

ÉTUDE SUR LES DIFFICULTÉS

EN SITUATION DE PRODUCTION D'ÉCRIT ET DE COMPRÉHENSION DE TEXTE CHEZ DES ÉLÈVES ET DES ADULTES SOURDS

Claire-Lise VELTEN

Dans le cadre de l'analyse des besoins pour le logiciel Vidéographix, nous avons réalisé des études afin de mieux prendre en compte les difficultés des personnes sourdes face à l'écrit. Deux catégories de population ont été soumises à des exercices : des adultes et des élèves de collège de deux établissements différents.

EXERCICE DE PRODUCTION D'ÉCRIT SOU MIS À CERTAINES CONTRAINTES

L'objectif était d'analyser la capacité des personnes sourdes à recréer de l'écrit selon différents modes de présentation de phrases.

Déroulement

Trois phrases complexes, constituées approximativement du même nombre de mots et du même nombre de propositions, ont été présentées selon trois modes différents: présentation à l'écrit, traduite en LSF ou contextualisée (condition de création et structuration interne) en LSF. Il était alors demandé de reconstituer les phrases à l'aide d'étiquettes créées avec le module étiquettes d'Idéographix.

Le mode de passation des phrases et leur ordre de présentation ont été pris en compte afin de ne pas induire d'effet d'ordre sur la reconstitution. Une phrase donnée pouvait donc être présentée en premier deuxième ou troisième rang et sa présentation pouvait être soit écrite soit traduite soit contextualisée.

Lorsque la phrase à reconstituer est présentée à l'écrit, on peut considérer être dans une procédure interne à un seul

langage : de l'écrit compris par la lecture à de l'écrit reconstitué à partir de ses éléments pour soulager et induire le travail d'écriture.

Lorsque la phrase est transmise en Langue des Signes, c'est-à-dire pensée par l'émetteur et reçue par le destinataire en LSF, le sujet comprend un message dans une langue (et donc une syntaxe propre) et doit le produire dans une autre langue et une autre syntaxe. On peut s'attendre, dans ce cas, à ce qu'il s'engage dans une sorte de traduction et qu'il cherche, avec les mots disponibles de l'écrit, à reconstituer peu ou prou, la syntaxe du message en LSF.

Lorsque la phrase est seulement contextualisée dans ses intentions et décrite sommairement dans sa structure, on peut faire l'hypothèse qu'il n'y aura pas (ou moins) de traduction (au mot à mot ?) d'une langue dans une autre mais directement production d'un texte à partir de la rencontre d'une intention et du matériel écrit proposé, une sorte d'incitation à penser directement avec de l'écrit sans l'écran d'un autre système linguistique.

Nous avons émis l'hypothèse que les meilleurs résultats seront obtenus par la présentation écrite (de l'écrit lu à l'écrit produit), puis par la présentation contextualisée (du sens à l'écrit) puis lors de la traduction en LSF puisque, dans ce dernier cas, l'hétérogénéité des syntaxes viendrait s'interposer.

● Nous avons dans un premier temps, testé **9 adultes** sourds. Les phrases qui leur ont été présentées étaient complexes :

1) Si tu étais libre samedi prochain, nous pourrions aller ensemble à la piscine qui a ouvert récemment dans le quartier où habite mon petit frère depuis qu'il a déménagé. (30 mots 5 propositions)

2) Quand tu auras fini de lire le livre qu'elle t'avait prêté, il faudrait que tu l'apportes à ma cousine puisqu'elle aime les romans policiers. (28 mots 5 propositions)

3) Bien qu'elle ait terminé ses valises et qu'ils soient enfin prêts à partir, il angoissait à l'idée qu'ils risquaient de rater le train qui ne les attendrait pas. (32 mots 5 propositions)

● Nous avons ensuite testé **16 élèves** de collège de niveau 6^{ème} et 5^{ème} de deux établissements différents (Noisiel et Nantes). L'exercice était le même que celui proposé aux adultes, cependant, la difficulté des phrases a été adaptée : comparativement aux adultes, les élèves avaient une connaissance et une pratique de l'écrit moins importante.

1) Si tu étais libre samedi prochain, nous pourrions aller ensemble à la piscine qui a ouvert récemment dans le quartier. (20 mots 3 propositions)

2) Quand tu auras fini de lire le livre qu'elle t'avait prêté, tu le rapporteras à ma cousine. (19 mots 3 propositions)

3) Bien que les valises soient terminées, Suzanne et son frère angoissaient à l'idée qu'ils risquaient de rater leur train. (21 mots 3 propositions)

Dépouillement des résultats

Dans tous les cas, les reconstitutions de phrases à l'aide des étiquettes ont été filmées. Les éléments vidéo capturés constituent 7 heures d'enregistrement. Les séquences ont été visionnées une première fois afin de prélever le temps de passation et l'état de la phrase finale, puis une deuxième fois suite aux analyses statistiques affinées qui avaient permis de dégager des axes d'observation particuliers : identifier des stratégies et des invariants dans la structuration des phrases, et tenter de comprendre ce qui était à l'œuvre dans la reconstitution des phrases.

Le temps de réalisation a été relevé et les phrases finales obtenues ont été notées selon 5 scores :

Score A : nombre de mots à leur place à l'intérieur de chaque proposition
Score B : nombre de propositions bien placées par rapport à celle qui la précède
Score C : nombre de mots qui se suivent correctement au-delà de 2
Score D : séquences de conjugaisons
Score E : séquences sémantiques

(Pour l'analyse on se reportera à la partie analyse statistique détaillée).

Remarques effectuées lors de l'observation des vidéos

L'observation des vidéos a été orientée par le traitement statistique détaillé des données. Il a été possible de relever différentes stratégies en fonction du mode de présentation de la phrase, et ce, aussi bien chez les adultes que chez les élèves :

● **L'utilisation de la mémoire photographique** lorsque l'individu est confronté à une présentation écrite : la phrase reconstituée respecte parfois exactement la mise en page de la feuille de présentation avec un retour à la ligne identique. Cette stratégie est un peu dangereuse dans la mesure où il est possible de s'appuyer complètement sur cette mémoire photographique sans chercher à vérifier si la phrase et son sens sont corrects. Dans le cas où cette mémoire n'est pas aussi fine, la reconstitution sera bancal et non travaillée par rapport au sens ou à la syntaxe.

Sur la photo ci-dessous, la mémoire photographique est performante, mais n'indique en rien que l'élève a compris le sens de la phrase.



● Face à la situation de présentation en traduction, **la reconstitution suit parfois la chronologie utilisée en LSF**, par exemple, pour la phrase finale de la phrase 2 on pourra obtenir : « *Quand ma cousine. prêté, le livre elle avait tu auras de lire fini qu'tu t'le rapporteras* ».

● En situation de présentation contextualisée, les mots inconnus et les mots outils posent plus de difficultés et semblent perturber davantage les élèves. Là encore, il ne sera pas rare qu'on les retrouve en fin de phrase.

● **Le phénomène de tâtonnements successifs :** construction par essais/erreurs. La phrase se construit d'abord avec des associations de petits groupes de mots, puis d'association entre groupes et enfin l'insertion des mots outils ou des mots inconnus en dernier.



● **Le tri par catégorie de mots observé uniquement chez les adultes.**



Difficultés spécifiques observées au niveau des phrases

L'observation des difficultés au niveau des phrases a été orientée par l'analyse statistique détaillée.

Alors que les phrases avaient été créées dans un souci d'équilibre et d'homogénéité (même nombre de subordinées, approximativement même nombre de mots...), force

est de constater que les deux premières phrases semblent être plus facilement reconstituées que la troisième. Ceci pourrait s'expliquer par le fait que cette dernière exprime la concession à l'aide de la conjonction « *Bien que* » et que les élèves ne connaissent pas cette locution. Ils connaissent le mot « *bien* » dans son sens de critère d'appréciation.

Ainsi, on retiendra qu'il paraît nécessaire de travailler les expressions conjonctives

De plus, l'expression « *il angoissait à l'idée que* » est une expression typiquement française qui s'exprimera en LSF par le fait d'angoisser : on ne trouve pas l'équivalent de « à l'idée que ». Du coup, tout paraît comme s'ils utilisaient les mots qui constituent ces expressions de manière séparée.

Il est donc à noter qu'il serait intéressant d'intégrer au logiciel un répertoire d'expressions en français écrit, qu'on ne retrouve pas en LSF (cela peut aussi s'envisager dans le sens inverse, d'expressions LSF qui n'ont pas leur équivalent à l'écrit).

L'adverbe « *récemment* » a aussi posé problème car son sens n'est pas toujours connu.

Il est important de se rappeler qu'un travail sur les mots et leurs racines (par exemple la constitution des adverbes, la construction de mots utilisant des suffixes et des préfixes) doit être intégré. De même en LSF, il existe des invariants dans la réalisation de certains signes.

À ceci s'ajoute pour la troisième phrase, la difficulté de la majuscule au début de la phrase : le prénom Suzanne commence par une majuscule et de ce fait, les élèves commençaient par ce mot plutôt que par « *Bien* ».

Des erreurs communes à une majorité d'élèves ont été relevées : « *Si tu étais libre le samedi prochain...* », « *...rater à train...* », « *...tu auras lire fini...* », « *...aller dans piscine...* ».

Des éléments d'explication et d'exercices sur la structure des phrases et des expressions doivent être intégrés. De ces erreurs il est possible, comme nous venons de le voir, d'envisager la mise en place d'aides et d'exercices qui permettront d'explicitier ou d'exercer ces constructions.

TRAVAIL DE COMPRÉHENSION

Déroulement

● Deux groupes d'**adultes** sourds ont travaillé sur un texte de Brigitte Garcia¹ qui est enseignante-chercheur au sein de l'UMR 7023 *Structures formelles du langage* (Université Paris 8).

Le support d'étude était extrait d'une conférence donnée à Reims le 10 mai 2005.

Le titre est : « *Les sourds et l'écrit : des difficultés sans lien avec le fait d'être sourd* ».

1. Voir *Sourds Lecteurs, mêmes démarches*, page 69.

● Les 16 **élèves** ont, quant à eux, été soumis à un exercice de compréhension portant sur un texte extrait de la « *Batterie d'Épreuves pour mesurer la lecture et l'orthographe* » ORLEC. Le test de lecture a été développé en 1967 par M. Lobrot, professeur agrégé au CNPS de Beaumont sur Oise. Il mesure l'efficacité de la lecture : l'information intégrée dans un laps de temps donné ou encore, la capacité à comprendre des mots ou des phrases écrites dans un minimum de temps.

Le texte comporte des difficultés dosées graduellement. Les quatre premières phrases sont formées de mots extraits des cent mots les plus courants de la langue française, qui ne présentent jamais plus de 2 syllabes ni de difficultés importantes. Les cinq phrases suivantes sont constituées de mots plus difficiles. Tout d'abord il n'y a pas de mots de plus de deux syllabes, il y a des mots-pièges c'est-à-dire des mots très semblables à d'autres mots du langage courant, et le mot à déchiffrer est le mot rare et non celui du langage courant. Les cinq dernières phrases sont encore plus compliquées, car il y a des mots de plus de deux syllabes qui comportent des difficultés de déchiffrement.

C'est le printemps. Les arbres sont en fleur. Depuis le matin, les oiseaux vont et viennent dans le soleil. Il est huit heures et je m'en vais de la maison. Je dis au revoir à mon père et à ma mère.

Je marche sur les dalles du jardin. Les pots sont casés le long de l'allée. Les pelouses ressemblent à de grosses tranches vertes. Le chat hante la margelle du puits. Des chemises sèchent sur le fil de fer.

Tout à coup, j'entends le vrombissement d'un avion. Je cours jusqu'à la vigne et je lui fais signe. Ahuri, je le vois atterrir sur le bord du chemin. Je grimpe et nous démarrons. En route pour l'aventure !

Les élèves doivent lire deux fois le texte, les deux temps successifs sont relevés, puis on pose six questions prédéfinies de compréhension. On note les réponses des élèves, attribuée 1 point si la réponse est correcte ou 0 si elle est incomplète ou fautive.

Dans l'épreuve ORLEC, les élèves devaient effectuer une lecture à « *haute voix* » ; dans notre cas, les élèves ont effectué une lecture silencieuse et on leur a posé en LSF les questions concernant le texte.

Dépouillement des données

Le travail des élèves a donné lieu à des documents écrits, le dépouillement a été effectué pour être intégré dans un tableau avec les éléments des exercices de remise en ordre de phrases.

Analyse des résultats

- Concernant les adultes (*voir article page 69*),
- Concernant les élèves.

Il a été possible de constater que le vocabulaire utilisé posait des difficultés : peu d'élèves connaissaient les mots « *aéroplane* » et « *vrombissement* ». Toutefois, certains, par des effets de contexte ont pu en inférer le sens.

Pour l'analyse on se rapportera à la partie analyse statistique détaillée.

EXERCICE DE BARRAGE DE LETTRES

Déroulement

Seuls les élèves ont passé ce 3^{ème} exercice qui consistait à barrer des lettres dans un texte d'une longueur d'une page.

Dans la littérature scientifique, il semble établi que, lors d'une épreuve de barrage de lettres, on en oublie toujours, mais inégalement et davantage dans les mots qui établissent la syntaxe de la phrase que dans les mots qui véhiculent un contenu sémantique. Et le rapport entre le taux d'oublis dans les mots syntaxiques et le taux d'oublis dans les mots sémantiques est fortement et positivement corrélé avec la qualité du lecteur (sa performance de lecture). On explique en général ce phénomène par l'anticipation de l'organisation syntaxique grâce à la vision parafovéale, anticipation qui

rend secondaire le traitement des mots syntaxiques en vision fovéale et permet alors de porter toute l'attention sur les mots « sémantiques ». D'où le lien avec la qualité de lecture.

Il a été demandé aux élèves de barrer la lettre U dans un texte. La consigne était d'abord présentée en LSF puis à l'écrit. La consigne écrite servait de support d'entraînement et les élèves devaient barrer la lettre U dans cette partie. Il leur était demandé d'effectuer une lecture normale du texte sans effectuer de retour en arrière, et de s'attacher tout de même à sa compréhension. Une fois la consigne comprise, le texte leur était présenté.

Les données recueillies sont : le temps de lecture du texte ainsi que la répartition entre les lettres barrées au sein des mots sémantiques et des mots syntaxiques et le rapport du nombre de lettres oubliées selon la nature des mots qui les contiennent.

Aujourd'hui c'était la journée parfaite. Il faisait beau et chaud et de plus, j'étais de bonne humeur. Le printemps avait déjà fait son apparition et l'été arrivait à grands pas. Je n'allais pas passer outre à mes obligations. La chose me faisait tout de même plaisir, j'en étais tout excité. J'étais décidé, peut-être même déterminé. C'était aujourd'hui que j'allais me trouver un appartement ! Je me levai donc et passai à la salle de bain. L'eau coula sur mon corps pendant que je rêvais aux pièces d'un nouvel appartement merveilleux. Je passai ensuite de la salle de bain à la salle à manger. La journée s'annonçait bien. Des croissants m'attendaient sur le comptoir. Je préparai rapidement, pour les accompagner, un grand café au lait. Il fumait maintenant devant moi avec une odeur exquise.

Quand on veut être bien disposé, on ne peut passer outre la nécessité d'un bon déjeuner ! De plus, j'avais déjà lu que c'était le repas le plus important de la journée. Je renchéris alors mon menu de noix, de fruits et de céréales. Aussi, je m'étais muni du journal « Voir ». J'allai donc à la section des petites annonces. Logements, logements, logements, ah voilà ! La rubrique n'était pas grosse mais tout de même.

J'allai tout de suite au quartier où j'espérais faire une trouvaille, Saint Jean-Baptiste. Les loyers n'y sont pas élevés, mais outre ces considérations financières, j'aime ce quartier que je trouve particulièrement sympathique. J'avais toujours pensé que certains quartiers avaient davantage d'âme. Saint Jean-Baptiste était l'un de ceux-là. Je scrutai donc chacune des annonces. Il ne fallut que quelques instants pour que la chance se pointât. Je me demandai un moment si j'avais bien lu. Je relus et fus convaincu. Mais qui pourrait renoncer à une telle aubaine ! Il s'agissait d'un 4-5 pièces, quatre cent euros tout compris. C'était rue Saint Olivier. On indiquait qu'il était disponible dès maintenant pour visiter.

Je sautai sur le téléphone afin de prendre un rendez-vous. J'étais un peu nerveux, mais sauf les palpitations de mon cœur, tout allait bien. En

moins de deux, une entente fut convenue. J'avais rendez-vous dans une heure avec le propriétaire de cet appartement qui promettait. Je terminai de manger et finis de m'habiller. J'allai ensuite chez ma copine pour lui apprendre la bonne nouvelle. Elle était enchantée. Pendant qu'on marchait, elle me raconta son rêve où elle se voyait prendre un verre sur le balcon d'un appartement de ce quartier. C'était notre première visite de logement. (Texte extrait du matériel de tests de la thèse de D. Foucambert)

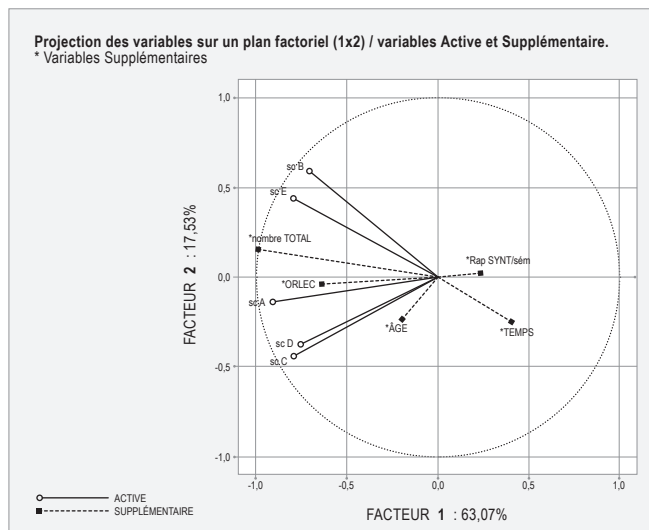
Dépouillement des données

Une grille d'analyse du texte a été mise au point afin de pouvoir comptabiliser les oublis de U au sein des mots syntaxiques et sémantiques.

ANALYSE STATISTIQUE DÉTAILLÉE DES RÉSULTATS

Une analyse statistique détaillée concernant les résultats des élèves et mettant ainsi en relation les différentes variables prélevées dans le cadre de l'exercice de la reconstitution des phrases ainsi que des autres exercices de compréhension et de barrage de lettres présentés, a été menée par Jean Foucambert, afin de déterminer quelle était la stratégie utilisée dans la reconstitution et de tester l'hypothèse. La méthode utilisée était une analyse en composantes principales (ACP), suivie d'un modèle de régression multiple.

Lors de l'ACP on obtient la figure suivante :



On voit que 80% des 5 scores sont bien rendus par les 2 premières composantes.

- Le facteur 1 exprime bien la cohérence entre les 5 scores, c'est le niveau général de la performance. Le score total (var. suppl.) en rend aussi très bien compte. La performance de lecture (ORLEC) va dans le même sens (est bien corrélée) avec le facteur 1 et avec le total. En bref, plus on est bon lecteur, plus on remet bien les phrases en ordre. Ce qui est surprenant, c'est le positionnement inverse du rapport d'oublis syntaxiques/sémantiques (var. sup.). Normalement, il est lié à une bonne lecture donc à la réussite dans la reconstitution.

- Le facteur 2 oppose 2 paires de scores : [C et D] contre [B et E], de part et d'autre de A.

Le score C mesure le nombre de mots qui se suivent correctement (au-delà de 2) et rend compte sans doute de la familiarité avec l'enchaînement des mots à l'écrit.

Le score D mesure les associations correctes dans les séquences sujets-verbès.

Ces deux scores C et D semblent rendre assez bien compte de la qualité des micro-liaisons de proximité entre les mots à l'intérieur de la phrase.

Le score B mesure le nombre de propositions bien enchaînées c'est-à-dire la relation qui existe entre elles au niveau du sens de la phrase.

Le score E mesure la présence de mots dont la proximité est nécessaire au plan sémantique, en quelque sorte des unités de sens incontournables pour la compréhension globale de la phrase.

Ces deux scores B et E semblent rendre assez bien compte de la composante compréhension, compréhension du fonctionnement de la phrase qui va de pair avec le repérage des associations sémantiques autour desquelles les phrases s'organisent.

Ce deuxième facteur introduit donc une différenciation entre les capacités qui participent à la constitution de la performance globale. On va trouver au sud de l'axe (sa partie négative) les sujets qui s'y prennent plutôt bien dans les micro-ajustements de mots et au nord de l'axe (côté positif) ceux qui ont une approche plus liée à la grammaire de la phrase et à la signification globale.

Dans une ACP, par construction les facteurs sont indépendants. On peut suggérer ici que le premier facteur, quelle que soit la manière dont ils s'y prennent, répartit les individus depuis les plus performants (à gauche) jusqu'aux plus faibles (à droite) pour la remise en ordre d'une phrase. Le second facteur ventile les individus, cette fois, quel que soit leur niveau de réussite, selon la nature de leurs stratégies (points d'appui), au sud ceux qui s'appuient davantage sur les micro-ajustements ; au nord ceux qui privilégient le sens de la phrase.

Finalement l'analyse tentera d'explorer :

- Le niveau de la performance, soit le 1^{er} facteur de l'ACP ;
- La nature de la stratégie, soit le second facteur de l'ACP. Cette analyse pourra être complétée en prenant séparément la performance dans les 2 paires [C+D] et [B+E].

Le niveau de la performance ²

Voici le tableau récapitulatif donnant la contribution des variables indépendantes dans un modèle de régression multiple qui tente d'expliquer le niveau de la performance.

Résultats univariés de chaque variable. Paramétrisation SIGMA restreint. Décomposition efficace de l'hypothèse.					
	Degrés de liberté	Niveau Somme des carrés	Niveau Moyenne des carrés	Niveau F	Niveau p
Intercept.	1	689,26	689,258	9,98249	0,019577
phrase	2	1110,36	555,180	8,04064	0,020062
passation	2	1649,91	824,956	11,94778	0,008084
rang	2	1285,57	642,786	9,30943	0,014476
SEXE	1	62,85	62,846	0,91019	0,376902
temps	1	561,98	561,976	8,13907	0,029071
ÂGE	1	447,47	447,473	6,48073	0,043742
Rap SYNT/sém	1	724,74	724,735	10,49630	0,017691
ORLEC	1	3912,25	3912,250	56,66089	0,000285
Erreur	6	414,28	69,047		
Total	17	11492,00			

Avec un R² ajusté de 0.90, il explique plus de 96% de la variance totale des résultats.

2. Pour une meilleure visualisation, l'axe 1 a été inversé afin d'associer les scores élevés aux bons résultats

Parmi les 8 variables explicatives proposées, 5 décrivent l'individu qui a fait l'exercice :

- Le sexe : il n'intervient pas ici de manière significative.
- L'âge intervient mais presque au seuil de significativité fixé à $p < 0.05$
- Le comportement de lecteur est cerné par le temps mis pour exécuter la remise en ordre (p significatif à .029) ; par une épreuve de barrage de lettres où est calculé le rapport entre les lettres oubliées dans les mots syntaxiques et dans les mots sémantiques, également significatif à .017 ; et par une épreuve standardisée (ORLEC) de compréhension d'un texte, contribution très significative (seulement 3 probabilités d'erreur pour 10 000 !)

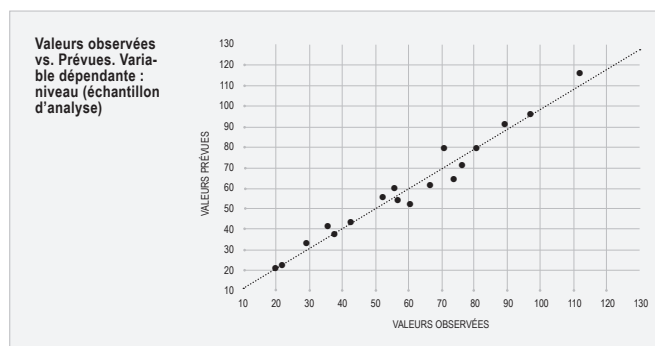
Une variable identifie la phrase à reconstituer : la variable « phrase » est significative à $p.02$

Une variable identifie le rang de la phrase selon que le sujet l'a traitée en premier, deuxième ou troisième position : effet également significatif à $p.014$.

Une variable enfin identifie le mode de présentation (passation) de la phrase à reconstituer selon qu'elle a été lue par l'enfant lui-même, qu'elle lui a été interprétée en LSF ou qu'elle a été seulement contextualisée en LSF par son intention et décrite sur le plan de sa composition : cette variable joue un rôle très significatif à $p.008$

Donc 7 variables sur 8 expliquent la performance et dans l'ordre décroissant d'importance : le niveau de compréhension en lecture, le type de présentation, le rang dans l'ordre de passation, etc.

Le pouvoir prédictif du modèle est puissant si on en juge à la comparaison entre les valeurs prévues et les valeurs réelles :



La régression multiple permet de décrire l'effet de chaque variable, toutes choses égales par ailleurs. Par exemple, lorsqu'on étudie l'effet « phrase », toutes les autres variables sont contrôlées : on a donc l'effet propre de chacune des 3 phrases et ceci quel que soit leur mode de présentation ou le niveau de lecture du sujet ou son âge, etc. On a donc accès au pouvoir explicatif de chacune des variables à l'intérieur du modèle. Nous allons donc passer en revue les 7 variables (sur 8) dont l'effet est significatif.

a) Variables qualitatives

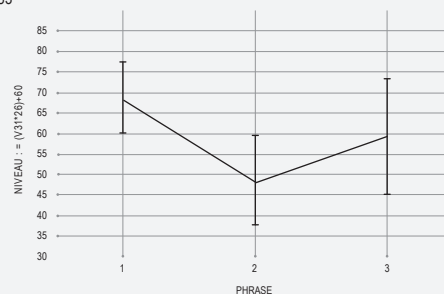
● **L'effet phrase.** Les trois phrases ne provoquent pas le même niveau de réussite. On voit à l'étendue des limites de confiance de la moyenne que la phrase 1 semble plus facile que les deux autres, lesquelles ne diffèrent pas entre elles de façon significative. On peut donc formuler deux remarques :

- *Même si nous avons essayé de proposer des phrases de « difficulté » et de « consistance » égales et malgré le soin apporté lors des pré-tests, on constate que, pour les sujets concernés, les phrases ne sont pas équivalentes. On observe aussi bien chez les élèves que chez les adultes, un effet de phrase. La phrase 1 semble être mieux réussie que les deux autres. Il est important de comprendre pourquoi elle offre moins de résistance à la reconstitution que la phrase 2.*

- *Grâce à la régression multiple, cette variance est intégrée dans le modèle et ne se répercute pas sur le rôle des autres variables (toutes choses égales par ailleurs)*

PHRASE : Moyenne pondérée / effet courant : $F(2, 6) = 8,0406, p = .02006$
(Décomposition efficace de l'hypothèse) / Barres verticales représentent les intervalles de confiance à 0,95.

Moyenne des covariables
temps : 3,162778
âge : 13,41667
Rap SYNT/sém : 2,923333
ORLEC : 3,166667

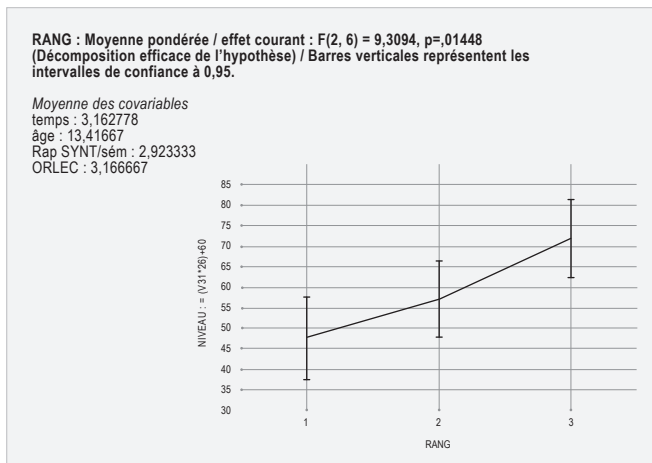


● **Le rang de passage.** Même si nous avons fait varier l'ordre de présentation et le mode de passation entre les sujets (principe du carré latin), l'analyse statistique des résultats des élèves a permis de montrer que, la dernière phrase quelle que soit sa nature et son mode de présentation, offrait une meilleure réussite.

On voit clairement qu'il y a un effet d'apprentissage, quelles que soient les autres variables. On constate un effet progressif qui permet d'affirmer que la 3^{ème} passation est significativement supérieure à la première.

Là encore, l'observation attentive des processus mis en œuvre par les sujets doit permettre de comprendre ce qui « s'apprend » en cours de route et qui pourrait peut-être alimenter quelques investissements pédagogiques.

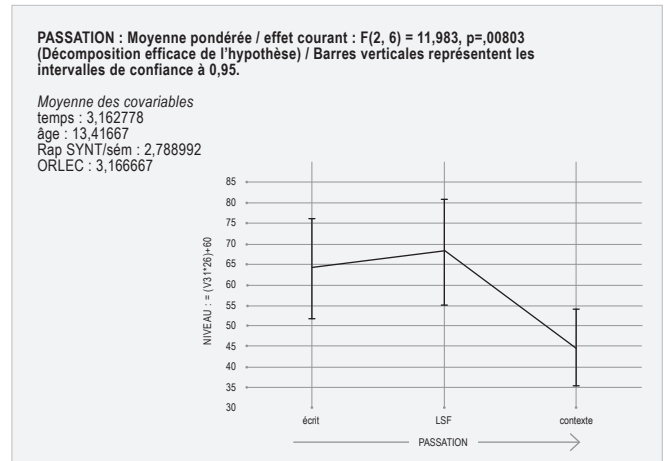
Pour autant, toujours grâce à la régression multiple, cette variation est contrôlée et n'oblitére en rien l'effet spécifique des autres variables.



● **L'effet du mode de présentation de la phrase.** Si les 2 variables précédentes étaient dans le modèle essentiellement pour en contrôler les éventuels effets, la variable mode de présentation est une de celles dont nous souhaitons pouvoir tester le rôle.

L'hypothèse était que de l'écrit vers l'écrit devrait donner de meilleurs résultats, suivi de contexte vers écrit, puis de traduction LSF vers écrit (puisqu'il vient s'interposer

l'hétérogénéité des syntaxes). Finalement, force est de constater que c'est faux :



Il est alors possible de faire une hypothèse sur l'échec de l'hypothèse : les élèves sont ici des sourds LSF qui en sont où ils peuvent dans leur niveau de maîtrise de l'écrit. Incontestablement, en production d'un message écrit seulement à partir d'un matériel mots et avec pour seule boussole communicative une information sur le *contexte* de production de la phrase cible, la tâche est difficile. Ce qui distingue les 2 premières modalités de cette troisième, c'est que la phrase est comprise préalablement, soit à travers une lecture directe de l'*écrit*, soit à travers son interprétation en *LSF*. Le fait que ces 2 modalités donnent à peu près le même résultat (avec un petit avantage à la source LSF) suggère que la lecture, non seulement n'a pas permis une meilleure compréhension que l'interprétation LSF mais n'a pas provoqué de familiarité exploitable avec la « grammaire » des mots et de la phrase écrite. Si l'hypothèse initiale n'est pas vérifiée, sans doute faut-il chercher des pistes d'explication du côté de la performance de lecture elle-même. D'où la nécessité de regarder (en revenant à l'enregistrement des passations) si chaque sujet met en œuvre des stratégies de reconstitution différentes selon le mode de présentation de la phrase.

b) Variables quantitatives

La distribution de ces quatre variables est donnée dans le tableau suivant :

Statistiques descriptives					
	Valid N	Mean	Minimum	Maximum	Std.Dev.
temps	18	3,16278	0,85000	5,50000	1,341526
ÂGE	18	13,41667	12,50000	15,00000	0,861770
Rap SYNT/sém	18	2,92333	1,29	6,02	1,573602
ORLEC	18	3,16667	2,00000	5,00000	1,098127

Le score niveau qu'on cherche à expliquer a une moyenne de 60 et un écart-type de 26 avec un minimum de 20 et un maximum de 112 points.

Voici les corrélations simples entre ces 4 variables explicatives et le score niveau :

	temps	ÂGE	Rap SYNT/sém	ORLEC
temps	1,00	- 0,16	- 0,14	0,03
ÂGE	- 0,16	1,00	- 0,67	0,20
Rap SYNT/sém	- 0,14	- 0,67	1,00	- 0,23
ORLEC	0,03	0,20	- 0,23	1,00
Niveau	- 0,41	0,19	- 0,24	0,63

On voit dans le tableau suivant que toutes les variables quantitatives contribuent de manière significative, dans ce modèle de régression multiple, à l'établissement du score niveau :

	Niveau paramètre	Niveau Std.Err	Niveau t	Niveau p	Niveau Beta
temps	- 11,0955	3,88921	- 2,85291	0,029071	- 0,572498
ÂGE	- 10,8900	4,27775	- 2,54573	0,043742	- 0,360948
Rap SYNT/sém	- 7,3720	2,27544	- 3,23980	0,017691	- 0,446175
ORLEC	14,4346	1,91762	7,52734	0,000285	0,609653

La colonne paramètre indique quelle contribution apporte à ce score l'augmentation d'une unité de la variable correspondante. Ainsi, l'âge étant donné en année, chaque année supplémentaire fait baisser le score niveau de 10.89 points. Et chaque point en plus en lecture (ORLEC) élève le résultat niveau de 14.43.

● **L'âge.** On voit dans le tableau des corrélations que l'âge est corrélé à .19 avec le niveau, donc positivement mais pas significativement. En revanche, dans la régression multiple, le lien est significatif (de justesse, le seuil étant p .05) mais négatif. Le fait pour le sujet qui reconstitue la phrase d'être plus vieux d'une année fait baisser le score de 10.89 points ! Ceci éclaire le rôle de la régression multiple et de l'analyse « *toutes choses égales par ailleurs* ». La régression simple (corrélation) montre faiblement que plus un élève est âgé et mieux il lit. Mais le fait d'être plus âgé, sans pour autant lire mieux (donc à niveau de lecture constant) témoigne de difficultés, ce qui apparaît dans le niveau de réussite.

● **Le temps.** Il en va de même du temps mis pour opérer cette reconstitution. À niveau de lecture équivalent, à âge équivalent, etc. plus on met de temps pour accomplir une tâche, plus on a de difficulté. Une minute de plus (sur 3.16 en moyenne) rend prévisible une baisse du niveau de 11.09 points. La rapidité est signe de maîtrise.

● **Compréhension de lecture (ORLEC).** Cette variable, comme toutes les autres dans une régression multiple, joue un double rôle. En premier lieu, elle « contrôle » le niveau de compréhension en lecture et des autres variables du modèle sauf celle dont on veut observer l'effet « *toutes choses égales par ailleurs* » donc. Et elle permet d'évaluer son effet spécifique lorsque toutes les autres sont à leur tour contrôlées. Sous ce dernier aspect, on n'apprend rien ici de surprenant : la compréhension de l'écrit contribue très fortement à la performance de remise en ordre des mots d'une phrase et ceci, quels que soient la phrase, le rang, le mode de présentation, l'âge, le temps mis, etc.

Pour autant, cette contribution est la plus forte de toutes les variables prises dans le modèle : chaque point gagné en lecture (entre 2 et 5 avec une moyenne de 3.16) élève le niveau de réussite de près de 14 point (sur un score de moyenne 60).

● **L'oubli de lettres** (rapport oubli syntaxique/oubli sémantique).

Ici, avec des sujets sourds pratiquant de manière prédominante la LSF, plus ce rapport est faible, c'est-à-dire plus les oublis de lettres se répartissent également entre mots syntaxiques et sémantiques et plus leur niveau d'efficacité dans la reconstitution de phrase est élevé.

S'il n'y a, dans le recueil et le traitement des données ni erreurs ni artefact, force est de s'interroger. Certes, les sourds, comme les entendants, oublient davantage de lettres dans les mots syntaxiques que dans les mots lexicaux. En moyenne, 2.9 fois plus. Lorsqu'on tente d'expliquer leur niveau de lecture (ORLEC) dans une régression multiple incluant l'âge, le sexe et ce rapport syntaxique/sémantique, rien n'est significatif même si la contribution de ce rapport au niveau de lecture est plutôt négative. En revanche, lorsqu'on introduit dans le modèle, le fait pour un sujet d'être appareillé (1 sur 2 le sont), l'effet de ce rapport syntaxique/sémantique devient significatif ($p=.0005$!) et positif, ce qui suggère que, une fois contrôlé l'effet « appareillage », les choses se passent de la même manière pour les sourds et les entendants. À défaut d'avoir l'expérience de la syntaxe de l'écrit à travers un enseignement systématique, tous les individus transposent tant bien que mal sur l'écrit leur expérience de la syntaxe de l'oral. Les sourds LSF non appareillés n'ont rien à transférer : l'écrit reste une succession de mots ayant chacun du sens (ou pas, par exemple pour eux la majorité des mots syntaxiques) mais leur organisation par une syntaxe reste un mystère et ce, d'autant plus, que la syntaxe de la LSF ne saurait les mettre sur une bonne piste. Les sourds appareillés, comme les entendants, n'ont pas ce désavantage et parviennent à se construire seuls quelques structures à partir de l'oral, même si l'écrit et l'oral n'ont pas la « même » grammaire. Toujours est-il que des sourds et des sourds appareillés peuvent bien avoir le même rapport d'oublis de lettres entre les mots syntaxiques et les mots lexicaux, la différence entre ces taux respectifs n'a pas la même origine. Pour les uns (qui peuvent transférer de l'oral une expérience « fonctionnelle » de sa syntaxe), l'oubli plus important de lettres dans les mots syntaxiques s'explique par leur fonction d'organiseurs de la structure de la phrase perçue alors en vision parafovéale, ce qui est une condition d'une lecture efficace ; pour les autres (en panne de syntaxe) le

même taux d'oublis s'expliquent par la fréquence beaucoup plus grande des « mots-outils », donc par leur familiarité mais cette identification sommative sans insertion dans une structure est un cul de sac pour la lecture.

Cette double étiologie de l'oubli différentiel de lettres pourrait expliquer pourquoi, lorsqu'on ne contrôle pas dans le modèle de régression multiple l'effet « appareillage » (en réalité l'effet « dispose ou non d'une certaine syntaxe de l'écrit à partir de l'oral ») mais qu'on contrôle la performance de lecture (laquelle inclut nécessairement un état donné de maîtrise de la syntaxe), la variable « rapport syntaxique/sémantique », lorsqu'elle augmente, témoigne simplement de plus de « fréquence » (stratégie nuisible) et non de plus de « syntaxe » (stratégie utile pour réussir la tâche) car la syntaxe disponible est déjà exprimée par le niveau de lecture. D'où la contribution négative et significative de ce rapport syntaxique/sémantique à la réussite de la tâche. Lorsqu'on inclut l'effet appareillage, c'est lui qui absorbe l'effet « dispose de la syntaxe ». Tout cela n'est pas forcément clair...

Ce qui est clair en revanche, c'est qu'avec des sourds non appareillés, il faut, encore plus qu'avec des entendants, faire un travail sur l'organisation de la phrase écrite et particulièrement sur tous ces mots qui ne se « traduisent » pas en LSF.

Il sera sans doute utile de regarder dans les enregistrements vidéo comment les sourds privilégient ou négligent les mots syntaxiques lors de leur reconstitution. À première vue, il semble en effet que la stratégie privilégiée porte sur les mots sémantiques importants autour desquels progressivement s'agglutinent les « petits » mots...

On en vient ainsi au second facteur de l'ACP initiale.

La nature de la stratégie ³

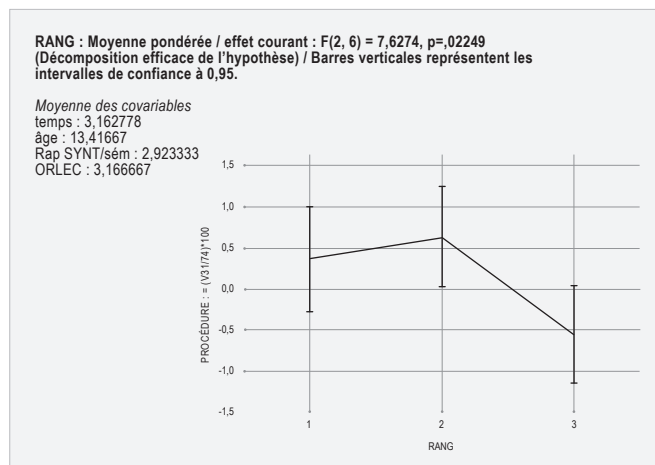
Ce deuxième facteur introduit une différenciation entre 4 scores partiels qui objectivent certaines capacités qui participent à la constitution de la performance globale. On va trouver au sud de l'axe (sa partie négative) les sujets qui s'y prennent plutôt bien dans les micro-ajustements de mots et au nord de l'axe (côté positif) ceux qui ont une approche plus liée à la grammaire de la phrase et à la signification globale.

3. L'axe 2 n'a pas été inversé car il n'a pas un pôle « meilleur » que l'autre

On utilisera le même modèle que précédemment. Il se révèle là encore assez puissant avec un R^2 ajusté de .706

Une seule variable est significative au seuil p .022

Il s'agit du rang de passage...



Le 3^{ème} passage (quelle que soit la phrase ou sa présentation) se différencie très significativement des 2 premiers : la performance lors du 3^{ème} passage recule sur l'axe des procédures de plus de 1.05 point entre la moyenne des 2 premiers passages (+0.496) et la moyenne du 3^{ème} (- 0.570).

Rappelons que ce facteur n'exprime pas un continuum (du faible au fort ou inversement) dans une compétence mais rejoint 2 pôles qui semblent chacun rendre compte de stratégies différentes dans le processus de reconstitution de phrases : pour simplifier, du micro au pôle négatif au macro au pôle positif. Or ici, au 3^{ème} passage, on observe un déplacement important du côté des micro-ajustements.

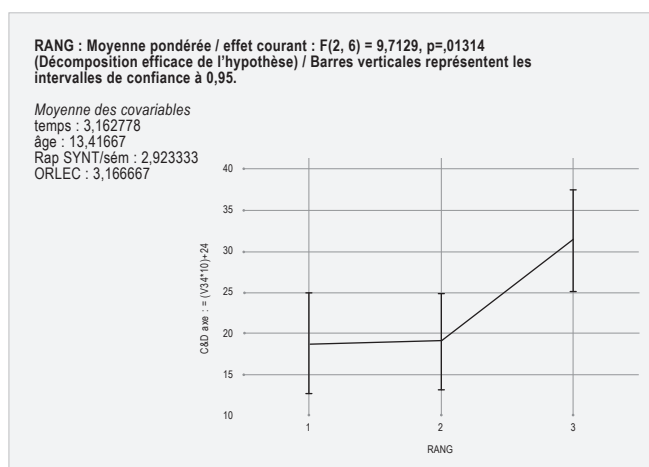
Si on se souvient que pour l'axe précédent, la 3^{ème} passation marquait un progrès dans le niveau de réussite qu'on pouvait imputer à un apprentissage ; et si la même 3^{ème} passation marque ici un déplacement des stratégies vers de meilleurs micro-ajustements, on peut être tenté d'en déduire, par transitivité, que l'amélioration de la performance résulte plutôt d'une gestion efficace de la proximité plutôt que d'un progrès dans l'économie générale de la phrase et du sens.

On doit pouvoir vérifier cette hypothèse en observant en ce sens les enregistrements. Mais également en analysant les scores [C+D] puis [B+E]

Du côté des micro-ajustements

On travaillera ici à partir des scores [C+D] traités préalablement dans une ACP et dont le premier axe est ensuite ramené à une distribution de moyenne 24 et d'écart-type 10. La régression multiple avec le modèle utilisé jusque là est puissant avec un R^2 ajusté de .737. Deux variables explicatives sur les 8 du modèle sont significatives.

● **Le rang de passage** (variable qui intervenait déjà pour l'axe des procédures.)



On vérifie ce que nous avons appelé précédemment l'effet apprentissage de la troisième passation et qui se manifeste par une forte élévation de plus de 12 points du score [C et D].

● **La compréhension de lecture (ORLEC).** Cette variable est significative au seuil p .007. Elle contribue positivement au score [C et D] qui rend précisément compte des procédures de micro-ajustements. Chaque point supplémentaire en compréhension de lecture élève le score des micro-ajustements de 4.7 points (presque 1/2 écart-type).

Grammaire de la phrase et signification globale

Il s'agit cette fois des scores [B et E] traités également dans une ACP et dont le premier axe est ramené à une distribution de moyenne 24 et d'écart-type 10.

Avec un R² ajusté de .870, ce modèle explique 95 % de la variance totale.

Résultats univariés de chaque variable. Paramétrisation SIGMA restreint. Décomposition efficace de l'hypothèse.

	Degrés de liberté	B&E axe Somme des carrés	B&E axe Carré moyen	B&E axe F	B&E axe p
Intercept.	1	151,852	151,8521	11,72500	0,014071
phrase	2	201,777	100,8886	7,78994	0,021493
passation	2	187,146	93,5732	7,22509	0,025256
rang	2	43,792	21,8960	1,69066	0,261615
SEXE	1	16,447	16,4465	1,26989	0,302829
temps	1	83,575	83,5752	6,45311	0,044063
ÂGE	1	103,563	103,5626	7,99641	0,030045
Rap SYNT/sém	1	112,273	112,2734	8,66900	0,025803
ORLEC	1	442,711	442,7110	34,18316	0,001104
Erreur	6	77,707	12,9511		
Total	17	1700,000			

6 variables jouent un rôle significatif.

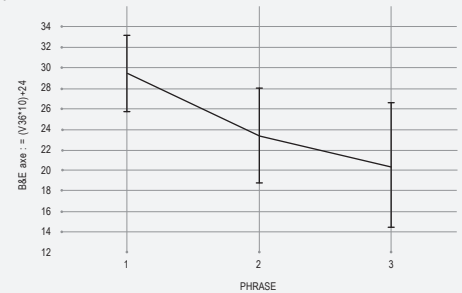
a) Les variables qualitatives

● **L'effet « phrase ».** C'est la phrase 3 qui donne les résultats les plus faibles comparés à ceux des phrases 1 et 2 qui ne se différencient guère (tests post-hoc). Il sera intéressant de comprendre cet effet (critères plus faciles ? phrase plus claire dans son organisation ?) et de voir si les sujets y mettent en œuvre des stratégies moins explicitement orientées vers une prise en compte « globale » de la phrase. (*voir premier schéma, colonne suivante*)

● **L'effet « passation ».** Ici, c'est la présentation par le contexte qui provoque des résultats plus faibles que les 2 autres. On peut supposer que l'absence d'informations explicites relatives au sens de la phrase conduit à des recherches d'ajustements de proximité puis d'ajustements

PHRASE : Moyenne pondérée / effet courant : F(2, 6) = 7,7899, p=.02149 (Décomposition efficace de l'hypothèse) / Barres verticales représentent les intervalles de confiance à 0,95.

Moyenne des covariables
 temps : 3,162778
 âge : 13,41667
 Rap SYNT/sém : 2,788992
 ORLEC : 3,166667

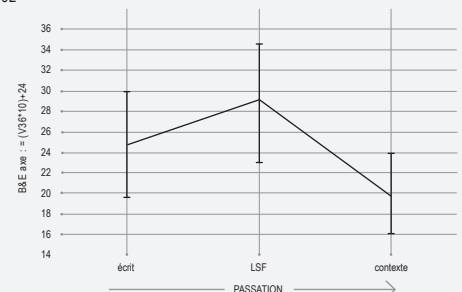


d'ajustements, etc., un peu comme lorsqu'on tente de reconstituer le puzzle d'une image qu'on n'a pas vue.

Toutes choses égales par ailleurs, en croisant les variables phrase et passation, les stratégies les plus intéressantes à observer devraient se rencontrer lorsque la phrase 3 est présentée par son contexte en LSF (c'est là que la variable étudiée a ses plus faibles scores) et lorsque la phrase 1 est présentée dans sa « traduction » en Langue des Signes (scores les plus élevés).

PASSATION : Moyenne pondérée / effet courant : F(2, 6) = 7,2251, p=.02526 (Décomposition efficace de l'hypothèse) / Barres verticales représentent les intervalles de confiance à 0,95.

Moyenne des covariables
 temps : 3,162778
 âge : 13,41667
 Rap SYNT/sém : 2,788992
 ORLEC : 3,166667



b) Les variables quantitatives

	B&E axe paramètre	B&E axe	B&E axe F	B&E axe p	B&E axe Beta
temps	- 4,2789	1,68439	- 2,54030	0,044063	- 0,574020
ÂGE	- 5,2390	1,85267	- 2,82779	0,030045	- 0,451478
Rap SYNT/sém	- 2,9016	0,98548	- 2,94432	0,025803	- 0,456591
ORLEC	4,8557	0,83051	5,84664	0,001104	0,533216

● **Le temps de la reconstitution.** Chaque minute supplémentaire pour faire l'épreuve fait baisser le score [B+E] de 4.28 points. (voir l'effet du temps sur la performance)

● **L'âge.** Chaque année supplémentaire du sujet fait baisser le score de 5.24 points. (voir l'effet de l'âge sur la performance)

● **Le niveau de lecture (ORLEC).** Chaque point supplémentaire du niveau de lecture élève le score de 4.86 points (voir l'effet de cette variable sur la performance)

● **Le rapport syntaxique / sémantique.** Chaque point supplémentaire du rapport syntaxique / sémantique fait baisser le score de 2.9 points. On se retrouve à faire les mêmes commentaires que pour l'effet de ce rapport sur le niveau général : plus le différentiel d'oublis de lettres dans les mots syntaxiques par rapport aux sémantiques est élevé et plus faible est la performance à ce score qui décrit l'efficacité des stratégies portant sur la grammaire de la phrase et les aspects sémantiques. C'est le contraire de ce que suggère l'observation du lecteur entendant.

Si on introduit dans le modèle la variable « appareillage » (en réalité l'effet « dispose ou non d'une certaine syntaxe de l'écrit à partir de l'oral ») - voir l'effet de cette variable sur la performance -, cet effet rapport cesse complètement d'être significatif et l'effet niveau de lecture (ORLEC) s'en trouve renforcé : désormais chaque point supplémentaire en lecture élève la performance « grammaire de la phrase et stratégie » de 6.41 points.

Bilan et apports pour le futur logiciel

D'une part, les 2 types de procédures aux extrémités du second facteur de l'ACP sont reprises dans les couples. L'observation des enregistrements vidéo doit permettre de décrire ces procédures en classant des comportements caractéristiques.

D'autre part, la nécessité avec des sourds de travailler méthodiquement la syntaxe et la grammaire de la phrase à partir de textes complexes. Car cette syntaxe ne peut leur venir que de l'écrit lui-même et pas de la syntaxe de l'oral ou de la LSF. Les observations précédentes et leur typologie pourraient permettre de concevoir des modes d'interventions. Chemin faisant, il a été donné dans ce compte-rendu quelques indications sur des présentations plus importantes que d'autres.

Claire-Lise VELTEN ■■■