XIX^èmes Journées Internationales de Psychologie Différentielle Marseille, 25-27 août 2010

Evolution des connaissances structurales du mot et des performances en lecture chez des enfants de 6 ans

Jean Ecalle, Jacques Juhel**, Annie Magnan**

** Laboratoire EMC EA 3082 - Université Lyon(2)*
*** CRPPC/LPE EA 1285 - Université Rennes 2*

De la théorie de la triple forme du mot (Berninger et al, 2009) …

- Apprendre à lire: identifier les mots écrits de façon autonome
- Trois connaissances linguistiques sont indispensables pour identifier et manipuler différentes unités constitutives des mots
  - les cces phonologiques
  - les cces morphologiques
  - les cces orthographiques

… aux 4 composantes de base nécessaires pour apprendre à lire les mots

- Les connaissances sur les lettres (nom et phonème):
  - un puissant prédicteur de la réussite en lecture (Foulin, 2005)
  - un indicateur de début de la compréhension du système alphabétique (Treiman, 2006)
- Apprendre à lire:
  - coordonner les différentes connaissances sur la structure linguistique du mot
  - savoir utiliser le système alphabétique (les lettres)

Une composante cognitive ?

- La capacité de dénomination rapide (RAN: Rapid Automatized Naming): un autre prédicteur de la réussite en lecture ?
- Des résultats controversés:
  - Oui: la contribution du RAN augmente avec l'âge (Kirby et al, 2003); oui chez les lecteurs débutants avec une faible contribution toutefois (Badian, 2000).
  - Non: dans l'étude de Roman et al (2009) avec des enfants de CM1, 6ème et 4ème où comparativement les cces phonologiques, morphologiques et orthographiques apportent une contribution significative

Une étude longitudinale au CP

- Mesures répétées au moment de l'apprentissage formel de la lecture (CP) pour étudier l'évolution des performances dans les 4 composantes (lettres, phonologie, morphologie, orthographe) + RAN
- Décrire simultanément l'évolution de ces capacités et celle de la lecture

Une étude longitudinale au CP

- Population: 60 enfants (pas de redoublants) au CP
- Plan à occasions fixes

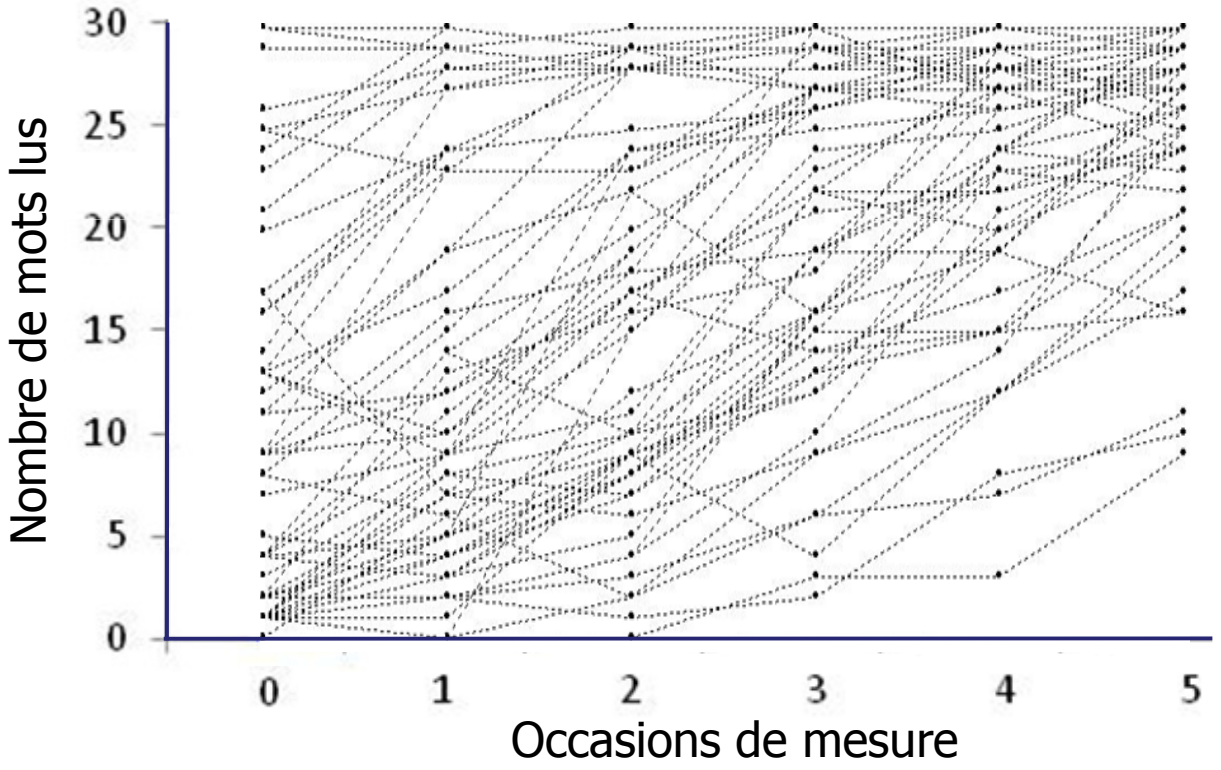
t_0	t_1	t_2	t_3	t_4	t_5
6-10/10	17-21/11	5-9/01	23-27/02	20-24/04	8-12/06
0	6 sem.	7 sem.	7 sem.	8 sem.	7 sem.

Tâches employées

Domaine	Tâche	Label	Exemple	Score
Phono	Détection d'intrus	PISY	sabot-sapin-mouton	/8
		PIPH	pomme-pince-ruche	/8
	Extraction d'unité	PESY	sabot-sapin	/8
		PEPH	pomme-pice	/8
Morpho	Jugement relation morphologique	MI	chaton-chat, riz-rideau	/8
	Extraction de base	ME	rejouer -> ; pêcheur ->	/8
Ortho	Détection d'intrus	OI	tillos-tiilos ; rtinos-trinos	/8
	Extraction d'unité	OE	tiilos	/8
Lettre	Dénomination de lettres	LET	en script, présentées isolément	/8
Dénom	Dénomination d'images	RAN		sec
Lecture	Lecture de mots voix haute	LECT	10 fréquents, 10 rares, 10 pseudo-mots	/30

+ 1 test d'intelligence non verbale administré avant l'expérimentation.

Evolution de la performance à l'épreuve de lecture de mots (N = 60)



Objectifs poursuivis

Identification :

1) des prédicteurs initiaux du niveau de performance et de la vitesse d'apprentissage (de t_1 *à* t_5) à la tâche de lecture de mots;

2) des variables endogènes dont les variations intra sont en lien avec celles de la variable-réponse LECT;

3) des variables endogènes susceptibles de rendre compte des variations intra dans la variable-réponse LECT.

Méthode d'analyse

Modèles multiniveaux de croissance

- Niveau 1 (intra-individuel : données répétées)

$$Y_{ti} = \pi_{0i} + \pi_{1i}\left(Temps_{ti}\right) + \pi_2\left(X_{ti}\right) + e_{ti}$$

$Temps_{ti}$ = (0, 7, 14, 22, 29)

π_{0i}, constante de la droite de régression pour l'individu *i*

π_{1i}, pente de la droite de régression pour l'individu *i*

X_{ti}, régresseur variant avec le temps (PISY, PIPH…, de t_1 à t_5)

e_{ti}, erreurs intra (écart entre l'observation au temps *t* chez un individu *i* et l'estimation correspondante de la droite de régression décrivant l'évolution temporelle de la performance de celui-ci)

Méthode d'analyse

Modèles multiniveaux de croissance

- Niveau 2 (inter: caractéristiques individuelles)

$$\pi_{0i} = \gamma_{00} + \gamma_{01} W_i + u_{0i}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11} W_i + u_{1i}$$

$$\pi_{2} = \gamma_{20}$$

γ_{00} et γ_{10}, constante et pente moyennes de la droite de régression décrivant l'évolution de Y_{ti} avec le temps

W_i, régresseur invariant avec le temps

u_{0i} et u_{1i}, effets aléatoires inter (écart d'un individu i à la constante et la pente moyennes dans la population de la droite de régression décrivant l'évolution de Y_{ti} avec le temps)

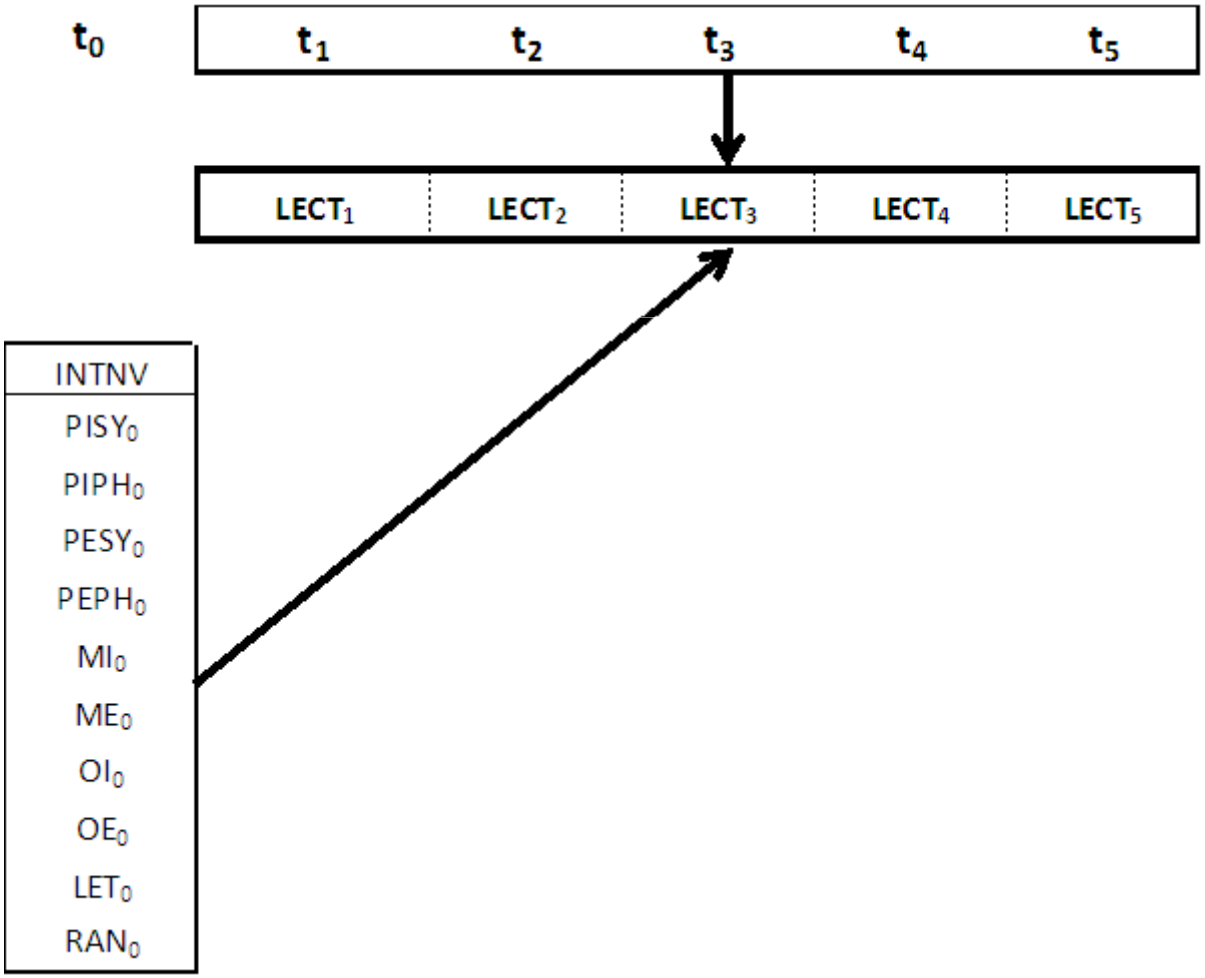
Méthode d'analyse

Modèles multiniveaux de croissance

- Combinaison des niveaux : substitution des équations de niveau 2 dans l'équation de niveau 1.
- <u>Hypothèses sur la structure de covariance des erreurs aléatoires intra et des effets aléatoires inter</u>
  - Niveau 1 : $e_{ti} \sim N(0, \sigma_e^2)$ (d'autres structures possibles : e.g. AR(1))
  - Niveau 2 : $\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u_0}^2 & \sigma_{u_{01}} \\ \sigma_{u_{10}} & \sigma_{u1}^2 \end{pmatrix}\right]$ (paramètres libres ou fixés)
- Estimation des paramètres par ML

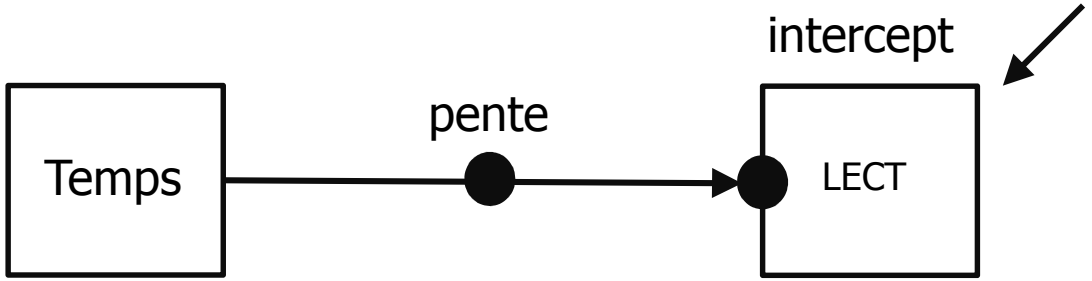
Modélisation 1

Identification des prédicteurs du niveau de performance et de la vitesse d'apprentissage à la tâche de lecture de mots;

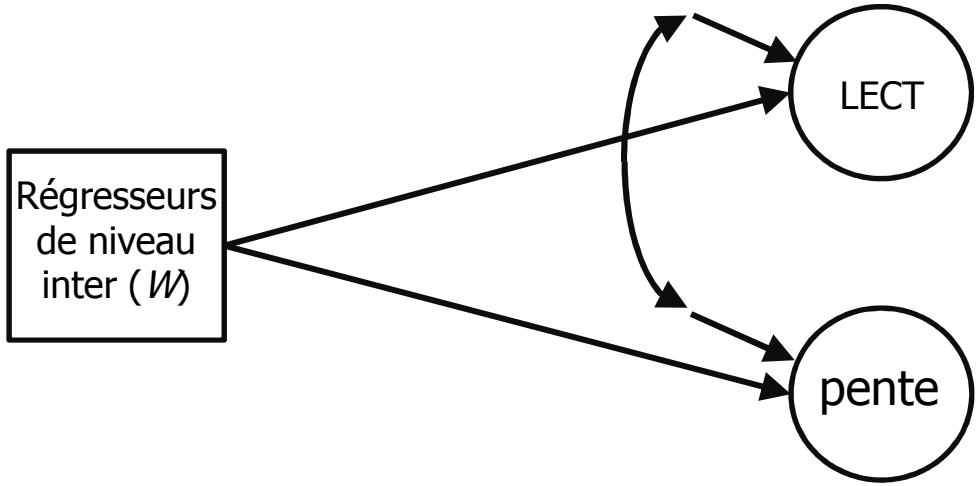


Modélisation 1

Représentation sous Mplus



1 : niveau intra



2 : niveau inter

Modélisation 1

Modèle de croissance avec effet aléatoire du temps et antécédents à t_0 (niveau inter), matrice de variance-covariance des effets aléatoires non structurée (estimation MLR)

Résultats du modèle

Corrélation intraclasse = 0,564

		Estimation	E.T.	*p*		Estimation	E.T.	*p*
Niveau 1	var. res. LECT	7,232	0,950	0,000				
Niveau 2	Intnv -> LECT	-	-	-	Intnv -> pente	0,089	0,042	0,034
	$PISY1_0$ -> LECT	-	-	-	$PISY1_0$ -> pente	-	-	-
	$PESY1_0$ -> LECT	-	-	-	$PESY1_0$ -> pente	-	-	-
	$PIPH1_0$ -> LECT	1,386	0,266	0,000	$PIPH1_0$ -> pente	-0,283	0,059	0,000
	$PEPH1_0$ -> LECT	1,386	0,266	0,000	$PEPH1_0$ -> pente	-0,283	0,059	0,000
	$MI1_0$ -> LECT	-	-	-	$MI1_0$ -> pente	-	-	-
	$ME1_0$ -> LECT	-	-	-	$ME1_0$ -> pente	-0,175	0,086	0,043
	$OI1_0$ -> LECT	-	-	-	$OI1_0$ -> pente	-	-	-
	$OE1_0$ -> LECT	-	-	-	$OE1_0$ -> pente	-	-	-
	$LET1_0$ -> LECT	0,707	0,146	0,000	$LET1_0$ -> pente	-	-	-
	$RAN1_0$ -> LECT	-	-	-	$RAN1_0$ -> pente	-	-	-
	LECT <-> pente	-5,112	1,362	0,000				
	var. res. LECT	30,051	5,820	0,000	var. res. pente	1,522	0,355	0,000

SABIC = 1685 (complet: 1712)

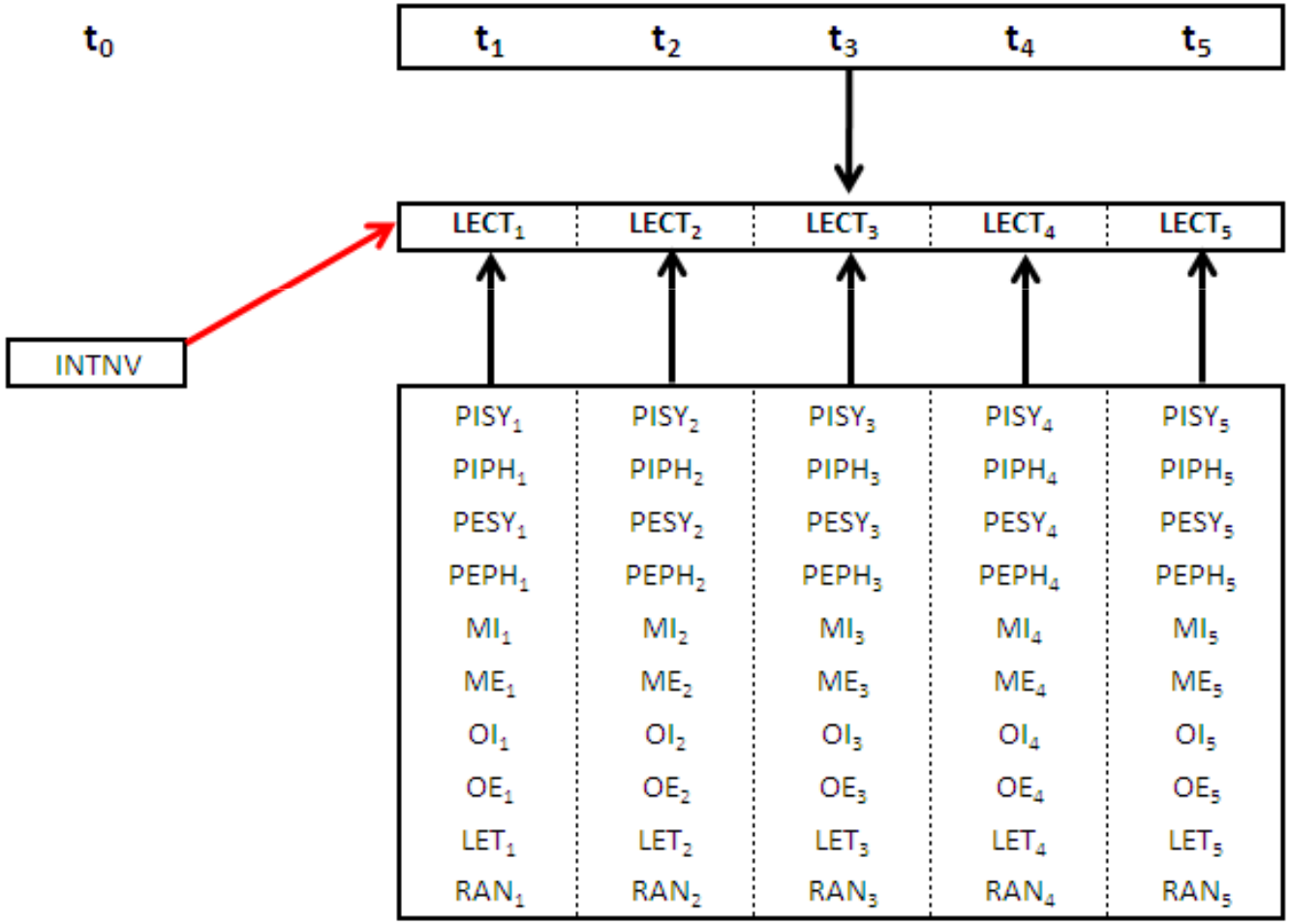
Modélisation 1

Discusssion

- La part de la variance interindividuelle de l'ensemble des mesures répétées (la variable-réponse LECT) représente 56,4% de la variance totale, celle de la variance attribuable à des variations temporelles chez un même élève 43,6%.
- Les niveaux initiaux de performance aux épreuves de connaissance phonologique (PIPH1 et PEPH1) et de connaissance des lettres (LET1) sont des prédicteurs significatifs du niveau initial de performance en lecture.
- Le niveau d'intelligence non verbale a un impact positif sur le rythme d'apprentissage de la lecture de mots.
- L'impact négatif du niveau de performance à certaines épreuves phonologiques (PIPH1 et PEPH1) et morphologiques (ME1) sur le rythme de croissance s'explique par la corrélation négative entre la constante et la pente de la droite qui décrit l'évolution temporelle de la performance à la tâche de lecture de mots.

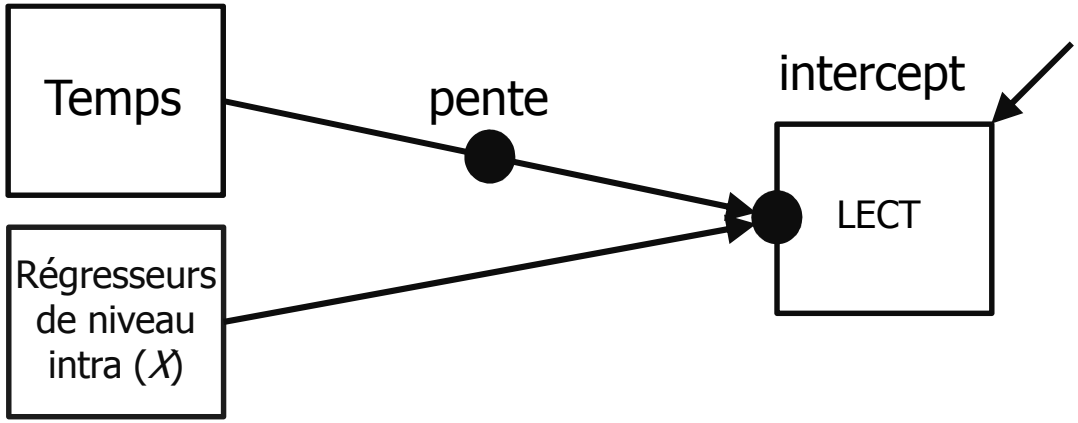
Modélisation 2

Identification des variables endogènes dont les variations intra sont en lien avec celles de la variable-réponse LECT;

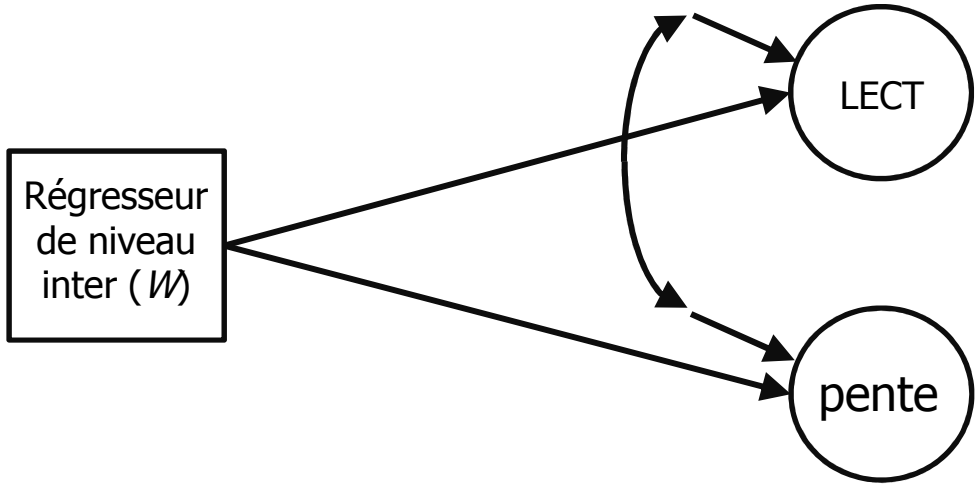


Modélisation 2

Représentation sous Mplus



1 : niveau intra



2 : niveau inter

Modélisation 2

Modèle de croissance avec effet aléatoire du temps, antécédent de niveau inter et régresseurs de niveau intra, matrice de variance-covariance des effets aléatoires non structurée (estimation MLR)

Résultats du modèle

Corrélation intraclasse = 0,160

		Estimation	E.T.	*p*
Niveau 1	$PISY_t$ -> $LECT_t$	0,431	0,138	0,002
	$PESY_t$-> $LECT_t$	-	-	-
	$PIPH_t$ -> $LECT_t$	-	-	-
	$PEPH_t$ -> $LECT_t$	0,690	0,140	0,000
	Mi_t-> $LECT_t$	-	-	-
	ME_t-> $LECT_t$	-	-	-
	OI_t -> $LECT_t$	-	-	-
	OE_t -> $LECT_t$	0,431	0,138	0,002
	LET_t -> $LECT_t$	0,431	0,138	0,002
	RAN_t -> $LECT_t$	-0,083	0,028	0,003
	var. res. LECT	7,339	0,991	0,000
Niveau 2	Intnv -> LECT	-	-	-
	Intnv -> pente	-	-	-
	LECT <-> pente	-9,169	2,076	0,000
	var. res. LECT	46,094	9,934	0,000
	var. res. pente	2,502	0,499	0,000

SABIC = 1699 (complet: 1711)

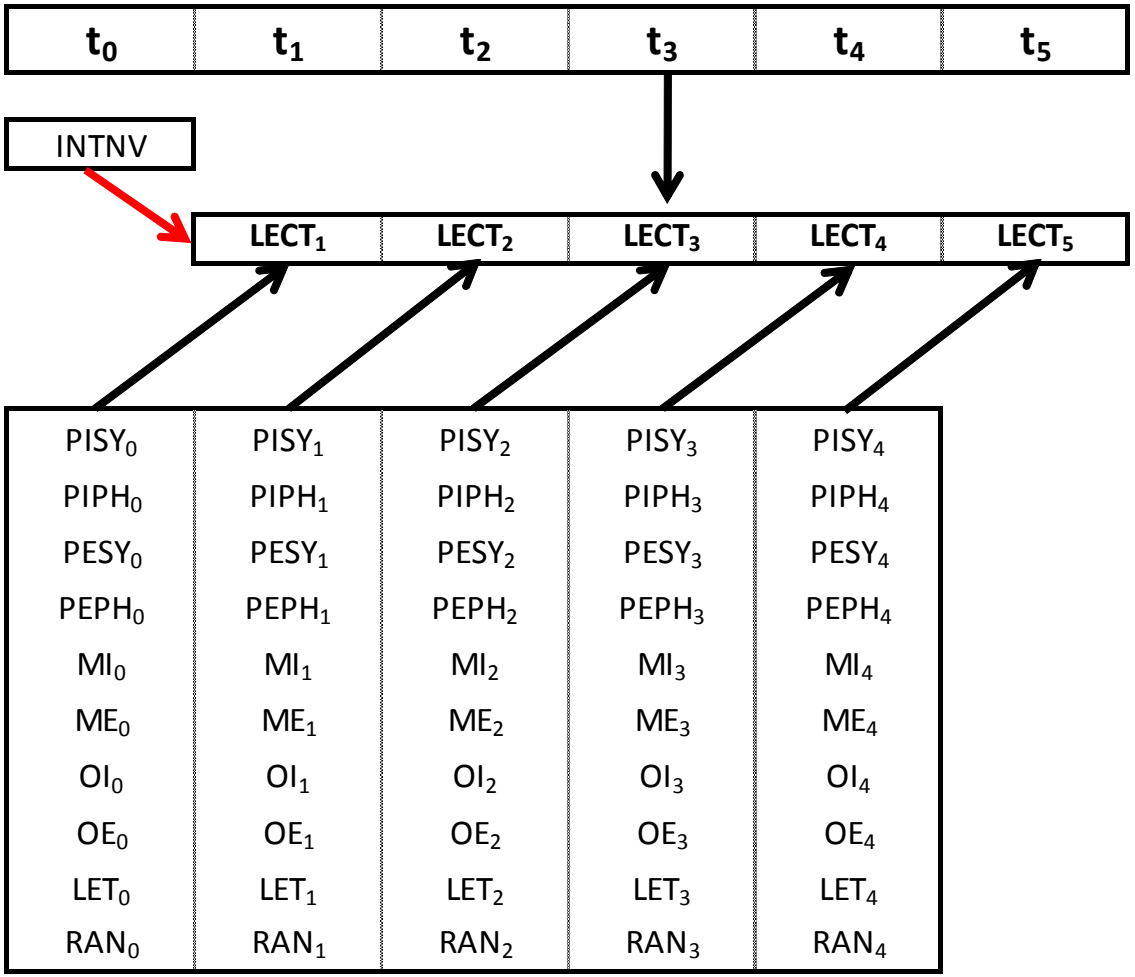
Modélisation 2

Discussion

- La part de la variance intra-individuelle de l'ensemble des mesures répétées (la variable-réponse LECT) représente 84% de la variance totale.
- Certaines variables mesurées à plusieurs reprises au cours de l'année sont significativement liées au niveau moyen de performance à la tâche de lecture. Les élèves dont les connaissances phonologiques (PEPH, PISY), des lettres (LET) et orthographiques (OE) ou la capacité de dénomination rapide (RAN) sont à *t* donné plus élevées que celles de leurs homologues, ont à *t* donné un meilleur niveau de performance en lecture que celui prédit par leur courbe de croissance.
- L'endogénéité des régresseurs interdit de conclure en termes de causalité (=> introduction d'un décalage temporel).

Modélisation 3

Identification des variables endogènes susceptibles de rendre compte des variations intra dans la variable-réponse LECT.



Modélisation 3

Modèle de croissance avec effet aléatoire du temps, antécédent de niveau inter et régresseurs de niveau intra avec décalage temporel, matrice de variance-covariance des effets aléatoires non structurée (estimation MLR)

Résultats du modèle

Corrélation intraclasse = 0,210

		Estimation	E.T.	*p*
Niveau 1	$PISY_{t-1}$ -> $LECT_t$	-	-	-
	$PESY_{t-1}$-> $LECT_t$	-	-	-
	$PIPH_{t-1}$ -> $LECT_t$	-	-	-
	$PEPH_{t-1}$ -> $LECT_t$	-	-	-
	Mi_{t-1}-> $LECT_t$	-	-	-
	ME_{t-1}-> $LECT_t$	-	-	-
	OI_{t-1} -> $LECT_t$	-	-	-
	OE_{t-1} -> $LECT_t$	-	-	-
	LET_{t-1} -> $LECT_t$	0,449	0,154	0,004
	RAN_{t-1} -> $LECT_t$	-	-	-
	var. res. LECT	7,339	0,991	0,000
Niveau 2	Intnv -> LECT	-	-	-
	Intnv -> pente	-	-	-
	LECT <-> pente	-11,828	2,188	0,000
	var. res. LECT	73,701	11,802	0,000
	var. res. pente	2,694	0,477	0,000

SABIC = 1733 (complet: 1748)

Modélisation 3

Discussion

- La variance attribuable à des variations temporelles chez un même élève représente 79% de la variance totale.
- Les élèves dont les connaissances phonologiques des lettres (LET) sont plus élevées que celles de leurs homologues au temps *t*, présentent au temps *t* +1 un meilleur niveau de performance en lecture que celui prédit par leur courbe de croissance.

Conclusions

- Cette étude montre la dynamique développementale de l'apprentissage de la lecture autour des 4 composantes fondatrices du développement des processus d'identification de mots écrits.
- L'importance de la connaissance des lettres est doublement réaffirmée en termes de niveau initial nécessaire pour apprendre à lire et prédictif de l'évolution des performances en lecture
- Les 4 composantes expliquent également l'évolution des performances en lecture, par ordre décroissant, phonologie, lettres, morphologie, orthographe
- Enfin, cette étude apporte un élément de réponse sur le rôle du RAN (capacité à dénommer rapidement des items) au cours du développement des processus d'identification du mot écrit en renforçant l'hypothèse d'un effet du niveau de performance en lecture sur la vitesse de dénomination.

Merci de votre attention

Les données de cette étude ont été recueillies par Sandra Berti, Romy Magnoli, Stéphanie Viel, étudiantes à l'ISTR (Ecole d'orthophonie, Université Lyon1)