sup>2</sup> Mieux diagnostique pour mieux enseigner : développement d'outils de diagnostic de la maîtrise, par les élèves, de compétences ciblées à l'issue de l'enseignement primaire.

<sup>3</sup> « Référentiel présentant de manière structurée les compétences de base à exercer jusqu'au terme des huit premières années de l'enseignement obligatoire et celles qui sont à maîtriser à la fin de chacune des étapes de celles-ci parce qu'elles sont considérées comme nécessaires à l'insertion sociale et à la poursuite des études » (Communauté française de Belgique, 1997, Art 5, 2°).

La figure 2, présente l'arbre de similarités relatif aux items destinés à évaluer la compétence « Lire » pour l'un des quatre cahiers. Le graphe se divise en trois parties. La première, à gauche, est séparée des deux suivantes et reprend l'ensemble des items jugés comme étant du premier niveau de complexité ainsi qu'un item jugé comme étant du second niveau de complexité. La deuxième partie du graphe comprend trois des quatre items ayant été définis comme étant du second niveau de complexité. Enfin, la troisième partie du graphe présente l'ensemble des items du troisième niveau de complexité. Bien que les deux dernières parties du graphe se rejoignent en une seule classe, celle-ci est créée à un niveau inférieur, on peut donc les considérer comme étant distinctes.



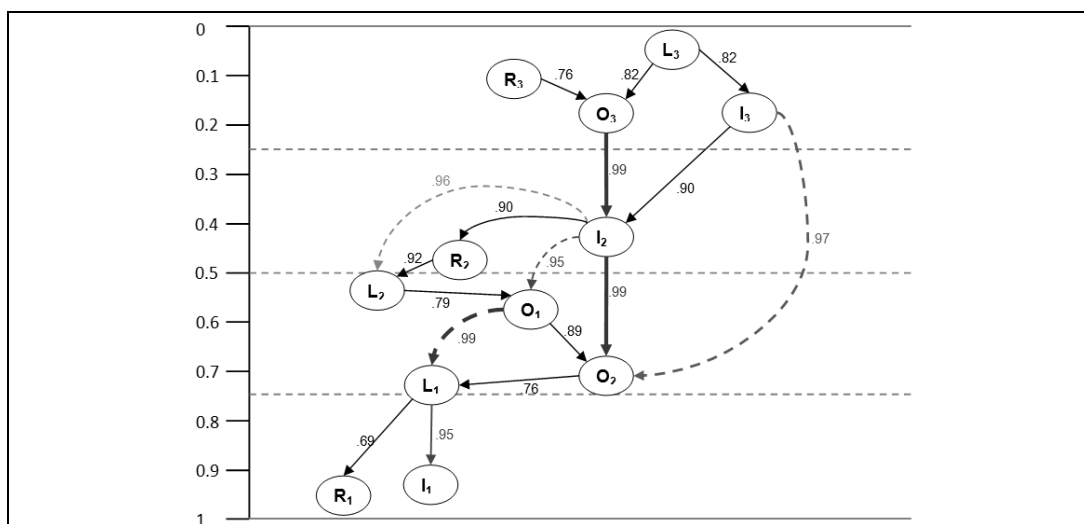
**Figure 2** : Arbres de similarités relatif à la compétence "Lire" d'un des quatre cahiers de test

De l'analyse de cet arbre de similarités, il apparaît que parmi les items relatifs à la compétence « Lire » un seul ait été mal jugé : l'item 14.c, identifié préalablement comme étant relatif à la combinaison « lire 2 », est associé à des items de niveau 1. Une analyse fine de l'exercice a ensuite permis d'expliquer ce glissement d'une classe à l'autre et ainsi de préciser les critères influençant la complexité d'une situation faisant appel aux combinaisons « Lire 1 » et « Lire 2 ».

Une fois la vérification du niveau de complexité de chaque item effectuée, la deuxième phase d'analyse a consisté à créer les graphes implicatifs relatifs à chacun des cahiers de test.

### 3.2.2. Hiérarchisation des combinaisons « compétence x niveau »

Dans un premier temps, il a été demandé au logiciel de ne faire apparaître que les liens d'implication significatifs. Dans un second temps, le seuil d'implication a été diminué en vue de faire apparaître toutes les combinaisons (compétence x niveau) dans le graphe. Pour chaque cahier, deux graphes implicatifs ont donc été construits. La figure 3 présente le second graphe implicatif issu de l'analyse de l'un des quatre cahiers.



**Figure 3** : Exemples de graphes implicatifs

Les combinaisons les plus faciles, c'est-à-dire, celles qui sont le plus souvent réussies, se situent dans le bas du graphe. A l'inverse, les combinaisons les moins réussies se situent dans la partie haute du graphe. Les flèches se lisent de manière suivante : « si l'élève maîtrise la combinaison située à l'origine de la flèche, alors, il est fort probable qu'il maîtrise également la combinaison située à l'arrivée de la flèche ». Le lien suivant «  $O_3 \rightarrow I_2$  » se traduit de manière suivante : « si l'élève maîtrise la combinaison "organiser 3", alors, il est fort probable qu'il maîtrise également la combinaison "Interpréter 2" ».

La lecture du graphe dans le sens opposé, permet l'identification de précurseurs. Dans ce cas, le lien «  $O_3 \rightarrow I_2$  » se traduit de manière suivante : « la maîtrise de la combinaison "Interpréter 2" est un précurseur à la maîtrise de la combinaison "Organiser 3" » ou encore « l'acquisition de la maîtrise de la combinaison "Organiser 3" sera favorisée si l'élève maîtrise la combinaison "Interpréter 2" ».

Ce graphe permet de constater que la combinaison « Interpréter 2 » occupe une place centrale dans la maîtrise du traitement de données. La maîtrise de la compétence « Interpréter » au deuxième niveau de complexité semble nécessiter la maîtrise préalable de l'ensemble des quatre compétences aux niveaux 1 et 2. La maîtrise de cette combinaison semble, à son tour, être un précurseur à la maîtrise des quatre compétences au niveau le plus complexe.

Dans le cadre scolaire, ce type de représentation hiérarchique peut être utilisé de deux façons. Il peut tout d'abord servir de guide pour la structuration des apprentissages : il est alors un outil aidant à planifier les différentes activités visant le développement des compétences de traitement de données (les activités visant le développement de la compétence « Organiser » de niveau 3 ne seront mises en place que lorsque la compétence « Interpréter » de niveau 2 sera elle-même maîtrisée). En outre, ce graphe peut être utilisé en tant que grille de diagnostic. Il est, par exemple, possible qu'une difficulté rencontrée par un élève lors de la mise en œuvre d'une tâche relative à la combinaison « Organiser 3 » soit liée à un manque de maîtrise de la compétence « Interpréter » au deuxième niveau. Dans ce cas, la remédiation devrait donc aussi porter sur cette combinaison.

Les liens mis en évidence dans les huit graphes implicatifs ont ensuite été synthétisés au sein d'un seul et même graphe destiné d'une part, à aider les enseignants à planifier leurs séquences d'apprentissage et, d'autre part, à déterminer, lorsqu'un élève rencontre une difficulté, quelles sont les compétences sur lesquelles il pourrait être nécessaire d'intervenir.

## **4. Hiérarchisation des composantes d'une épreuve de français langue étrangère**

### **4.1. Le Test d'évaluation de français**

Le Test d'évaluation de français (TEF) est un test standardisé d'évaluation des compétences de communication en langue française des adultes non francophones (Noël-Jothy et Sampsonis, 2006 : 74-75). Ce test, conçu et diffusé par la Chambre de commerce et d'industrie de Paris (CCIP), est composé de cinq épreuves :

- Trois épreuves sous forme de questionnaires à choix multiples (compréhension écrite, compréhension orale et lexique/structure), corrigées par lecture optique ;
- Deux épreuves ouvertes d'expression écrite et d'expression orale, évaluées par des correcteurs et des examinateurs au moyen d'une grille analytique.

Le TEF fournit aux candidats une photographie linguistique de leur niveau de compétence en langue française en les situant sur une échelle de 7 niveaux, qui s'étend du niveau 0+ (capacités les plus élémentaires) au niveau 6 (maîtrise complète). La CCIP a procédé à l'alignement de l'échelle de niveaux du TEF avec les échelles des standards internationaux (Demeuse *et al.*, 2004, Casanova *et al.*, 2010) que sont le Cadre européen commun de référence pour les langues (Conseil de l'Europe, 2002) et les Niveaux de compétence linguistique canadiens (Citoyenneté Immigration Canada, 2006).

Les jeux épreuves au format QCM du TEF exposent les candidats à un ensemble d'items permettant de tester différents objectifs pédagogiques à différents niveaux. Une table de spécification permet de concevoir des formes parallèles du test en précisant quel est le type d'items attendu à chaque emplacement du test. Les items sont inclus dans les jeux d'épreuve après avoir été validés sur le plan pédagogique puis mis à l'épreuve dans le cadre de pré-tests (Demeuse *et al.*, 2005). La difficulté des items pré-testés est estimée au moyen de la théorie de réponse à l'item (modèle de Rasch) après élimination des items dont le pouvoir discriminant n'est pas jugé satisfaisant ou qui ne s'ajustent pas correctement au modèle. Des analyses a posteriori sont mises en œuvre sur des échantillons plus larges de candidats, en situation réelle de tests et non lors de pré-tests sans enjeu, qui permettent d'actualiser et de contrôler la stabilité des paramètres psychométriques des items au cours du temps et de s'assurer que l'assemblage des items réalisé offrait un bon niveau de fidélité de l'évaluation.

Si le modèle de Rasch permet d'estimer les différences de difficulté entre items (et donc de savoir si l'item *A* a davantage tendance à être réussi que l'item *B*), il ne permet pas d'indiquer dans quelles mesure les candidats qui réussissent les items de type *A* ont une forte probabilité de réussir les items de type *B* (notamment lorsque les items tests des objectifs pédagogiques différents au sein d'un même domaine). Or la mise en évidence de telles relations de quasi-implication présente un intérêt indéniable sur le plan pédagogique dans la mesure où elle permet de valider empiriquement certaines hypothèses.

### **4.2. Méthode**

Contrairement à la recherche évoquée en première partie de l'article, les travaux présentés ici sont essentiellement exploratoires, l'objectif de cette recherche étant d'évaluer l'intérêt de l'utilisation d'un logiciel tel que C.H.I.C., à partir des résultats de candidats, dans un processus de validation des épreuves du TEF. Les analyses présentées dans le cadre de cette communication portent essentiellement sur l'épreuve de lexique/structure du TEF.

Afin de disposer d'informations généralisables et de pouvoir raisonner sur un nombre limité de composants, nous avons regroupé les items de tests au sein de classes, chaque classe d'items (ou unité-type) correspondant à l'ensemble des items d'une épreuve visant le même objectif

pédagogique au même niveau de difficulté<sup>4</sup> à partir du même type de support. Ainsi, pour chacune des classes constituées, une valeur 1 a été attribuée au candidat lorsqu'il réussissait plus de 50% des items de la classe et une valeur 0 sinon. Toujours par souci de généralisation, nous avons cumulé les résultats des candidats à une dizaine de formes parallèles du test (soit un échantillon total de 5.513 candidats). Si ce procédé peut introduire un « bruit » dans les valeurs attribuées aux classes comportant peu d'items (deux formes d'un même test ne sont jamais exactement parallèles), il permet de tirer des conclusions générales là où les conclusions relatives à une seule forme nécessiteraient une réplication de l'analyse pour en éprouver la stabilité.

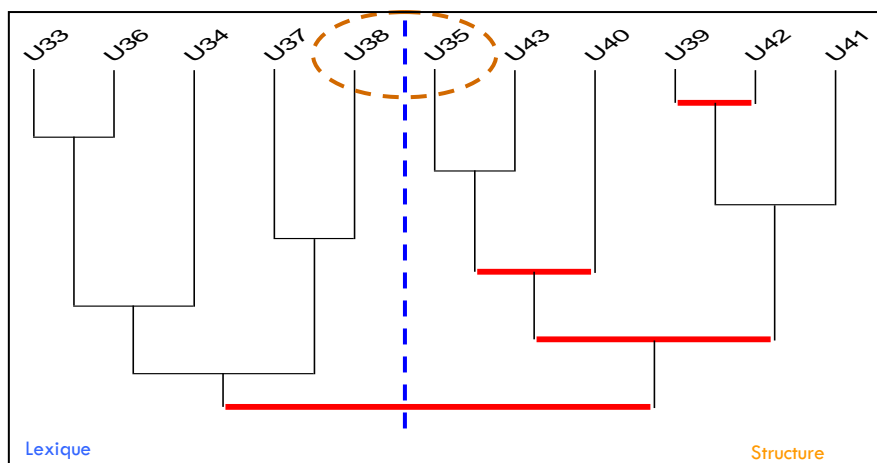
Compte-tenu de la taille de l'échantillon, nous avons utilisé la méthode entropique (Gras, Kuntz & Briand, 2001), fortement conseillée lorsque le nombre d'individus dépasse quelques centaines.

### 4.3. Analyses

L'épreuve de lexique/structure présente plusieurs particularités :

- Elle combine, au sein d'un questionnaire, des items relatifs à deux domaines de connaissance différents (les connaissances lexicales et les connaissances grammaticales) ;
- Elle est principalement composée d'items indépendants les uns des autres où les candidats doivent compléter une phrase en sélectionnant l'unité lexicale ou la structure syntaxique appropriée. Seules une unité-type en lexique (U37) et une unité-type en structure (U43) sont constituées de plusieurs items en relation avec un document support.

L'arbre de similarité des différentes unités-types composant cette épreuve (Cf. figure 5) montre que, alors qu'on constate des similarités entre unités-types relatives à un même domaine (les unités relatives au lexique sont situées dans la partie gauche de la figure et les unités relatives à la structure dans la partie droite), il n'y a pas, à deux exceptions près<sup>5</sup>, de relations de similarité entre unités-types de domaines différents. Cela confirme la dimension double de l'épreuve. Cependant on constate également que la relation entre la classe des items de la moitié gauche (donc principalement les items de lexique) et la classe des items de la moitié droite (donc principalement les items de structure) est fortement significative, traduisant un degré élevé de similarité entre les résultats aux deux parties de l'épreuve.



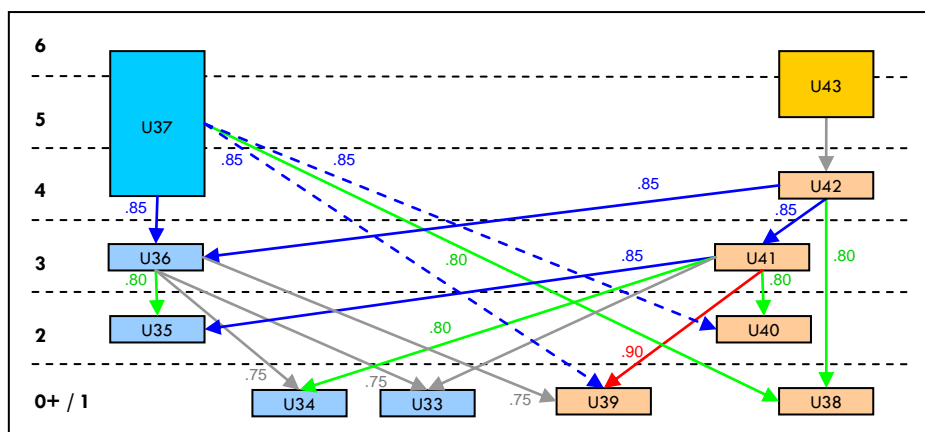
**Figure 4** : Arbre de similarité des unités composant l'épreuve de lexique/structure

<sup>4</sup> A l'exception de quelques unités-types, constituées de plusieurs items se rapportant à un même document support et partageant un même objectif pédagogique mais à des niveaux différents.

<sup>5</sup> L'unité lexicale U35 se situe dans la partie « syntaxe » de la figure et l'unité syntaxique U38 se situe dans la partie « lexique ».

Le graphe implicatif permet d'explorer les relations d'implication existant entre les différents types d'unités. Selon le tutoriel du logiciel C.H.I.C., une intensité de l'ordre de .75 avec la méthode entropique est aussi bonne que de fortes intensités dans le cas classique.

On y voit sans grande surprise que la réussite des items de niveau élevé à l'une des parties du test implique en général la réussite aux items des unités de niveau inférieur de la même sous-partie. Il est plus surprenant de constater que l'absence de relation d'implication entre l'unité de type U38, censée être des plus simples (niveau 0+) et l'unité-type U41, de niveau 3. Cela est probablement dû au fait que, comme la plupart des candidats se sont vu attribuer la valeur 1 à l'unité-type U38, l'intensité de la relation sera très sensible à la présence de contre-exemples à la relation (par exemple, des erreurs d'étourderies pour U38 ou des réponses justes par hasard pour U41).



**Figure 5 :** Graphe implicatif des composantes de l'épreuve de lexique/structure

On constate également dans ce graphique que la réussite des unités de structure d'un niveau implique la réussite des unités de lexique des niveaux inférieurs. Cela n'est pas surprenant dans la mesure où la réponse à l'item suppose la compréhension au moins partielle de la phrase par le candidat, et donc du lexique correspondant. C'est la raison pour laquelle il est difficile de dissocier les dimensions de lexique et de structure en proposant deux sous-tests distincts : dans un test communicatif, où les items testent des connaissances ou des compétences dans le contexte de phrases ou de textes porteurs de sens, l'aspect structure ne peut être évalué isolément.

On constate en revanche peu de relations d'implication entre les unités de lexique les plus élémentaires, qui s'appuient sur des phrases simples du point de vue de la syntaxe, et les unités de structure. Le cas particulier de l'unité type U37 s'explique par le fait que les items de cette unité testent la capacité des candidats à trouver le sens d'un mot à partir du contexte<sup>6</sup> mobilisant pour cela leurs connaissances syntaxiques.

Les relations ainsi mises en évidence confirment donc l'interdépendance des deux parties de l'épreuve de lexique et structure. L'orientation de ces relations (qui toutes partent d'un niveau supérieur pour aboutir à un niveau inférieur) montre d'ailleurs que les composantes ciblant un niveau donné en lexique ne sont pas plus difficiles ou plus faciles que les composantes ciblant le même niveau en structure. Des analyses complémentaires relatives à la dimensionnalité de l'épreuve (analyse factorielle, dimtest) sont actuellement à l'étude.

<sup>6</sup> Les candidats doivent en effet trouver, parmi 4 options de réponse, celle dont le sens est le plus proche du mot (ou groupe de mot) souligné dans le texte.



## 5. Conclusion

L'analyse statistique implicative permet d'analyser, à partir d'échantillons de données plus ou moins importants (méthode entropique ou méthode classique), les relations entre les différentes composantes d'un test, que ce soit en phase d'élaboration, suite à une première expérimentation d'une forme de test, ou a posteriori. La mise en évidence de relations hiérarchique entre compétences peut aider, comme dans le cas de la première étude, à l'organisation des séquences d'enseignement/apprentissage. Dans le cas d'un test standardisé, elle donne la possibilité d'interroger la pertinence de choix pédagogique, que ce soit en phase d'élaboration, à partir des résultats de l'expérimentation d'une première forme du test, ou dans le cadre d'un processus de validation a posteriori, à partir de données de candidats en situation réelle de test avec enjeu.

La restitution graphique des résultats par le logiciel C.H.I.C. permet par ailleurs de partager les constats avec des pédagogues qui ne sont pas nécessairement familiarisés avec les analyses statistiques et de s'assurer avec eux de la cohérence des résultats avec les présupposés pédagogiques. Les constats qui en résultent peuvent conforter ou remettre en cause les choix initiaux, mais ils peuvent également contribuer à la formulation de nouvelles hypothèses. En tout cas, ils contribuent à une meilleure connaissance, par les concepteurs, des caractéristiques de l'outil d'évaluation et offrent donc un complément utile aux analyses psychométriques plus classiques.

## 6. Références et bibliographie

- Alexandre, V. (1971). *Les échelles d'attitude*. France : Editions universitaires.
- Carette, V. (2007). Les implications de la notion de compétence sur l'évaluation. *Education-Formation*, 286, 51-61.
- Casanova, D., Crendal, A., Demeuse, M., Desroches, F. & Holle, A. (2010). *Validation empirique d'un test de français langue étrangère en regard du Cadre européen commun de référence pour les langues*. Actes du 22e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (Braga, Portugal, 14-16 janvier 2010).
- Communauté française de Belgique. (1997). *Décret du 24-07-1997 définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre*. Moniteur belge.
- Couturier, R. (2001). Traitement de l'analyse statistique implicative dans CHIC. *Actes des Journées « Fouille des données par l'analyse statistique implicative »*, 33-50.
- Couturier, R., Bodin, A. & Gras, R. (2010). *CHIC v5.0 Classification Hiérarchique Implicative et Cohésitive : Notice d'utilisation*. Ecole Polytechnique : Nantes.
- Dehon, A., Demierbe, C., Derobertmeasure, A. & Malaise, S. (2008). *Mieux diagnostiquer pour mieux enseigner : Développement d'outils de diagnostic de la maîtrise par les élèves de compétences ciblées à l'issue de l'enseignement primaire*. (Rapport final 1<sup>ère</sup> année). Mons : Université Mons, Institut d'Administration scolaire, service de Méthodologie et Formation, non publié.
- De Ketele, J.-M. & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par compétence. *Mesure et évaluation en éducation*, 28(3), 1-26.
- Delhaxhe, A. & Crahay, M. (1983). La compréhension du fonctionnement de la balance : Une analyse hiérarchique. *European Journal of Psychology of Education*, 4(3), 349-375.
- Demeuse, M., Desroches, F., Crendal, A., Renaud, F., Oster, P. & Leroux X. (2004). *L'évaluation des compétences linguistiques des adultes en français langue étrangère dans une perspective de multiréférentialisation*. Actes du 17e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (Lisbonne, Portugal, 18-20 novembre 2004).

- Demeuse, M., Desroches, F., Crendal, A., Renaud, F. & Casanova, D. (2005). *La fiabilité de l'évaluation des compétences linguistiques pour des adultes non francophones : présentation d'un protocole d'évaluation*. Actes du 18e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe). Reims, France, 24-26 octobre 2005.
- Gras, R. (1996). *L'implication statistique : Nouvelle méthode exploratoire de données*. France : La Pensée sauvage éditions.
- Gras, R. & Régnier, J.-C. (2009). Fondements de l'Analyse statistique Implicative – A.S.I. In Gras, R (dir.) *Analyse Statistique Implicative – Une méthode d'analyse de données pour la recherche de causalité*, 17-41. Toulouse : Cépaduès.
- Gras, R, Kuntz, P & Briand, H. (2001). Les fondements de l'analyse statistique implicative et quelques prolongements pour la fouille des données. *Mathématiques et Sciences humaines*, 154-155, 9-29.
- Laurier M., Toussignant, R. & Morissette, D. (2005). *Les principes de la mesure et de l'évaluation des apprentissages*. Montréal : Gaëtan Morin.
- Lempereur, A., Demeuse, M., Straeten, M.-H. (2000). *Évaluer des compétences à l'école*. Liège : Service de Pédagogie expérimentale de l'Université.
- Loubet Del Baye, J.-L. (2000). *Initiation aux méthodes des sciences sociales*. Paris : L'Harmattan.
- Malaise, S. (2011). Classification hiérarchique de compétences : comparaison d'une méthode basée sur l'analyse de fréquences de réussite et d'une méthode d'analyse statistique implicative. *Le Bulletin de l'ADMEE-Europe*, 2011(1), 8-18.
- Malaise, S., Dehon, A. & Franquet, A. (2009). *Mieux diagnostiquer pour mieux enseigner : Développement d'outils de diagnostic de la maîtrise par les élèves de compétences ciblées à l'issue de l'enseignement primaire*. (Rapport final 2<sup>ème</sup> année). Mons : Université Mons, Institut d'Administration scolaire, service de Méthodologie et Formation, non publié.
- Malaise, S. (2010). Classification hiérarchique de compétences par l'intermédiaire du logiciel C.H.I.C. fondé sur la méthode d'analyse statistique implicative : *Actes du congrès international d'Actualités de la Recherche en Education et en Formation AREF 2010* (Genève, 13 – 16 septembre 2010).
- Malaise, S. & Dandoy, A. (2010). *Mieux diagnostiquer pour mieux enseigner : Développement d'outils de diagnostic de la maîtrise par les élèves de compétences ciblées à l'issue de l'enseignement primaire*. (Rapport intermédiaire 3<sup>ème</sup> année). Mons : Université Mons, Institut d'Administration scolaire, service de Méthodologie et Formation, non publié.
- Noël-Jothy, F., & Sampsonis, B. (2006). *Certifications et outils d'évaluation en FLE*. Paris: Hachette.
- Ottaviani, M.-G. & Zannoni, S. (2001) Implication statistique et recherche en didactique : utilisation d'un outil non symétrique d'analyse de données pour l'interprétation des résultats d'un test d'évaluation. *Mathématiques et Sciences humaines*, 154-155, 61-79.
- Perrenoud, P. (1997). *Construire des compétences dès l'école*. Paris : ESF.
- Rey, B., Carette, V., Defrance, A. & Kahn, S. (2006). *Les compétences à l'école: apprentissage et évaluation*. Bruxelles : De Boeck.
- Scallon, G. (2007). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck & Larcier.