
La construction des standards en Suisse

Au-delà des engagements politiques : le travail dans les disciplines

Daniel Elmiger, Martine Wirthner, Jean-François de Pietro, Luc-Olivier Pochon

*Institut de recherche et documentation pédagogique IRDP
Fbg de l'Hôpital 43
CH-2002 Neuchâtel
daniel.elmiger@irdp.ch
martine.wirthner@irdp.ch
jean-francois.depietro@irdp.ch
luc-olivier.pochon@irdp.ch*

RÉSUMÉ. En Suisse, quatre consortiums réunissant des chercheuses et chercheurs des différentes régions linguistiques sont chargés d'établir un modèle de compétences et une proposition de compétences minimales pour la langue de scolarisation (L1), les langues étrangères (L2, L3), les mathématiques et les sciences. Le travail concret d'élaboration des modèles de compétence et des outils permettant de les valider soulève un grand nombre de questions méthodologiques et théoriques, et il a aussi une influence sur le l'égalité des élèves par rapport à ce qui est visé par les standards.

MOTS-CLÉS : Suisse, standards, niveaux de compétence, mathématiques, langue de scolarisation, langues étrangères.

1. Ouverture

La contribution « Harmonisation et équité du système éducatif suisse » a mis en évidence les enjeux et modalités de divers changements actuellement en cours dans le système éducatif suisse, et en particulier de ceux liés au projet HarmoS qui est explicitement centré sur la recherche d'une plus grande harmonisation structurelle des systèmes cantonaux, d'une mesurabilité des performances atteintes permettant un monitoring du système et d'une plus grande égalité de ces performances, qui serait assurée par l'établissement de standards minimaux que l'ensemble de la population scolaire devrait atteindre.

1.1. Quelques remarques sur la question des classes sociales face aux mathématiques

La question de l'égalité des chances n'est – bien sûr – guère nouvelle. Sans vouloir remonter à Mathusalem, le problème de l'égalité des classes sociales face aux mathématiques, par exemple, est récurrent. Il a été rapidement constaté, voire prévu, lors du renouvellement de l'enseignement des mathématiques dans les années 1970, ceci notamment en liaison avec les « réclamations » des parents qui n'étaient pas (plus) capables d'aider leurs enfants soit de façon informelle, soit dans le cadre des devoirs à domicile. Et ceci d'autant moins qu'ils étaient plus éloignés socialement du monde scolaire ou académique. Ce problème est réapparu lors de modifications ultérieures des programmes et des contenus.

Des solutions partielles ont été trouvées à travers l'organisation de cours dans le cadre de l'école des parents, la diffusion de dépliants informatifs et de supports multimédia. De même, des moyens d'enseignement complémentaires avaient été créés pour les classes rurales.

Puis la question s'est posée de savoir si ce problème d'une certaine inégalité face aux contenus se reflétait également à travers les niveaux de réussite à des tests de connaissance. De fait, l'effet de la classe sociale sur le niveau de réussite à des tests de connaissance est une constante. Il n'a pas manqué d'être mis en évidence lors des évaluations menées par l'IRDPA que ce soit lors de la première vague d'évaluation ou plus récemment dans le projet MathEval (Antoniotti, 2003, 2005).

Lors de la première phase d'évaluation (Rosenberg, 1982), une étude plus approfondie a cependant été entreprise, pour la fin du cycle moyen, qui nuance un peu le propos. Si globalement, la tendance est toujours marquée, les différences peuvent varier selon les domaines et les types de tâches. L'habillage verbal notamment peut avoir une influence sur les écarts mesurés. Lorsque l'habillage se réfère à des activités proches de la vie courante, les écarts étaient moindres. Il en va de même lorsque les données des problèmes étaient épurées.

A l'époque, il avait été souhaité de procéder à des vérifications des hypothèses élaborées à partir de ces premiers résultats. Cela n'a pas été fait. Par contre l'attention d'un groupe chargé de mettre à jour les plans d'études, a été attirée sur le

problème. La notion de cadre fonctionnel a été introduite comme moyen d'équilibrer les contextes des activités mathématiques proposées aux élèves.

Par la suite, ces propositions n'ont été que partiellement retenues, les auteurs préférant vraisemblablement offrir une approche plus homogène de l'enseignement mathématique. Ils ont également misé sur les retombées d'une approche didactique rigoureuse dans le processus de génération des savoirs mathématiques, qui empêcherait en quelque sorte les effets sociaux. La prise en compte des aspects socio-affectifs dans la nouvelle génération de moyens d'enseignement est un sujet qui semble pour le moment avoir été peu abordé.

L'introduction de standards minimaux va assurément remettre cette question à l'ordre du jour. Dans le cadre de l'opération HarmoS, il s'avère en effet qu'il est difficile de s'en tenir à la définition d'un noyau de savoirs minimaux maîtrisés par tous. Les modèles utilisés pour la définition de standards nécessitent d'introduire des niveaux de difficultés dans les tâches proposées. Il s'ensuivra inmanquablement la mise en évidence de différences de niveau de réussite entre différentes populations (peut-être liées à des différences de classes sociales, mais dues aussi à d'autres facteurs : sexe, milieu, langue d'origine, etc.). S'ensuivront, aussi, différentes questions à propos des écarts qui ne manqueront pas d'être constatés :

- quelle est la part des écarts due à l'artefact ?
- quelle est la part « essentielle » (relevant du pur domaine mathématique) ?
- que savons-nous des mécanismes de production de ces écarts ?
- que sera-t-il entrepris pour y remédier ?

Il est d'autant plus urgent de s'intéresser à ces questions que les résultats obtenus aux tests pourraient avoir une influence sur le passage des élèves d'un cycle à l'autre...

Dans notre intervention, nous nous intéressons plus spécifiquement au travail concret réalisé au sein des disciplines – en particulier les langues secondes et la « langue de scolarisation » ainsi qu'on nomme la langue locale dans ce contexte – dans le but d'élaborer les outils permettant la réalisation du projet HarmoS. Dans ce projet, si la recherche doit se mettre au service du politique, cela ne permet cependant pas – même si certains le souhaiteraient peut-être – de faire l'économie de « vraies » interrogations portant sur les difficultés de l'opération :

- Comment définir ce que les élèves doivent savoir (faire) ? Quelle(s) modélisation(s) proposer ?
- Comment définir des niveaux de standards ?
- Peut-on définir des standards minimaux sans entrer dans une logique de la différenciation de plusieurs niveaux ?
- Comment observer ces niveaux de manière fidèle ?

Et nous évoquerons également d'autres interrogations, subséquentes, qui portent elles sur les effets des données « scientifiques » ainsi obtenues : comment les interpréter ? comment décider des mesures – politiques ? pédagogiques ?

didactiques ?... – à prendre ? Que devient la question de l'égalité des chances et/ou des résultats ? Etc.

2. Mesurer pour assurer l'égalité, mais mesurer quoi ? comment ?

2.1. Vers des modèles de compétence

Comme dans la plupart des approches actuelles, le projet HarmoS met en avant la notion de compétence (plutôt que celle de ressource, ou de capacité, ou de savoir) pour cerner ce qu'il s'agit de mesurer dans les différentes disciplines. Les différents consortiums (mathématique, sciences naturelles, L2, L1) ont par conséquent dû commencer leur travail en tentant d'élaborer un tel modèle de compétences valable pour leur discipline, voire en reprenant un modèle déjà existant. Il s'agissait donc – ce qui n'est cependant pas une mince affaire – de se mettre d'accord sur une définition de la compétence. Le projet HarmoS s'est finalement appuyé sur un concept de compétence unifié, issu des réflexions formulées par Weinert (1999, 2001), que les différents consortiums disciplinaires n'ont guère eu l'occasion de discuter... En accord avec Weinert (2001, p. 27-28), il faut donc entendre par compétences « les capacités et les savoir-faire cognitifs dont un individu dispose ou qu'il peut acquérir pour résoudre des problèmes spécifiques, ainsi que les dispositions et les capacités motivationnelles, volitives et sociales qui y sont liées pour appliquer les solutions aux problèmes avec succès et de manière pleinement responsable dans diverses situations. »

Selon cette définition, la compétence est une disposition qui permet aux personnes qui la possèdent de résoudre avec succès certains types de problèmes, c'est-à-dire de maîtriser les exigences situationnelles concrètes d'un certain type (Klieme *et al.*, 2003, p. 72). Une telle définition, mettant en avant la compétence comme disposition et le produit qui en résulte plutôt que le processus de réalisation de tâches complexes mériterait une discussion approfondie. Elle se justifie vraisemblablement, voire s'impose, dans le cadre d'un projet dont la finalité est de définir des modèles à propos desquels des mesures peuvent être effectuées et sur la base desquels des standards minimaux peuvent ensuite être fixés et vérifiés. En effet, dans la perspective du projet HarmoS, les standards doivent être décrits en termes de performances par rapport à des tâches et / ou situations ; ils correspondent au produit de la compétence dans une situation donnée ; leur établissement passe par conséquent par la définition de tâches se référant au modèle de compétence et l'observation des performances des élèves qui permettra en définitive de vérifier la validité du modèle même. Comme on le voit déjà ici, sans qu'une véritable discussion ait eu lieu, ce sont les finalités politiques du projet – et l'exigence de la mesure qui les sous-tend – qui ont orienté dans une large mesure le choix d'un modèle théorique.

2.2. *L'exemple du modèle de la langue de scolarisation*¹

L'élaboration du modèle s'est appuyée d'abord sur une conception de la langue. Celle-ci peut être décrite comme un complexe constitué de deux domaines qualitativement distincts: le premier domaine est désigné habituellement par le concept de « langue en tant que système », le second par celui de « langue en usage ». Tous deux sont concernés par les processus de développement et d'apprentissage, le premier parce qu'il délimite – hors contexte – ce qui est potentiellement possible, le second parce qu'il décrit ce qui est effectivement réalisé, sur un double continuum allant l'un d'un pôle « échec » vs « réussite » par rapport à la tâche à effectuer, l'autre d'un pôle « conforme à l'usage » à un pôle « non conforme à l'usage » sur un axe relevant des normes sociales (qui doivent bien sûr être définies en tenant compte des situations de communication).

En liaison avec la conception mentionnée ci-devant de la compétence, c'est ce second domaine, des usages effectifs, que nous avons privilégié. Le modèle de compétences en « langue l » qui en découle s'appuie donc sur une approche pragmatique et fonctionnelle de la langue et de son usage. Et la compétence de langue, en tant que disposition, est dès lors définie comme une disposition pluridimensionnelle et socialement située de la cognition humaine, caractérisée par une interaction constante de plusieurs domaines cognitifs (Fodor, 1975; Jackendoff, 1997). Une telle définition a immédiatement des conséquences importantes sur la mesure de la compétence langagière qui ne saurait dès lors être évaluée en additionnant des mesures de savoir-faire partiels (Teilfertigkeiten). De plus, les compétences se manifestent concrètement dans des situations thématiquement et socialement signifiantes. Ce fait doit également être pris au sérieux lors de l'élaboration du modèle (puis des outils de mesure), qui ne peut prétendre « refléter » les compétences sans tenir compte du contexte situationnel et social d'une activité langagière, de son contenu et de sa forme spécifique (genre).

Le modèle de compétences élaboré par le consortium L1 comporte trois niveaux. Les actions langagières constituent le point de départ de la modélisation. Elles sont réalisées dans différentes modalités (production ou réception) et font appel à différents médias (oral et écrit, communication directe et indirecte). L'action langagière peut être décrite par les différents aspects qui la constituent, à des degrés de nécessité divers selon les modalités et les médias. Dans cette perspective, nous considérons d'abord les aspects situer, planifier et réaliser. A cela s'ajoute que, en cours même de réalisation ou une fois réalisée, l'action langagière peut être évaluée et réparée. Le modèle comporte donc cinq aspects qui devraient permettre de

1. La langue de scolarisation n'est pas la langue maternelle ! C'est ici celle qui est utilisée à l'école, celle de la société locale, sachant que pour bon nombre d'élèves, il ne s'agit pas de la langue principale de la famille. Entre autre, il faut savoir qu'en Suisse alémanique, des dialectes sont utilisés dans la vie courante et non l'allemand standard, appris et utilisé principalement à l'école. Dans le cadre de HarmoS, trois langues nationales sont prises en considération : l'allemand, le français et l'italien.

structurer l'ensemble des manifestations possibles de l'activité langagière en situation². C'est le premier niveau du modèle.

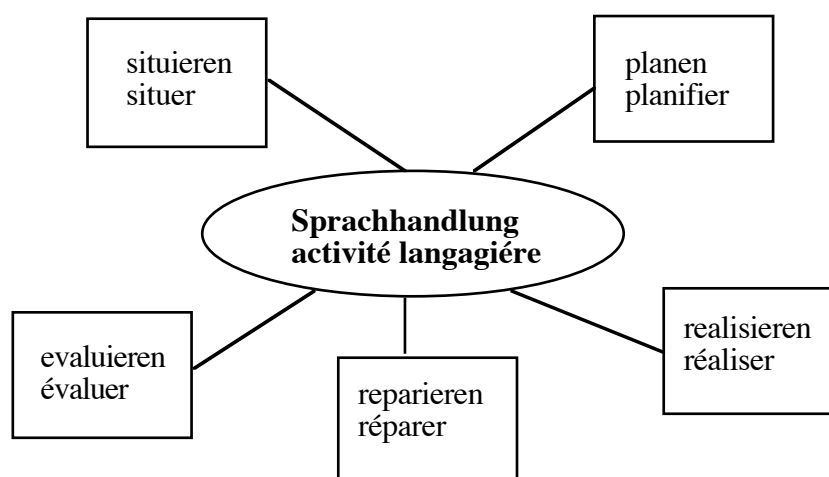


Figure 1. Les cinq aspects de l'activité langagière

Ces aspects (abstrait) de l'activité langagière doivent dans un second temps être mis en relation avec les modalités matérielles de l'action langagière, réparties traditionnellement en quatre sous-domaines: parler et écouter, lire et écrire³. Ainsi s'élabore le deuxième niveau du modèle⁴.

2. Il faut cependant se garder d'interpréter – au sens de la psychologie cognitive – ces cinq aspects comme les étapes successives d'une action langagière complexe.

3. Pour des raisons qui tiennent essentiellement aux difficultés pratiques de mise en œuvre et de mesure, il a été renoncé, pour l'oral, à faire de l'interaction un sous-domaine spécifique.

4. D'autres paramètres, secondaires dans notre approche, influent cependant eux aussi sur la configuration, socialement située, des compétences langagières : le fait que la réalisation langagière, au-delà de sa matérialité effective, soit issue d'une « conceptualisation » orale ou écrite, permettant d'estimer la pertinence de l'activité langagière en fonction d'un degré de proximité à la situation de communication ; le processus-le produit, précisant si l'accent est porté sur le résultat de l'activité ou davantage sur la manière dont le locuteur, le lecteur ou le scripteur s'y prend pour réaliser cette activité ; les genres de texte, mettant en avant la compétence à tenir compte des caractéristiques langagières et du contexte socio-discursif propres aux textes présents dans la société ; la réflexion sur la langue, amenant le locuteur, le lecteur ou le scripteur à être capable de dire quelque chose de l'activité langagière dans laquelle il se trouve ou du langage qu'il utilise ou auquel il est confronté, etc.

2.2.1. D'autres modélisations possibles

Il importe de relever que d'autres modélisations auraient été envisageables, prenant davantage en compte, par exemple, les connaissances hors contexte des sujets. Mais ce modèle, à base fonctionnelle et socio-cognitive, nous est apparu comme le mieux approprié aux exigences du projet. En mathématique, en revanche, il est intéressant de voir que le modèle élaboré repose sur une logique apparemment différente : il croise des aspects de compétences – capacités mathématiques telles que comparer, mesurer, etc. –, correspondant dans une certaine mesure aux aspects retenus pour les langues, et des domaines mathématiques – un peu comme si on associait les aspects des compétences langagières aux contenus verbaux concernés... De telles différences sont intéressantes et – une fois encore – nous interrogent, épistémologiquement, sur les modalités d'élaboration des modèles.

2.2.2. Des modèles aux tâches / tests qui permettent d'en évaluer la validité

A un troisième niveau du modèle en L1 (mais sur ce point il en va de même pour toutes les disciplines) se situent les tâches dans lesquelles les compétences se déploient, se réalisent (cf. conception pragmatique, fonctionnelle de la langue). C'est donc à ce niveau qu'il devient possible d'évaluer les élèves et de valider le modèle de compétences préalablement construit.

Aspects de l'activité langagière (niveau 1)		Domaines spécifiques de l'activité langagière (niveau 2)			
		Ecouter	Parler	Lire	Ecrire
	Situer	Tâches (niveau 3)	Tâches	Tâches	Tâches
	Planifier	Tâches	Tâches	Tâches	Tâches
	Réaliser	Tâches	Tâches	Tâches	Tâches
	Réparer	Tâches	Tâches	Tâches	Tâches
	Evaluer	Tâches	Tâches	Tâches	Tâches

Tableau 1. *Modèle de description des compétences de langue à trois niveaux*

Pour L1, dans les différents sous-domaines – lire, écrire, écouter, parler et interagir –, des épreuves ont alors été élaborées pour les 6e et 9e années de la scolarité obligatoire. Les tâches s'inscrivent dans des genres de textes divers, en tenant compte de leur regroupement en trois groupes : narratif – argumentatif –

expositif (expliquer, relater, décrire des actions)⁵. Les items des différentes épreuves renvoient à l'un ou à l'autre des aspects décrits pour chacun des sous-domaines.

Ceci est vite dit et paraît bien simple, comme s'il suffisait désormais de passer à l'acte... Pourtant, c'est en fait là que tout commence ! Car la difficulté consiste précisément à trouver des situations, des tâches et des contenus qui remplissent toutes les conditions pour devenir de bons ingrédients des épreuves : pertinence sociale, recevabilité, mesurabilité, faisabilité, etc.

Si de premiers essais ont pu être réalisés sur un petit échantillon de classes et d'épreuves, ils n'ont toutefois pas permis une réelle régulation du travail d'élaboration des tests. Les délais très courts imposés par le projet ont obligé à produire de très nombreux items sans avoir suffisamment le temps de les discuter, de les essayer et de les améliorer. Le choix méthodologique d'une organisation selon les 4 sous-domaines issus de la tradition et l'impossibilité d'une collaboration entre les groupes (précisément organisés par sous-domaines), en raison du manque de temps, ont empêché d'établir, pour plusieurs épreuves, des liens entre lire et écrire, entre parler et lire, etc. Et à ces obstacles bien prosaïques viennent s'ajouter d'autres contraintes encore, d'ordre méthodologique, liées aux modèles statistiques privilégiés (choix du modèle de Rasch ; cf. infra).

Dans ces conditions, en fonction des tâches que nous sommes parvenus à mettre en place dans le cadre des tests, il apparaît clairement que ni les différents sous-domaines, ni les différents aspects du modèle de compétences ne peuvent être évalués de manière totalement égale et comparable. Tous ne se prêtent en effet pas de manière égale à une évaluation dans le cadre des tâches élaborées, qui doivent satisfaire aux exigences de la mesure. La mise à l'épreuve de notre modèle ne peut par conséquent avoir l'ambition de prétendre à une validation totale du modèle. Cela est inévitable et certainement normal. Toutefois, un certain nombre de questions plus problématiques restent posées, sur lesquelles nous reviendrons en conclusion.

Au préalable, deux mots sur la passation de ces épreuves et sur les modèles de mesure auxquels il est fait appel. Mais, auparavant encore, une remarque : nous entrons ici dans la présentation d'une phase de la recherche qui est en cours. Nous sommes alors présentement trop engagés dans l'urgence pour pouvoir prendre la distance suffisante à une réflexion épistémologique. C'est pourquoi nous ne ferons – pour le moment – qu'évoquer les points qui suivent, réservant une réflexion plus approfondie lors de notre communication.

Une validation empirique large (environ 13'000 élèves) s'est déroulée en mai 2007, pour des classes de 6e et de 9e. Une validation moins large (1'400 élèves) a eu

5. Cette notion de regroupement de genres s'inspire de celle développée dans le moyen d'enseignement romand « S'exprimer en français » (2001) qui propose une didactique de la production écrite et orale basée sur le travail de genres de textes. Pour des raisons essentiellement pratiques, cinq regroupements de genres y sont réalisés autour des capacités dominantes suivantes : narrer, relater, argumenter, exposer et décrire des actions.

lieu en 2e année primaire. Les résultats doivent permettre de valider le modèle de compétence et de créer les standards (minimaux).

2.2.3. *Ce que dit la mesure : le modèle de Rasch*

Le traitement des données sera effectué en recourant au modèle de Rasch. Selon ce modèle statistique, les tâches retenues doivent présenter des items discriminants (permettant d'obtenir des résultats de mauvais à très bons), ce qui permettra une double distribution : des items d'une part (faciles à difficiles), des élèves d'autre part (avec de bons, de moyens et de mauvais résultats). Ainsi, il sera possible de déterminer pour chaque item son niveau de difficulté et le pourcentage de réussite et d'échec qu'il occasionne. Aux chercheurs ensuite d'indiquer lesquels retenir pour construire les standards (pour chaque sous-domaine, pour la 2P, la 6e et la 9e année).

Dans ces conditions, la tâche des chercheurs sera pour le moins délicate : sur quels critères fixer le standard minimal ? Devra-t-on admettre qu'il existe des élèves qui ne l'atteignent pas ? Combien ? Et comment tenir compte des différences entre les filières de l'enseignement secondaire ? Les résultats de l'enquête PISA (Nidegger, 2001 ; 2005) avaient montré qu'il existe des zones de recoupement des résultats des élèves de ces différentes filières ; comment procéder alors ?

Par ailleurs, le nombre relativement élevé de questions à choix multiples – elles aussi suscitées par les modalités de la mesure – témoigne que bien des tâches ne sont pas – loin s'en faut – des tâches complexes de résolution de problème, telles que la définition de la compétence définie a priori les évoquait.

Comme on le voit, l'interprétation statistique des résultats soulève elle aussi de nombreuses interrogations.

2.3. *Le modèles pour les langues étrangères (L2, L3)*

En ce qui concerne le consortium chargé d'élaborer un modèle de compétence et des propositions de standards pour les langues étrangères (L2 et L3), il se voit confronté, dans l'ensemble, à des difficultés qui ressemblent en de nombreux points à ceux du consortium L1. Cependant, l'établissement d'un modèle de compétence pour les langues L2 et L3 (en général une langue nationale et l'anglais) s'avère à la fois plus simple et plus complexe que celui pour la langue de scolarisation. D'un côté – étant donné que le consortium pour les langues étrangères peut (et a même pour mandat de) s'appuyer sur un modèle existant, le *Cadre européen commun de référence pour les langues* (Conseil de l'Europe, 2000) –, il peut s'appuyer sur des travaux fondateurs qui ont fait leurs preuves au niveau théorique et pratique.

De l'autre côté, le consortium pour les langues étrangères doit affronter un certain nombre de problèmes pratiques et organisationnels particuliers. Cela concerne d'une part l'enseignement que les élèves impliqués dans la phase de validation ont suivi durant leur scolarité : actuellement, les items de test doivent être

validés avec une population d'élèves qui a bénéficié d'une durée d'enseignement très variable d'une région linguistique (et, parfois, d'un canton) à l'autre.

En plus, l'enseignement même des langues en Suisse se trouve dans une phase de réforme visant à moyen terme un enseignement de deux langues étrangères dès la 3^e et la 5^e année de l'école primaire. Cette réforme, aboutie dans certains cantons, n'en est encore qu'à la phase de préparation dans d'autres. Il s'avère ainsi que le modèle de compétence – ainsi que les standards qui en découlent – est actuellement validé à l'aide d'une population d'élèves qui ne correspond que fragmentairement à celle qui aura suivi le cursus prévu par la nouvelle politique linguistique. Il est dès lors difficile à prévoir dans quelle mesure les résultats HarmoS pour les langues étrangères peuvent déboucher sur des standards (minimaux) généralisables.

D'autre part, les tests même élaborés par le consortium L2/L3 ont dû être élaborés en deux langues scolaires (allemand, français), pour trois langues étrangères (français, allemand, anglais). Outre le fait que la multiplication des configurations possibles pose des problèmes pratiques considérables, elle a aussi une incidence sur la mesure même, car il n'est pas possible d'exclure certains effets parasites, causés par la traduction des consignes ou des items de test. Cela a pour conséquence que tous les items de test doivent être codés en fonction de leur provenance (texte original/texte traduit), et l'analyse statistique aura comme tâche supplémentaire de déceler les possibles effets imputables à la traduction.

Une autre différence notable par rapport au modèle de compétence prévu pour la L1 réside dans la construction même du *Cadre européen commun de référence*. Si le premier niveau du modèle pour la langue de scolarisation, l'activité langagière, se retrouve plus ou moins au niveau des descripteurs des niveaux de compétences du *Cadre* et si les quatre sous-domaines se recoupent largement avec ceux pertinents pour les langues étrangères (qui différencie, à l'oral, entre l'expression orale continue et les échanges en interaction), le niveau suprême du modèle de compétence élaboré pour la langue de scolarisation ne figure pas dans celui pour les L2 et L3.

Il va de soi que les éléments y figurant (la situation, la planification, la réalisation, l'évaluation et la réparation, cf. tableau 1 et figure 1) sont inhérents à toute action langagière, quelle que soit la langue. Cependant, les aspects de l'activité langagière, fondateurs pour la L1, demeurent complètement implicites dans le modèle pour les langues L2/L3. Cela n'est évidemment pas sans avoir une influence sur les modèles heuristiques et les résultats pratiques dans les deux consortiums. Ainsi, au niveau des standards, on peut légitimement se demander jusqu'à quel degré un modèle pour la langue de scolarité peut différer de celui pour les langues étrangères. Cette question n'a non seulement des incidences au niveau des plans d'étude (lisibilité, comparabilité des objectifs poursuivis), mais aussi à celui des objectifs didactiques. Depuis quelques années, les spécialistes de l'enseignement des langues se prononcent en faveur d'une didactique intégrée des langues, dépassant les frontières traditionnelles entre les langues et leurs enseignements. Ce but serait à poursuivre à plusieurs niveaux : la formation des enseignant-e-s, les plans d'étude, les curricula des élèves, le travail concret en classe, etc. L'objectif d'une didactique

intégrée des langues n'est pas facilitée si les modèles de compétences pour les langues (L1 vs. L2/L3) reposent sur des principes divergents. Dans le domaine des langues, les travaux de concrétisation des résultats du projet HarmoS (p. ex. au niveau des plans d'étude) devront ainsi accorder un souci particulier à l'harmonisation des modèles de compétence et des standards.

3. Des perspectives intéressantes... mais beaucoup de questions

Nos préoccupations se déclinent en plusieurs questions :

Les standards devront donc être définis sur la base des résultats obtenus aux tests et réinterprétés selon le modèle de compétence et les items utilisés afin de ne pas se limiter à un simple score. Ensuite, il n'est en fait pas encore très clair s'il s'agit de standards minimaux ou de standards de formation et ce qui les distingue. En effet, la méthode de mesure appliquée permet de répartir la population étudiée sur un continuum qui, comme dans le cas de PISA, contiendra différents niveaux de compétence. Selon Mons et Pons (2006), le standard du projet HarmoS est fortement polysémique ; par exemple, ces standards doivent certes faciliter la rédaction de plans d'études (visée curriculaire du projet HarmoS), mais aussi améliorer l'évaluation par l'entremise de tests standardisés. Selon le projet, ils peuvent encore – et surtout – remplir la fonction d'indicateurs pour un monitoring du système scolaire à niveau national.

Quelle sera la forme finale des standards ? Une valeur abstraite ? Une liste de compétences ? Que deviendront les contenus d'enseignement qui s'inscrivent moins bien dans cette logique de la mesure ?...

Quelles incidences aura le choix des aspects de compétence finalement retenus ?

— Comme la plupart des capacités, leur nature extrêmement transversale veut qu'ils ne relèvent pas d'apprentissages scolaires uniquement, ce qui peut révéler des différences socio-culturelles influençant fortement la réussite des élèves.

Et le choix des domaines mathématiques ou des situations communicatives finalement retenues ?

— Selon les visées des domaines mathématiques retenus, des situations langagières, la distance par rapport aux pratiques de référence de chaque élève peut là encore induire une part d'iniquité.

Et quelles seront les conséquences pour les élèves ne répondant pas aux standards : des appuis ? des drills ? Est-ce là l'assurance vers plus d'équité ?...

D'autres interrogations portent plutôt sur les liens entre cette approche nouvelle, centrée sur la mesure de l'output et sur les standards, et les plans d'étude : quelle cohérence entre les modèles de compétences, les standards et les contenus des plans d'études ? Si, dans la logique de HarmoS, les standards doivent servir (entre autres) à la rédaction du plan d'études suisse alémanique, la distance entre eux et les exigences des plans en vigueur en Suisse romande pourrait être sensible. S'agira-t-il

alors d'adapter ceux-ci à ceux-là, alors même que le plan romand est en cours d'élaboration ? Une distance conséquente entre des standards nationaux, dont l'atteinte est mesurée par des épreuves elles aussi nationales, et les exigences scolaires explicitées dans les plans d'études régionales serait regrettable, nuisant autant à l'enseignement qu'à l'apprentissage. Elle pourrait conduire à un glissement de l'enseignement/apprentissage vers la seule résolution de tests type HarMoS, au détriment des orientations didactiques actuellement en vigueur dans les classes.

Dans ce sens, que dire d'un modèle de compétences élaboré d'un point de vue théorique, sans référence aux documents de l'école, voire aux pratiques scolaires ? De quoi rend-il compte réellement ? Sa validation repose sur les performances des élèves aux tests construits sur sa base ; ce qui signifie que les tests mesureraient bien ce que le modèle présente en matière de compétences. Or, les contraintes inhérentes à la forme des tests et à leur passation ne distordent-elles pas les compétences que l'on cherche à mesurer ? Une suite de tâches diverses – mélange entre L1 et L2, mélange entre compréhension orale, écrite et écriture, mélange d'items à QCM, à réponses courtes ou longues – que les élèves doivent effectuer en un temps limité (20 ou 30 minutes pour chacune des tâches) rend-elle véritablement compte des compétences des élèves à lire, écouter, écrire des textes de genres variés ? Et que dire pour les élèves qui ne sont pas familiarisés avec de telles tâches et qui sont habituellement en difficulté dans les disciplines évaluées par ces tests ? N'observe-t-on pas ainsi une distance entre modèle de compétences et épreuves ?

C'est toute la question de la place des didactiques disciplinaires qui est implicitement posée ici. Où situer la didactique des disciplines dans la création du modèle de compétences ? Et quelles conséquences les standards pourront avoir sur la didactique ?

A l'heure actuelle du projet, il est encore trop tôt pour tirer un bilan de l'opération. Les difficultés ne manquent pas, comme nous l'avons vu, qui font du projet HarMoS un véritable défi pour les chercheurs.

3.1. Bibliographie

- Antonietti, J.-P. (Ed.) (2003). *Evaluation des compétences en mathématiques en fin de 2e année primaire, résultats de la première phase d'enquête Mathéval*. Neuchâtel : IRDP.
- Antonietti, J.-P. (Ed.) (2005). *Evaluation des compétences en mathématiques en fin de 4e année primaire, résultats de la première phase d'enquête Mathéval*. Neuchâtel : IRDP.
- Conseil de l'Europe (2000). *Cadre européen commun de référence pour les langues. Apprendre, enseigner, évaluer*. Strasbourg : Didier
- Fodor, Jerry A. (1975). *The language of thought*. Cambridge Mass. : Harvard University Press.
- Jackendoff, Ray (1997). *The architecture of the language faculty*. Cambridge Mass. : MIT Press.
- Klieme, E. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards, eine Expertise*, Bonn : BMBF.

- Mons, N., Pons, X., (2006). *Les standards en éducation dans le monde francophone, une analyse comparative*, Neuchâtel : IRDP .
- Nidegger, Ch. (Coord.) (2001). *Compétences des jeunes romands. Résultats de l'enquête PISA 2000 auprès des élèves de 9e année*. Neuchâtel : IRDP.
- Nidegger, Ch. (Coord.) (2005). *PISA 2003 : Compétences des jeunes romands. Résultats de la seconde enquête PISA auprès des élèves de 9e année*. Neuchâtel : IRDP.
- Rosenberg, S. (1982). *Acquisitions en mathématiques selon les sections scolaires et l'origine socio-économique*. Neuchâtel : IRDP.
- Weinert, F. E. (1999). *Concepts of competence (Contribution within the OECD project Definition and selection of competencies: Theoretical and conceptual foundations (DeSeCo)*, Neuchâtel : DeSeCo.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In Weinert, F. (Ed.) *Leistungsmessung in Schulen*. Weinheim und Basel : Beltz Verlag.