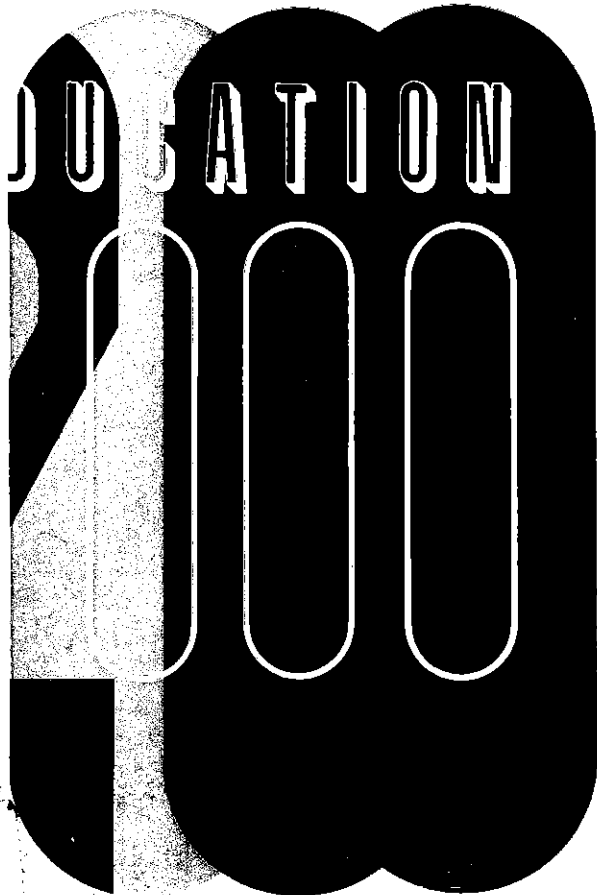


203
Gilbert DE LANDSHEERE

LE TEST DE CLOSURE



EDITIONS LABOR - Bruxelles FERNAND NATHAN - Paris

DANS LA MEME COLLECTION :

EDUCATION 2000

dirigée par G. DE LANDSHEERE

**LA PEDAGOGIE PALEOLITHIQUE
OU PREHISTOIRE DE LA CONTESTATION,**

de H. BENJAMIN.
Traduction de G. DE LANDSHEERE.

PSYCHOPEDAGOGIE DES MOINS DOUES,

de L. MASSARENTI.

PSYCHOLOGIE DE L'ENFANT ET DE L'ADOLESCENT,

de F. HOTYAT - 10^e édition.

**EVALUATION CONTINUE ET EXAMENS
PRECIS DE DOCIMOLOGIE,**

de G. DE LANDSHEERE - 4^e édition.

TECHNOLOGIE EDUCATIVE ET AUDIO-VISUEL,

de T. DECAIGNY - 3^e édition.

LE TEST DE CLOSURE,

de G. DE LANDSHEERE - 4^e édition.

COMMUNICATION AUDIO-VISUELLE ET PEDAGOGIE,

de T. DECAIGNY - 2^e édition.

CONSTRUIRE UN COURS PROGRAMME,

de D. DECLERCQ, J. DONNAY, R. DE BAL
et P. LAMBRECHT - 2^e édition.

COMMENT MESURER LA LISIBILITE,

de G. HENRY.

LA RELATIVITE EDUCATIONNELLE,

de A. CLAUSSE.

L'ETUDE EN EQUIPE,

de G. POIRIER.

CONCEPTS ET METHODES DE LA STATISTIQUE - Vol. 1 et 2,

de L. D'HAINAUT.

**LA MISE EN PLACE DES POLITIQUES EDUCATIVES :
ROLE ET METHODOLOGIE DE LA CARTE SCOLAIRE,**

de J. HALLAK.

INTRODUCTION A LA PSYCHOPEDAGOGIE DE L'EXPRESSION, Vol. 1 et 2,

de J. FABRY.

COMMENT SE DOCUMENTER,

de J.-E. HUMBLET.

LE TEST DE CLOSURE
MESURE DE LA LISIBILITE ET DE LA COMPREHENSION

DU MEME AUTEUR

Introduction à la recherche en éducation, Paris, A. Colin-Bourrellier ; Liège, Thone, 1970, 3^e éd. revue et augmentée. — Traductions : *Einführung in die pädagogische Forschung*, Weinheim, Julius Beltz Verlag, 1971, 2^e éd. ; *La investigación pedagógica*, Buenos Aires, Estrada, 1971.

Les tests de connaissances, Bruxelles, Editest, 1965.

Rendement de l'enseignement des mathématiques dans douze pays (en collaboration avec T.N. Postlethwaite), Paris, Institut Pédagogique National, 1969.

Comment les maîtres enseignent. Analyse des interactions verbales en classe, Bruxelles, Ministère de l'Education nationale, 1969. (avec la collaboration de E. Bayer).

H. BENJAMIN, *La Pédagogie paléolithique ou Préhistoire de la contestation*, Préface et adaptation française, Collection « Education 2000 », Paris, F. Nathan ; Bruxelles, Labor, 1970.

Evaluation continue et examens. Précis de docimologie, Collection « Education 2000 », Paris, F. Nathan ; Bruxelles, Labor, 1972, 2^e éd.

203
Gilbert DE LANDSHEERE
Professeur à l'Université de Liège

LE TEST DE CLOSURE

Mesure de la lisibilité
et de la compréhension

Deuxième édition

Fernand NATHAN
Editeur
PARIS

1 9 7 8

Editions LABOR
Rue Royale 342
1030 BRUXELLES

© 1978, Editions Labor, Bruxelles

D/1978/258/19.

ISBN 2-8259-0063-X.

INTRODUCTION

Exploité intensivement aux Etats-Unis, depuis une quinzaine d'années déjà, le cloze test n'a pas encore fait l'objet de recherches approfondies dans les pays de langue française.

Certes, la technique des textes mutilés est connue depuis longtemps, mais elle trouve ici une systématisation nouvelle pour être mise principalement au service de la mesure de la lisibilité ou de la compréhension de la lecture.

Instrument de recherche de grande valeur, le test de closure se prête aussi à des utilisations pratiques précieuses pour l'enseignement, dès la troisième ou la quatrième année primaire.

En effet, il permet à tous les enseignants de construire, en quelques minutes, un test de compréhension de la lecture d'une validité certaine et, au besoin, des formes parallèles, sans recours aucun aux analyses subjectives et aux manipulations statistiques.

Dans une première partie, nous avons tenté de faire la synthèse des études américaines les plus importantes afin d'esquisser un cadre théorique pour les chercheurs et les praticiens de langue française qui ne comptent pas encore le test de closure parmi leurs outils habituels.

La seconde partie rend compte d'une recherche portant sur des populations belges distribuées dans neuf années d'études : du milieu de l'école primaire à la fin de l'enseignement secondaire supérieur (9-18 ans). Cette recherche constitue un premier aboutissement d'essais limités que nous faisons depuis 1969.

Bien qu'elle porte sur presque douze mille épreuves, cette étude garde un caractère exploratoire.

Notre effort a principalement porté sur l'enseignement secondaire auquel nous voulions initialement nous cantonner. Lors de l'analyse des premiers résultats, deux observations nous ont conduit à opérer un sondage dans le primaire : la faible discrimination entre presque tous les textes, dès la première année du secondaire et, pour le degré inférieur du secondaire, la différence marquée de résultats entre l'enseignement général et le technique. Deux questions se posaient ainsi. Y aurait-il discrimination entre textes à l'école primaire ? A quel niveau se situent les résultats de l'enseignement technique du cycle inférieur par rapport à l'école primaire ?

Nous remercions chaleureusement les quarante-trois classes d'école primaire et les trente et une écoles secondaires qui nous ont ouvert leur porte, ainsi que les nombreux instituteurs et professeurs de langue maternelle qui, intéressés par notre essai et les perspectives qu'il ouvre, ont aimablement accepté de faire subir les épreuves à leurs élèves.

Notre reconnaissance va aussi à nos collaborateurs, MM. G. Henry, R. De Bal et P. Lambrecht, qui ont apporté une aide considérable pour l'exécution de la recherche ainsi que pour le traitement des données.

**

Liste des écoles qui ont participé à la recherche.

- Classes d'école primaire des communes de Chênée (16 classes), de Dalhem (3 classes), de Grivegnée (22 classes), de Liège (2 classes).
- Ecoles secondaires :
 - Alleur : Ecole moyenne de l'Etat
Ecole technique (soir)
Athénée royal
 - Andenne : Ecole moyenne de l'Etat
Ecole normale de l'Etat
 - Auvelais : Ecole industrielle
 - Esneux : Athénée royal
 - Gouvy : Ecole moyenne de l'Etat
 - Herstal : Athénée royal
 - Herve : Institut de la Providence
 - Huy : Institut Provincial d'Enseignement Technique
 - Jumet : Ecole moyenne de l'Etat
 - Liège : Athénée royal de l'Etat - Liège 1
Institut technique Don Bosco
Institut technique Saint-Laurent
Lycée royal de l'Etat
 - Morlanwelz : Institut d'Enseignement Technique
 - Ottignies : Athénée royal
 - Saint-Georges : Athénée royal
 - Saint-Gilles : Athénée
 - Stavelot : Athénée royal
 - Tamines : Athénée royal
Lycée royal
 - Verviers : Ecole normale de l'Etat
Institut Provincial d'Enseignement Technique
 - Visé : Ecole moyenne de l'Etat
 - Waremme : Institut Provincial d'Enseignement Technique
 - Woluwé-Saint-Pierre : Athénée royal

PREMIERE PARTIE

LE TEST DE CLOSURE

1. DEFINITION ET HISTOIRE

Le mot *closure*, employé en français comme en anglais, doit être rapproché du verbe *clôturer*. Il a été employé par Thurstone pour désigner une *structuration perceptive*.

Henri Piéron écrit : « La vitesse de "closure", de *saisie*, est celle avec laquelle on devine une lettre ou un dessin, incomplètement esquissés. La souplesse (*flexibility*) de closure correspond à la facilité avec laquelle on peut substituer à une certaine structure perçue une structure différente (quand on passe, par exemple, de la perception du dessin d'un vase à celle d'un profil humain dessiné par les contours du vase). Le concept de clôture est d'origine gestaltiste, une "bonne forme" devant toujours être fermée (Koffka) ¹. »

Dans les notes qui suivent, nous entendons uniquement par *closure* l'acte subjectif de combler des lacunes ménagées dans un texte, par suppression de lettres, de mots ou de symboles.

Le test de *closure* (*cloze test*) a été inventé par W.L. Taylor, en 1953 ². Il le définit comme : « Un outil psychologique permettant de jauger le degré de correspondance totale entre les habitudes d'encodage d'émetteurs et les habitudes de décodage des récepteurs. » E.F. Rankin écrira plus tard : « Le test de closure est donc une mesure de la correspondance entre le langage du lecteur et celui de l'écrivain ³. » La procédure ne doit pas être confondue, soit avec les épreuves de complétion de phrase (par exemple, le test de Paul, de D. Bonnet), soit avec des épreuves de closure utilisées, notamment, dans certains tests de maturité spécifique pour l'apprentissage de la lecture (par exemple, lettres ébauchées à compléter, figures lacunaires).

Le test de closure, conçu par Taylor comme nouvel instrument de mesure de la lisibilité, consiste à supprimer un mot sur cinq dans un texte. Les sujets doivent reconstituer le texte original.

1 H. PIERON, *Vocabulaire de la psychologie*, Paris, P.U.F.

2 W.L. TAYLOR, Cloze Procedure : A New Tool for Measuring Readability, in *Journalism Quarterly*, Fall 1953, 115 sq.

3 E.F. RANKIN, The Cloze Procedure. Its Validity and Utility, in R. FARR, *Measurement and Evaluation of Reading*, New York, Harcourt, Brace and World, 1970, 237-253.

La même technique est aussi utilisée pour mesurer la compréhension des textes. On trouve en effet une corrélation élevée entre les résultats aux meilleurs tests de lecture et au test de closure ¹.

Enfin, Rankin, puis Bloomer ont repris l'instrument à des fins d'enseignement, en particulier pour améliorer la capacité en compréhension de la lecture.

J.B. Carroll ² signale que la technique de closure a déjà été utilisée vers 1897 par H. Ebbinghaus ³. A l'origine, il essaye, par ce moyen, de mesurer la fatigue chez l'enfant (qui devait restituer des syllabes supprimées dans une traduction allemande des *Voyages de Gulliver*). Il constata que l'exercice de closure était une pauvre mesure de la fatigue, mais que ses résultats étaient hautement corrélés avec l'âge et les performances scolaires ⁴.

Quelque cinquante ans plus tard (1939), J.B. Carroll a démontré, par analyse factorielle, qu'un exercice de complètement du type

Quant à

où il faut fournir la première réponse qui vient à l'esprit et où le degré de probabilité d'apparition du mot manquant (déterminé sur un échantillon normatif) constitue le critère de correction :

1° est une bonne mesure de la connaissance de la langue maternelle ;

2° manque, au niveau individuel, de fidélité.

1 Parmi les travaux fondamentaux de J.R. BORMUTH, on se référera d'abord à *The Implication and Use of Cloze Procedure in the Evaluation of Instructional Programs*, Los Angeles, U.C.L.A., 1967, 47 p.

2 J.B. CARROLL et al., *An Investigation of Cloze Items in the Measurement of Achievement in Foreign Languages*, Cambridge, Mass., Harvard University Laboratory for Research in Instruction, 1959.

3 H. EBBINGHAUS, Über eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern, in *Zeitschrift für Psychologie*, 1897, 13, 401-459 (cité par Carroll).

4 J. VAN BIERVLIET, *Pédagogie expérimentale*, Paris, Alcan, 1912, p. 259, reprend le test d'Ebbinghaus. D'après Binet, — qui, lui aussi, s'est intéressé à la technique, — il propose l'exemple suivant : « Depuis plus un mois, la santé mini tou chancel!, était profond al ; c'était de lit malu en proie cru dou! qu'il diri à la les armées et pro de Cinq-Mars. »

On observe qu'ici, tantôt des mots, tantôt des syllabes sont supprimés : l'intervalle entre lacunes est irrégulier. J. Van Biervliet conclut aussi au peu d'intérêt de la technique pour la mesure de la fatigue.

2. DESCRIPTION DU TEST

I. Les lacunes

Comme on vient de l'indiquer, Taylor, actuellement suivi par la majorité des chercheurs, supprime un mot sur cinq.

Les mots manquants sont remplacés par des lacunes de longueur uniforme (quinze espaces). L'attention est attirée sur les lacunes, soit par des cadres, soit par un soulignage¹.

Ce procédé simple permet la construction aisée de cinq formes d'épreuves parallèles. Dans la première, on supprime les mots 1, 6, 11, 16, etc. ; dans la deuxième, les mots 2, 7, 12, 17, ... ; et ainsi de suite. Quand cinq formes possibles ont été administrées, tous les mots des textes ont ainsi été testés.

Si l'on utilise une seule forme, la suppression 5, 10, 15, ... doit être préférée car, de cette façon, le premier mot est précédé et suivi d'une même longueur de contexte. Pour fournir aux sujets un contexte plus sûr, certains chercheurs ne mutilent ni la première ni la dernière phrase du passage testé. G. Klare² n'estime pas cette précaution nécessaire, sauf pour de jeunes élèves, pour des sujets à peu près illettrés et, en général, pour des sujets qu'il importe d'aider beaucoup.

Pour des textes assez longs, M. Hater (700 mots), R. Klare et d'autres auteurs n'ont pas observé de différences significatives entre les résultats obtenus aux diverses formes.

Pour les textes courts, une différence significative de rendement peut cependant exister. La présente étude a notamment pour objectif de préciser ce problème. Dès maintenant, on peut cependant prendre pour règle générale qu'il ne faut pas se contenter d'une seule forme pour mesurer la lisibilité d'un texte, mais qu'une seule forme suffit pour évaluer la capacité en lecture d'un sujet.

La règle adoptée par Taylor n'est pas suivie par tous. Certes, il n'est guère possible de supprimer plus d'un mot sur cinq. W. Mac Ginnitie a montré³ qu'en dessous de cette limite, la réussite à un *item* donné devient trop dépendante de la réussite aux *items* voisins. Autrement dit : un sujet incapable de combler une lacune risque d'être mis, par ce fait, dans l'impossibilité de combler la

1 Un test réel figure en annexe.

2 G.R. KLARE et al., *The Cloze procedure : a convenient readability test for training materials and translations*, Arlington, Virginia, Institute for Defense Analysis, 1971, p. 15.

3 W. Mac GINNITIE, Contextual Constraint in English Prose Paragraphs, in *Journal of Psychology*, 51, 1961, 121-130.

suiuante. Par contre, il ne semble pas exister de différence importante entre 1/5 et 1/7¹. D'autres auteurs ont travaillé avec des lacunes de 1/10.

Pour tester, on a auantage à obtenir le plus grand nombre d'*items* dans le temps le plus réduit possible ; il est ainsi aisé d'augmenter la fidélité et, souvent aussi, la validité de l'épreuve. C'est pourquoi la proportion minimale de 1/5 est la plus généralement adoptée.

Deux autres méthodes, utilisées par certains linguistes, constituent de bons outils de recherche, mais seruent beaucoup moins bien les objectifs pratiques que nous poursuivons. La première consiste à supprimer tous les mots d'une classe donnée : par exemple, tous les aduerbes. On peut ainsi déterminer le rôle spécifique joué par les différentes classes de mots dans la difficulté du texte. La seconde méthode est parfois appelée « closure totale ». C'est le fameux *guessing game* où le sujet est invité à deviner tous les mots d'un texte. L'essai porte d'abord sur le premier mot ; s'il n'est pas trouvé, on le fournit. On passe de même au deuxième mot, puis au troisième et ainsi jusqu'à la fin. Le nombre de réussites pour cent mots constitue le résultat. Dans cette procédure, la contrainte est unilatérale ; elle convient spécialement au message parlé qui, à un moment donné, ne peut jamais s'éclairer que par ce qui précède.

II. Règles pour la suppression des mots

Quelques règles simples doivent être respectées pour assurer l'uniformité des mesures. La ligne de conduite suivante s'inspire des règles que nous auons adoptées pour calculer les scores de lisibilité Flesch :

1. En principe, on considère comme mot tout ensemble séparé des autres par des espaces blancs (*U.N.E.S.C.O.*, 1972, *l'hirondelle, aujourd'hui* ne comptent donc que pour un seul mot).
2. Deux mots unis par un trait d'union ne sont traités séparément que s'ils peuvent être utilisés isolément dans la langue. On traitera donc *co-président* comme un seul mot, mais *bateau-mouche* comme deux mots. Dans ce dernier cas, le trait d'union devra être maintenu dans le texte mutilé.
3. Pour un nombre de plusieurs chiffres écrits en lettres, chaque partie constitue un mot indépendant.

Dans plusieurs recherches faites à Liège, dont les importants travaux de G. Henry, les formes élidées ont été traitées séparément : *l'hirondelle y* est donc considéré comme deux mots. Cette convention simplifiée a d'ailleurs été adoptée dans le présent travail.

1 Ibid.

III. Les consignes ¹

Les sujets sont invités à parcourir d'abord tout le texte mutilé, sans combler les lacunes. Ils acquièrent ainsi une connaissance globale du contenu.

Après un court entraînement, on passe ensuite au travail de complètement. Le plus souvent, le temps n'est pas limité d'avance.

IV. Correction

Comme déjà indiqué, seul le mot existant dans le texte original est considéré comme correct. Accepter les synonymes conduit rapidement à des divergences entre correcteurs et allonge considérablement le travail. Ce risque et ce surcroît ne sont même pas compensés par un avantage appréciable. E. Rankin a montré qu'accepter les synonymes n'augmente ni la validité ni la fidélité du test ², tandis que W. Taylor n'observe pas d'accroissement de la discrimination dans la mesure de la lisibilité ³.

On accepte les fautes d'orthographe d'usage qui n'empêchent pas d'identifier le mot avec certitude.

Les fautes d'orthographe grammaticale entraînent la perte du point.

Comme pour les tests courants, une grille facilite la correction.

E. Jongsma ⁴ estime que, dans le cas où la technique de closure est utilisée à des fins d'enseignement, la façon de calculer les scores devrait évoluer. On accepterait d'abord comme correct tout mot appartenant à la même classe grammaticale que le mot supprimé, puis les synonymes seulement ; enfin, on exigerait le mot original.

Nous manquons d'expérience pour nous prononcer sur l'intérêt de cette suggestion.

V. Scores

Le score est égal au pourcentage d'*items* réussis.

L'*indice de difficulté* d'un texte est égal à la moyenne des scores obtenus par un échantillon représentatif d'une population donnée.

1 Voir exemple en annexe.

2 On peut augmenter légèrement la fidélité en acceptant des synonymes dont on a déterminé la fréquence d'occurrence au cours d'une recherche préliminaire. Mais le gain de fidélité n'est pas suffisant pour justifier ce lourd travail préparatoire.

3 E.F. RANKIN, *The Cloze Procedure*, o.c., p. 251.

4 E. JONGSMA, *The Cloze Procedure as a Teaching Technique*, Newark, Del., International Reading Association, 1971, p. 24.

Selon le critère proposé en 1917 par Thorndike (confirmé par Betts en 1946 et par Harris en 1962), on peut considérer que, pour le travail dirigé, un texte convient à un élève quand celui-ci répond correctement à au moins 75 % des questions de compréhension ; pour un travail indépendant, la limite est portée à 90 %.

Dans une première recherche, J. Bormuth a observé que si l'on évalue un même texte par un test de compréhension de la lecture (choix multiple ou courtes réponses construites) et par un test de closure, un score d'environ 44 % en closure correspond à 75 % de succès au test de lecture classique ; 55 % de réussite en closure correspondrait à 90 % de réussite au test de compréhension¹. En outre, Bormuth observe que 33 % de réussite en closure correspond à 95 % de lecture orale correcte des mots. Ce chiffre élevé s'explique par la possibilité de prononcer correctement des mots incompris.

La signification des 44 % de réussite est confirmée à la même époque par E. Rankin et J. Culhane².

En juin 1971, J. Bormuth termine, pour le *U.S. Office of Education*, une nouvelle recherche qui apporte, nous semble-t-il, le progrès le plus décisif, depuis Thorndike, en matière de critères de performance souhaitable pour des textes didactiques³.

En un passage d'une remarquable densité, Bormuth indique les principales variables dont il importe de tenir compte pour déterminer l'adéquation des textes : « Une taxonomie assez complète de ces variables (...) comprendrait des variables cognitives telles que l'apprentissage, la rétention et le transfert des informations contenues dans le texte, des variables d'habileté comme la vitesse de lecture et la latence des réponses apprises par cette lecture ; des variables affectives telles que la préférence du sujet pour le contenu, le style, la difficulté du texte et l'empressement à l'étudier ; des facteurs économiques tels que la variation du coût de l'instruction selon le niveau de capacité en lecture auquel on veut conduire l'étudiant et le coût de la préparation du matériel de lecture adéquat ; des facteurs psychosociaux, comme l'influence que l'on exerce sur les attitudes et la perception de soi-même (*self-concept*) chez l'élève, selon le degré de difficulté des textes qu'on lui fait lire, considéré par rapport à sa capacité en lecture⁴. »

1 J. BORMUTH, Cloze Test Readability: Criterion Reference Scores, in *Journal of Educational Measurement*, 5, 1968.

2 E. RANKIN et J. CULHANE, Comparable Cloze and Multiple Choice Comprehension Test Scores, in *Journal of Reading*, 13, 1969, p. 8.

3 J. BORMUTH, *Development of Standards of Readability: Toward a Rational Criterion of Passage Performance*, Chicago, Univ. of Chicago ; Washington, U.S. Office of Education, 1971.

4 *Idem*, p. VIII.

Pour obtenir les critères dont il va être question, Bormuth n'a pas pu tenir compte de toutes ces variables. Il a retenu : le gain d'information, la vitesse de lecture, l'empressement à l'étude, la préférence pour le contenu, le style et le degré de difficulté ¹.

Bormuth aboutit à des scores-critères beaucoup plus nuancés que ceux proposés par ses prédécesseurs et par lui-même, en 1968. Non seulement, il fait une distinction entre les manuels à utiliser en classe, les livres à lire seul et les livres de référence, mais, en outre, il distingue chaque année scolaire, de la 3^e primaire à la fin du secondaire supérieur.

Le tableau ci-dessous indique les niveaux de difficulté optimum, exprimés en scores de closure moyens ².

		Manuels scolaires	Livres de référence	Livres à lire seul
Primaire	3 ^e année	54	52	62
	4 ^e année	56	54	58
	5 ^e année	57	54	55
	6 ^e année	57	52	52
Secondaire	1 ^{re} année	54	50	50
	2 ^e année	51	47	47
	3 ^e année	48	44	45
	4 ^e année	46	42	42
	5 ^e année	45	41	39
	6 ^e année	48	41	36

Le fait que, presque dans tous les cas, le score-critère soit nettement supérieur aux 44 % proposés antérieurement, ne doit pas surprendre en raison même de la multiplicité des variables dont on a tenu compte, au lieu de s'en tenir à la simple compréhension, mesurée d'ailleurs de façon fort approximative.

Il paraît néanmoins évident que tous les pourcentages qui viennent d'être cités ne peuvent être pris de façon absolue. D'abord, parce que les sciences humaines ne connaissent pas de limites aussi nettes. Ensuite, parce que toutes les recherches citées ont été faites en anglais, sur des populations américaines.

1 On constate, en effet, qu'un sujet donné lit de préférence des textes qui se situent dans une bande de difficulté déterminée. Ce phénomène est encore mal étudié en psychologie. Peut-être, les théories de McLelland mettraient-elles sur le chemin de l'explication.

2 Cf. Bormuth, *Development...*, o.c., pp. 138-139.

D'après les travaux faits actuellement au laboratoire de Pédagogie expérimentale de l'Université de Liège et les résultats du présent travail, les scores-critères à retenir pour le français se situeraient en dessous des critères américains.

VI. Etalonnage

Le test de closure s'étalonne de la même façon que les autres.

Il n'est guère possible d'administrer les cinq formes parallèles à un même groupe de sujets. Dans ce cas, les *items* utilisés ne sont qu'un échantillon de tous les items possibles¹.

VII. Echantillonnage des textes

Pour des textes de moins de 500 mots, il n'est pas conseillé d'échantillonner. Pour de plus gros volumes, G. Klare *et al.* conseillent² :

1. De choisir au hasard au moins deux échantillons par section ou chapitre. L'homogénéité est ainsi mieux appréciée.
2. De toujours prendre les échantillons à partir du début d'un paragraphe. Sans cette uniformité, les structures lexicales peuvent varier d'un échantillon à l'autre et l'évaluation de la difficulté peut ainsi être faussée.
3. Chaque échantillon devrait permettre au moins 50 lacunes.³

VIII. Fidélité

Pour un même individu, on observe parfois une fidélité légèrement inférieure à celle des tests de compréhension classiques ; néanmoins, la fidélité est presque toujours très élevée.

Il semblerait qu'au niveau individuel, le test de closure soit particulièrement sensible à certains traits de personnalité. E. Rankin

1 Pour la mesure de la lisibilité, il faut donc tenir compte de l'erreur d'échantillonnage des *items*, qui vient s'ajouter à l'erreur d'échantillonnage de la population.

Si l'on peut administrer au moins deux formes parallèles :

- a) On calcule la variance des moyennes des deux formes ;
- b) En soustrayant, du résultat obtenu, la variance de l'erreur d'échantillonnage de la population, on trouve une estimation de l'erreur d'échantillonnage des *items*.

La formule 21 de Kuder-Richardson, pour calculer la fidélité, tient compte de l'erreur de l'échantillonnage des *items*.

2 G.R. KLARE *et al.*, *The cloze procedure : a convenient readability test for training materials and translations*, Arlington, Virginia, Institute for Defense Analysis, 1971.

3 L'erreur d'échantillonnage des textes est trouvée en déterminant la difficulté de chaque passage retenu, en calculant la variance des scores obtenus et en soustrayant, du résultat trouvé, la variance de l'erreur d'échantillonnage de la population et la variance de l'erreur d'échantillonnage des *items*.

a montré, en particulier, l'incidence négative d'une anxiété supérieure à la moyenne¹. Il remarque toutefois que l'influence de la personnalité sur la validité et la fidélité des autres types de tests est mal connue.

En administration collective, W. Taylor (1957) signale une fidélité test-retest allant de .74 à .88 pour une épreuve de 80 *items*, administrée en 40 minutes. G. Klare *et al.* constatent de leur côté une fidélité entre .76 et .94 tant pour des enfants que pour des adultes².

Dans sa recherche de 1962³, J. Bormuth a trouvé une fidélité équivalente pour neuf tests de 31 *items* à choix multiple et pour neuf tests de closure de 50 *items* qu'il avait construits.

J. Bormuth estime que la fidélité du test de closure peut être légèrement inférieure à celle des tests classiques en raison d'un certain nombre d'*items* très difficiles ou très faciles. Il écrit : « Toutefois, le grand nombre d'*items* très difficiles ou très faciles rencontrés dans les tests de closure offre un avantage, car le test est ainsi applicable à des sujets de capacités très différentes. Des scores nuls ou maximum et des distributions anormales sont rarement observés quand les épreuves sont soigneusement administrées. »

Le champ d'application n'est cependant pas illimité. Gallant a démontré la faible fidélité du test de closure au début de l'école primaire⁴.

On sait que R. Bloomer estime que le test de closure ne peut pas être subi par des sujets d'un âge mental inférieur à dix ans⁵. S. Ramanauskas a cependant pu descendre de façon répétée jusqu'à six ans d'âge mental (le *skill* de la lecture étant évidemment acquis). Il est d'ailleurs intéressant de noter que Ramanauskas a essentiellement travaillé avec des handicapés mentaux, ce qui semble ouvrir des perspectives nouvelles pour l'enseignement spécial⁶.

1 E.F. RANKIN, o.c., p. 252.

2 G. KLARE *et al.*, o.c., p. 11.

3 Voir J. BORMUTH, *The Implication and Use of Cloze Procedure in the Evaluation of Instructional Programs*, document ronéotypé, s.d., p. 12.

4 R. GALLANT, An Investigation of the Use of Cloze Tests as a Measure of Readability of Materials for the Primary Grades, in J.A. FIGUREL, Ed., *Reading and Inquiry*, Newark, Del., International Reading Ass., 1965, 286-287.

5 R. BLOOMER, Cloze Symposium, in G. SCHICK et M. MAY, *Multidisciplinary aspects of college-adult reading*, 17th Yearbook of the National Reading Conference, 1968.

6 S. RAMANAUSKAS, Oral Reading Errors and Cloze, in J. CAWLEY *et al.*, *The slow learner and the reading problem*, Springfield, Ill., Chas. Thomas & Co, 1972.

Comme on le verra par la suite, des mesures particulières doivent de toute évidence être prises quand on a affaire à de jeunes élèves. Pour l'usage courant, la limite fixée par R. Bloomer paraît sage.

Quoi qu'il en soit, et pour en revenir au problème de la fidélité, c'est encore R. Bloomer qui propose un sage repère en recommandant que les textes testés comptent au moins 250 mots et donc 50 *items*¹.

Discutant de la fidélité de leur test de closure portant sur des textes de mathématiques, M. Hater et R. Klane² apportent des informations importantes qui justifient la longue citation suivante :

« Dans tous les cas, les coefficients de fidélité ont été supérieurs à 0,90 (formule 20 de Kuder-Richardson). On peut donc conclure que les tests de closure ont une fidélité très élevée.

Nous avons comparé la fidélité des réponses fournies aux mêmes tests de closure par des groupes peu et très nombreux. (...) La fidélité obtenue pour des groupes de 30 est similaire à celle que l'on trouve pour des groupes de 200 sujets.

La longueur du test a certainement contribué à la fidélité élevée. Quelques *items* seulement présentaient une difficulté de 1 ou de 0. Après élimination de ces extrêmes, chaque test comptait encore plus de cent *items*. Avec un aussi grand nombre d'*items* offrant un large spectre discriminatif, les fidélités devaient normalement être élevées.

Contrairement aux tests à n choix multiples où un certain nombre de réponses correctes sont dues simplement au choix aveugle, une réponse correcte au test de closure résulte d'un choix correct parmi des mots existant dans l'esprit du sujet ou d'indices fournis par le contexte proche de la lacune. Comme l'éventail des réponses possibles à partir de l'expérience du sujet et des indices existant dans le test est grand, l'erreur due à la possibilité de deviner la réponse au pur hasard est petite. (...)

Pour quatre des cinq tests de closure, le pourcentage moyen des sujets répondant aux cinq derniers *items* de chaque test était similaire au pourcentage moyen des sujets répondant à cinq *items* choisis au

1 R. BLOOMER, *Non-overt reinforced cloze procedure*, Storrs, Univ., of Connecticut, 1966.

2 Leur recherche est synthétisée plus loin.

hasard dans le reste du test. Pour un test, les pourcentages étaient différents. Le fait que certains sujets abandonnent avant d'autres peut augmenter à la fois la valeur discriminative des derniers *items* du test et la variance ; il en résulte une augmentation apparente de la fidélité.

Bref, les caractéristiques intrinsèques du test de closure qui semblent contribuer à une fidélité élevée sont : la longueur du test, l'absence presque totale du facteur de conjecture aveugle et la distribution de la difficulté des *items*. Un facteur extrinsèque qui pourrait avoir augmenté artificiellement la fidélité, pour un des cinq tests, est le fait que certains élèves n'ont pas terminé l'épreuve. »

IX. Champ d'application

Pour les principaux types d'application qui seront soit illustrés par la suite, soit discutés au chapitre de la validité, nous nous limitons à de brèves indications.

A. Mesure de l'intelligibilité

1. Mesure de la lisibilité :

a) Mesure directe.

b) Mesure-critère utilisée pour la construction de formules.

2. Mesure de l'audibilité :

Le test de closure est administré oralement.

3. Mesure de l'audio-visibilité :

N.B. Les domaines 2 et 3 ont, jusqu'à présent, fait l'objet de peu de recherches.

B. Mesure de la capacité de lecture

1. Mesure de la compréhension générale

Par compréhension générale de la lecture, on entend une sorte de niveau moyen de capacité, déterminé chez le lecteur mis en face d'un échantillon de tâches de lecture supposé représentatif de toutes les tâches possibles. C'est ce niveau moyen que les tests de lecture classiques essaient de déterminer.

2. Mesure de la compréhension spécifique

La capacité de comprendre un texte particulier (un article, un manuel scolaire) est ici mesurée.

Le test de closure est unanimement considéré comme très efficace pour ce type de mesure.

C. Instrument pronostique

A. Kingston et W. Weaver¹ ont examiné le pouvoir prédictif de la technique de closure utilisée en première année primaire.

L'instituteur familiarise progressivement ses élèves avec la technique en supprimant des mots lexicaux des textes de lecture figurant au tableau et en invitant les élèves à proposer tous les mots susceptibles de combler significativement la lacune. Au début, des images sont placées comme indices au-dessus des lacunes. Petit à petit, on passe à des textes lacunaires dactylographiés.

Quatre types de tests de closure sont utilisés :

1. Cinq textes présentant chacun 10 lacunes du n^{me} mot.
2. Closure structurale : suppression de mots-fonctions et réponse à deux choix, le distracteur appartenant à la même classe grammaticale que le mot supprimé.
3. Closure lexicale (noms, verbes, adjectifs) à choix multiple comme en 2.
4. Test de closure oral avec suppression de tous les n^{mes} mots.

L'instituteur lit à haute voix ; les élèves suivent sur le texte écrit.

A chaque lacune, le maître s'arrête 30 secondes et les élèves doivent écrire le mot manquant.

Pendant l'année scolaire, les élèves subissent une série de tests : deux tests de *readiness* (*Ginn Pre-Primer Achievement Test* ; *Lee-Clark Readiness Test*), administrés la première semaine de la rentrée scolaire, les divers tests de closure (subis en février-mars), et le *Ginn Primary Achievement Test*. En fin d'année, on utilise la partie « langage » du *California Achievement Test* et les résultats à ce dernier servent de mesure de critère pour une analyse de régression (step-wise).

Le meilleur prédicteur unique est le test de closure lexical (R multiple = .68) ; avec, en 2^e rang, le test de closure portant sur tous les n^{mes} mots, on obtient un R = .73.

Autre résultat intéressant de cette recherche : la technique de closure se révèle utilisable dès le début de l'apprentissage de la lecture, pour autant que l'on y entraîne systématiquement les élèves.

Comme le note pertinemment E. Jongsma², l'introduction de choix multiples dans certains des tests de closure altère considérablement la technique initialement proposée par Taylor.

1 A. KINGSTON et W. WEAVER, *Feasibility of Cloze Techniques for Teaching and Evaluating Culturally Disadvantaged Beginning Readers*, University of Georgia, R. and D. Center in Early Educational Stimulation, 1970.

2 O.c., p. 18.

D. Instrument diagnostique

En le combinant avec la technique de la réflexion parlée, le test de closure devient un instrument diagnostique. Invité à décrire sa démarche mentale, à s'introspecter pendant qu'il complète le test, le sujet peut donner des indications précieuses sur ses réactions personnelles (tension, anxiété, ...), sur sa capacité à utiliser des indices syntaxiques, à combiner les significations, à dégager les implications ; des observations sur la connaissance de la signification des mots et des structures du langage ainsi que sur la sensibilité aux nuances de style sont aussi possibles¹.

La simple analyse des *items* d'un test de closure permet aussi de déterminer si, pour une catégorie particulière de mots ou pour une relation à établir, le sujet obtient des résultats normaux par rapport à un étalonnage approprié.

A. Etienne² a observé individuellement 24 élèves de 11 à 15 ans passant le test de closure. Ils avaient été choisis en fonction de leur différent degré de réussite à des tests de closure précédents, de façon à couvrir une gamme de capacité aussi large que possible. Les commentaires ont été enregistrés.

Les élèves lisaient à haute voix et indiquaient simultanément les mots lus (les élèves interrompent la lecture orale avant chaque lacune, mais lisent encore silencieusement quelques mots au-delà, d'où l'intérêt des deux indications).

A. Etienne observe que les élèves faibles ne tentent généralement pas de saisir d'abord le sens global d'un paragraphe ou d'une phrase, mais ne considèrent qu'un groupe restreint de mots précédant la lacune. La mémoire immédiate joue donc à plein.

S'ils ne trouvent pas de solution, ou bien les élèves sautent le mot ou bien ils retournent en arrière pour trouver de nouveaux indices. Mais ils ne tentent pas de structurer systématiquement la phrase et encore moins d'analyser les diverses relations. Il n'y a donc pas de transfert entre l'analyse textuelle systématique telle que les professeurs la pratiquent avec leurs élèves et la solution du problème de closure.

Une seconde lecture spontanée n'apporte pas un progrès sensible.

Si la réponse est incorrecte, le mot fourni peut néanmoins appartenir à la même catégorie grammaticale que le mot man-

1 Voir à ce propos la recherche de M. JENKINSON, *Selected Processus and Difficulties in Reading Comprehension*, Univ. of Chicago, 1957, cité par E.F. RANKIN, o.c., p. 250.

2 A. ETIENNE, *Analyse des démarches et des types d'erreurs dans le test de closure*, Université de Liège, 1972 (mémoire inédit).

quant. En 6^e primaire, les élèves reconnaissent exactement la catégorie grammaticale de la réponse attendue dans 71,5 % des cas pour un texte difficile (score Flesch \pm 22) et dans 82,5 % des cas pour un texte légèrement supérieur à leur niveau (score Flesch \pm 47). En 4^e moderne, les résultats sont de 80 % pour un texte légèrement supérieur à leur niveau et de 91,5 % pour un texte inférieur à leur niveau.

A l'école primaire, le pourcentage de réponses grammaticalement correctes est :

- pour le groupe : noms, déterminants, verbes, adjectifs, pronoms, prépositions 84,5 %
- pour le groupe : conjonctions et adverbes 46 %

Dans l'enseignement secondaire, les pourcentages correspondants sont respectivement : 90 % et 60 %.

Les élèves ne tiennent pas toujours compte des indices fournis par la ponctuation. Il faut toutefois remarquer que la présentation du test de closure rend peut-être la perception de la ponctuation plus difficile.

On observe aussi que les indices proches relatifs au genre et au nombre ne sont pas toujours respectés en raison de l'influence de contraintes sémantiques très fortes.

Les meilleurs résultats sont naturellement obtenus par les rares élèves qui considèrent la phrase entière ou le paragraphe. On semble donc en droit d'opposer un comportement mécanique à un comportement plus intelligent.

Enfin, les erreurs inhérentes à des faiblesses dans la technique de lecture (omission d'un mot, insertion d'un mot, substitution d'un mot à un autre, conduites de persistance) modifient les indices dont les élèves disposent et expliquent aussi un certain nombre d'erreurs.

L'ensemble de ces observations confirme qu'à partir d'exercices de closure, on pourrait probablement procéder à des exercices systématiques d'utilisation d'indices et d'analyse textuelle susceptibles d'améliorer la lecture. Des tests de closure étalonnés serviraient à évaluer les progrès accomplis.

E. Recherche linguistique

Les travaux de J. Bormuth, dont nous reproduirons un extrait par la suite, montrent qu'en administrant les cinq formes possibles d'un texte dont on supprime tous les cinquièmes mots, on peut étudier le rôle joué par les différentes catégories de mots. Il faut toutefois insister sur le fait que les observations ainsi recueillies

n'ont pas de valeur causale. Les indices obtenus sont de simples indicateurs de la contribution de la catégorie considérée à la difficulté.

F. Test d'intelligence verbale

En 1959, J.B. Carroll, un des meilleurs spécialistes contemporains de l'analyse factorielle, conclut que le test de closure permet une des plus pures mesures qui soient du facteur v^1 .

En 1957, W. Taylor² avait établi une corrélation de .85 avec le sous-test de compréhension des mots et de .76 avec le sous-test de raisonnement arithmétique de l'*Air Force Qualification Test*.

Un test de closure a été inclus par F. Weinfeld³ dans une batterie de 28 épreuves administrées à 190 garçons et 154 filles de 9^e, 10^e et 11^e années d'études.

Cette recherche a montré une corrélation assez élevée entre le test de closure et plusieurs facteurs cognitifs bien connus : r , v , fluidité d'expression et résultats en rédaction.

Cette dernière observation est corroborée par une de nos constatations : le test de closure est bon prédicteur des résultats en langue maternelle dans notre enseignement secondaire.

Dans le même ordre d'idées, R. Guice⁴ a trouvé une corrélation de .79 entre les résultats à deux tests de *créativité* de Guilford et al.⁵ et seize exercices de closure portant de façon aléatoire sur des noms, des verbes, des adjectifs et des adverbes.

Cette observation ne nous surprend pas, étant donné le rôle que le raisonnement logique (r) et la fluidité verbale (w) doivent normalement jouer dans l'épreuve de closure. Des recherches ultérieures, — aisément concevables, — devront préciser ce point.

Enfin, une nuance d'un grand intérêt est apportée par J. Schneyer⁶ qui a construit des exercices de closure de deux types : dans le premier, on supprime tous les 10^{es} mots ; dans le second,

1 J.B. CARROLL, An Investigation of Cloze Items..., o.c.

2 W.L. TAYLOR, Cloze readability scores as indices of individual differences in comprehension and aptitude, in *Journal of Applied Psychology*, 1957, 41, 19-26.

3 F.D. WEINFELD, A Factor Analytic Approach to the Measurement of Differential Effects of Training : An Evaluation of Four Methods of Teaching English Composition, University of Harvard, Thèse doctorale inédite, 1959, citée par J.B. Carroll.

4 R. GUICE, The use of the cloze procedure for improving reading comprehension of college students, in *Journal of Reading Behavior*, 1, 1969, 81-92.

5 J. GUILFORD, MERRIFIELD et CHRISTENSEN, *Tests of Creativity : Consequences - Alternate uses*.

6 J.W. SCHNEYER, Use of the Cloze Procedure for Improving Reading Comprehension, in *Reading Teacher*, 19, 1965, 174-179.

on supprime les noms et les verbes (dichotomie lexicale-structurale).

Schneyer constate une corrélation de .63 entre les exercices de premier type et l'intelligence verbale, et une corrélation de .42 pour les exercices du second type ; ces derniers testeraient donc mieux la capacité linguistique. Cette dernière constatation est confirmée par B. Guice qui obtient une corrélation de .40 dans des conditions similaires.

G. Technique d'enseignement

Plusieurs chercheurs ont cru aussi trouver, dans la technique de closure, un précieux instrument d'enseignement et un exercice correctif efficace. Ces espoirs ont, jusqu'à présent, été le plus souvent déçus. L'explication première paraît simple : on a eu la naïveté de croire à l'efficacité d'exercices de complètement, effectués sans encadrement d'enseignement ou de traitement.

Cependant, tant l'expérience empiriquement acquise par les enseignants que certains résultats expérimentaux portent à croire que la technique de closure peut être un utile auxiliaire pédagogique. Ici aussi, il appartient aux chercheurs d'en préciser les possibilités. Avant d'engager de nouveaux travaux en ce sens, la lecture de l'ouvrage critique de Jongsma est conseillée. Nous en extrayons les quelques considérations suivantes¹.

Parmi les auteurs ayant cru, en quelque sorte, au pouvoir automatique de la closure, on compte Roossinck (1962), Bloomer (1962), Friedman (1964), Schneyer (1965), Blumenfeld et Miller (1966), Heitzman et Bloomer (1967).

Par exemple, R. Bloomer² tente d'améliorer la capacité de lecture de 127 étudiants universitaires de la façon suivante. Il crée trois groupes :

1. Le premier fait des exercices de closure où tous les dixièmes mots sont supprimés.
2. Le second fait des exercices correctifs classiques.
3. Le troisième ne subit aucun traitement.

Les textes utilisés pour la closure étaient extraits de manuels élémentaires de sciences naturelles et de sciences sociales ; ils étaient de difficulté croissante et, pour pouvoir passer au texte supérieur, le sujet devait réussir au moins 96 % des *items* du texte précédent. Les sujets corrigeaient eux-mêmes les exercices.

1 E. JONGSMA, *The Cloze Procedure as a Teaching Technique*, o.c.

2 R. BLOOMER, *The Cloze Procedure as a Remedial Reading Exercise*, in *Journal of Developmental Reading*, 5, 1962, 173-181.

En utilisant le *Diagnostic Reading Test* comme prétest et post-test, Bloomer constate que les étudiants qui ont fait les exercices de closure ont progressé plus significativement en compréhension de la lecture que ceux du deuxième groupe. Mais on peut difficilement accorder une valeur à cette observation. Outre plusieurs faiblesses expérimentales, l'effet à moyen terme du traitement n'a pas été contrôlé.

Avec la recherche de R. Martin¹, on arrive sur un terrain pédagogique plus sûr. Ici, après chaque exercice de closure, les élèves discutent de leurs réponses avec le professeur. Les résultats sont nettement positifs.

R. Martin travaille avec 142 étudiants de 1^{re} année d'université : 50 étudient *l'Introduction à la grammaire transformationnelle* de Jacobs et Rosenbaum ; 50 font 18 exercices de closure ; 42 suivent simplement le cours d'anglais donné habituellement aux élèves du Carolina State College.

Ces trois groupes étaient de force équivalente en intelligence verbale (SAT). Les lacunes portaient sur les noms, les verbes, les adjectifs et les adverbes. Pendant la première semaine, les deux exercices de closure effectués comprenaient deux choix de réponse pour chaque *item* ; ensuite, on ne proposait plus de réponses pour combler les lacunes.

Au bout de neuf semaines, *l'Iowa Silent Reading Test* révèle des progrès pour les deux premiers groupes, entre lesquels il n'y a pas de différence significative. Le troisième groupe est inférieur.

Si encourageante soit-elle, cette recherche présente néanmoins une grave faiblesse expérimentale : la variable « professeur » n'a pas été contrôlée.

Pour la recherche future, E. Jongsma fait une série de recommandations que l'on peut résumer ainsi :

1. Le maître devrait discuter de chaque solution avec les élèves, pris individuellement ou non. Le travail en groupe pourrait aussi être fructueux.
2. Quatre manières de présenter l'exercice devraient être mieux étudiées :
 - a) Comblent simplement les lacunes ;
 - b) Comblent les lacunes et répondent à un questionnaire à choix multiple ;
 - c) Lire le texte complet puis faire l'exercice de closure ;
 - d) Faire l'exercice de closure — lire le texte complet — refaire l'exercice de closure.

1 R. MARTIN, *Transformational Grammar, Cloze, and Performance in College Freshmen*, Syracuse University, Thèse de doctorat inédite, 1968.

3. Les exercices de closure devraient s'étendre sur une longue durée (plusieurs mois).
4. Pour l'enseignement, il importerait d'étudier beaucoup plus systématiquement que par le passé les suppressions lexicales ou structurales.
5. En particulier, la suppression de noms et de verbes pourrait permettre de vérifier les connaissances (surtout factuelles) des étudiants dans certaines branches comme les sciences naturelles.
6. Pour l'enseignement des langues étrangères, les exercices de closure pourraient faciliter les acquisitions structurales.

3. VALIDITE DE LA TECHNIQUE

1. Aspect théorique

Le test de closure sert principalement à évaluer la lisibilité d'un texte et à mesurer le niveau de compréhension qu'un individu possède d'une langue. Comme l'exercice consiste à reconstruire un texte, la connaissance active de la langue est aussi mesurée.

Dans tous les cas, le point de départ théorique est cependant le même : l'existence de contingences statistiques dans une langue donnée. Ces contingences sont de mieux en mieux mises en lumière par la linguistique contemporaine, et elles sont de deux ordres :

- 1° Elles concernent la nature des mots et des lettres qui se succèdent. Par exemple, la probabilité qu'un article soit suivi d'un substantif ou d'un adjectif est très élevée en français, alors que la succession article-adverbe ne l'est pas. Le groupe littéral *str* est presque certainement suivi d'une voyelle.
- 2° Elles concernent les mots mêmes. Par exemple, comparer la probabilité d'apparition des adjectifs *vert* et *bleu* dans : l'herbe *verte*, l'herbe *bleue*. Il est clair que l'emploi de certains mots est à la fois conditionné, d'une part, par la nature du phénomène en question et la connaissance que nous en avons (l'herbe que nous percevons est le plus souvent verte dans notre pays) et, d'autre part, par des clichés culturels (ce type de cliché veut que la *mer* soit surtout *bleue* — la grande bleue ! — alors que nous la voyons presque toujours d'une autre couleur dans nos régions. Idem pour le *beau Danube bleu*).

Selon sa connaissance de la langue, des phénomènes décrits et, plus généralement, de son expérience vécue, un individu s'attend donc plus ou moins à ce qui va suivre dans le discours. Exploitant systématiquement les variations interindividuelles ou intra-individuelles quant à la validité de la prédiction, la technique de closure semble donc théoriquement valide.

Toutefois, l'individu hésite souvent entre plusieurs solutions adéquates — c'est-à-dire présentant des niveaux de probabilité sans différences très significatives — et se décide plus ou moins au hasard. Ainsi s'explique la fidélité relativement peu élevée de l'épreuve de closure utilisée pour mesurer la compréhension individuelle (rappelons qu'on n'admet pour correct que le seul mot utilisé par l'auteur). Ainsi s'explique aussi la valeur du test de closure comme mesure de la lisibilité : si l'on attend tel mot dans un texte et qu'on en trouve un autre, le processus de la lecture s'en trouvera ralenti, voire arrêté. Pour mesurer la lisibilité, il est

souhaitable que les sujets soient de même intelligence et de même culture, conditions pratiquement impossibles à atteindre de façon parfaite. On peut cependant s'en rapprocher en utilisant des populations très homogènes, par exemple, les étudiants d'une même année et d'une même section.

A. Mesure de la capacité en lecture

Il n'est sans doute pas inutile de discuter brièvement d'une objection parfois soulevée à propos de la technique de closure : le morcellement du texte serait incompatible avec le globalisme de la lecture courante.

Pareille objection semble reposer sur la croyance qu'on ne lirait jamais « qu'en diagonale », c'est-à-dire en recherchant dans le texte écrit un petit nombre de mots clés permettant de déduire le sens général (écrémage).

Les étudiants connaissent bien la nuance entre parcourir rapidement un cours par repérage et écrémage, pour s'en faire une première idée, et le lire de façon approfondie pour pouvoir le résumer sans négliger aucune des idées essentielles.

L'important n'est pas seulement le sommaire, mais aussi et surtout, à mesure que l'on s'élève dans l'échelle de l'intelligence et du savoir, tout ce qui s'y ajoute de finesse, de relativité.

Le bon lecteur, beaucoup plus rare qu'on ne l'imagine souvent, est celui qui possède un degré élevé de flexibilité dans la stratégie de la lecture : il modifie aisément sa façon de lire en fonction du but poursuivi¹.

Le test de closure vise à vérifier la capacité de comprendre un texte dans son intégralité.

On sait le rôle que joue dans le décodage l'empan du lecteur, c'est-à-dire sa capacité de mémoire immédiate mesurée en mots. En particulier, si la distance linguistique — ou nombre de mots qui séparent deux mots d'une phrase directement dépendants l'un de l'autre — est supérieure à l'empan du lecteur, celui-ci ne comprendra probablement pas le texte. Or, selon F. Richaudeau² — qui prend une marge de sécurité —, l'empan est de :

1 Sur les variations de comportement de lecture selon le but, voir H.K. SMITH, *The Responses of Good and Poor Readers when Asked to Read for Different Purposes*, in *Reading Research Quarterly*, 3, 1967, 53-83.

Selon McDonald, seulement 10 % des lecteurs auraient un comportement vraiment flexible. Voir A.S. Mc DONALD, *Factors Affecting Reading Test Performance*, in *Ninth Yearbook of the National Reading Conference*, Fort Worth, Texas, Texas Christian Univ. Press, 1960, 29-35.

2 F. RICHAUDEAU, *La Lisibilité*, p. 114.

8 mots pour le lecteur lent,
13 mots pour le lecteur moyen,
16 mots pour le lecteur rapide.

Le test de closure où l'on supprime un mot sur cinq présente le double avantage de se situer unitairement sous l'empan minimum alors que ses deux multiples (10 et 15) restent respectivement sous l'empan moyen et supérieur.

La technique peut, grâce à cela, être appliquée à des sujets de capacité très variée : du début de la lecture courante à l'art le plus achevé. On constate toutefois que le rôle du contexte immédiat diminue à mesure que le sujet se développe et que sa capacité de lecture augmente ; des indices de signification, parfois fort éloignés, sont alors utilisés.

Jouant sur un contexte immédiat de huit mots au maximum, quatre devant et quatre derrière, la technique de closure semble, a priori, porter plus sur la compréhension de la proposition ou de la phrase, que sur le paragraphe ou le passage entier.

Disons tout de suite que, même si sa validité — certaine — s'arrêtait là, on aurait de toute façon fait un progrès considérable dans le testage de la lecture. On sait depuis longtemps tester la compréhension de mots pris isolément ; on a maintenant gagné la phrase. Resterait l'ensemble de phrases.

Mais une recherche récente semble démontrer que la compréhension testée par la closure dépasse les limites du contexte immédiat. Et cela est capital.

Comme le rappelle S. Ramanauskas à qui l'on doit l'étude dont il va être question¹ : « Il est établi que le discours écrit (consistant en un ou plusieurs paragraphes) est plus que la somme des phrases qui le composent et possède sa propre réalité psychologique ou Gestalt. » Il est bien connu que des contraintes syntaxiques ou sémantiques s'exercent, non seulement à l'intérieur de la phrase, mais aussi sur des mots situés en dehors d'elle.

S. Ramanauskas a sélectionné deux textes comptant respectivement 302 et 304 mots et les a présentés chacun sous deux formes : la forme originale et une seconde où l'ordre des phrases avait été modifié au pur hasard. Sur chacune des formes ainsi obtenues, un test de closure a été construit.

L'hypothèse était que, tous les autres facteurs restant constants, un plus grand nombre de réponses correctes seraient fournies pour les textes présentés dans l'ordre normal des phrases que pour les autres. (On se rappellera, en particulier, qu'avec les

1 S. RAMANAUSKAS, The responsiveness of cloze readability measures to linguistic variables operating over segments of text longer than a sentence, in *Reading Research Quarterly*, VIII, 1, 1972, 72-91.

formules de lisibilité classiques, le texte original et le texte modifié auraient ici été affectés du même score de lisibilité. En effet, dans les anciennes formules, tous les calculs se font à l'intérieur de la phrase.)

Un schéma expérimental en carré latin a été appliqué pour contrôler les principales variables comme l'ordre de présentation des épreuves, le sexe, la capacité en langue maternelle et le Q.I.

Le résultat expérimental est net : la différence entre les scores de closure obtenus pour les textes « naturels » et les textes « brouillés » est extrêmement significative : $P < .0005$.

L'expérience de Ramanauskas devrait être répétée, d'autant plus que la population sur laquelle il a travaillé est spéciale : 58 sujets accusant un handicap mental léger ; ils fréquentent l'enseignement secondaire inférieur, mais ont un âge mental moyen de 10 ans.

Les rôles respectifs des indices syntaxiques et sémantiques dans la compréhension interphrase devront, par ailleurs, être mieux cernés. Or, comme le note Ramanauskas, ils sont difficiles à séparer : « ... car les changements de structure affectent probablement la signification. Brown (1970) écrit que les attentes syntaxiques sont guidées par des informations sémantiques antérieures, par exemple dans la recherche du sujet logique ou de l'attribut. Treisman (1965) conclut que la capacité d'utiliser correctement des parties du discours est considérablement affectée par la cohérence conceptuelle du passage. » (p. 83).

Si les sujets sont invités à combler les lacunes sans avoir essayé préalablement de se faire une idée globale du sens du texte, il arrive qu'au terme d'un complètement assez satisfaisant, l'ensemble ne soit néanmoins pas bien compris. Le morcellement de la pensée causé par les arrêts à chaque lacune explique une bonne partie, sinon la totalité du phénomène. Une relecture du texte complété est donc souhaitable.

Par ailleurs, tel qu'il est souvent appliqué, le test de closure combine les avantages de l'écrémage (*scanning*) et de la lecture approfondie. En consultant les consignes reproduites en annexe II, on verra que le sujet est invité à lire l'ensemble du texte avant de commencer à combler les lacunes. Vu le taux de redondance généralement très élevé, — A. Moles l'évalue à 55 % pour le français¹, — on peut ainsi se faire une idée générale du contenu et de ses grandes articulations.

1 A. MOLES, *Théorie de l'information et perception esthétique*, Paris, Flammarion, 1958, p. 50.

C. Shannon estimait d'abord la redondance moyenne de l'anglais à 75 % (c'est-à-dire qu'en supprimant 75 % des lettres d'un texte, on peut encore en deviner le sens). Cette estimation a été nuancée par la suite (A. Chapanis ; G. Miller ; E. Friedman). Un taux de 50 % de suppression est généralement considéré comme acceptable.

Il semble, en outre, évident que moins un texte est redondant, moins on peut en sauter des parties tout en continuant à le comprendre.

Moyennant quelques précautions élémentaires, la closure paraît donc indiscutablement tester aussi bien la lisibilité que la compréhension de l'ensemble considéré.

Comment s'explique la supériorité que l'on attribue souvent au test de closure par rapport au test de lecture classique, à choix multiple ou à courtes réponses ouvertes ?

Cette supériorité est de deux ordres : l'élimination de l'action perturbatrice de la question et le meilleur échantillonnage du contenu.

Considérons d'abord l'influence des questions.

Comme E. B. Coleman le rappelait récemment encore : « Un chercheur peut poser des questions faciles sur un texte difficile et vice versa. De plus, les questions posées sur un même texte par tel examinateur peuvent différer considérablement des questions posées par tel autre ¹. »

On sait, depuis longtemps, que comprendre les questions (ou les problèmes posés) constitue une première épreuve de lecture qui s'ajoute à la compréhension du texte et dont il est fort difficile, sinon impossible, d'évaluer l'incidence sur le résultat.

Les deux expériences suivantes mettent cette difficulté en lumière.

Davis et Haggard ² ont formulé la même question, avec réponse à choix multiple, en utilisant, d'une part, un vocabulaire de fréquence élevée et, d'autre part, un vocabulaire plus rare. Alors que, pour la première forme de l'*item*, la différence de réussite entre groupes socio-économiques supérieur et inférieur fut de 12 %, elle passa à 32 % pour la seconde forme.

En fait, la difficulté d'une question dépend à la fois de la langue dans laquelle elle est formulée et de la nature du problème posé ³.

Ayant calculé la lisibilité de 120 textes (formule de Gray-Leary) et la lisibilité de chaque série de questions posées sur chacun des

1 COLEMAN, E.B., *Developing a Technology of Written Instruction*, document ronéotypé, p. 3.

2 A. DAVIS, *Education for the Conservation of Human Resources*, in *Progressive Education*, 27, 1950, 221-224.

3 Pour la discussion de cet aspect, nous suivons P. VAN HAUWERMEIREN, *De Leesbaarheidsmeting*, Univ. de Louvain, 1972, thèse doctorale inédite.

textes, I. Lorge observe une corrélation très basse entre les deux et estime que les questions n'explorent qu'un tiers de la difficulté du texte¹.

S'attaquant à l'autre aspect du problème, J. Tuinman ne le résout pas, mais remarque que les questions devraient être telles qu'on ne devrait pas pouvoir y répondre, sans avoir lu le texte. Il écrit :

« Un *item* a une dépendance de lecture optimale si la probabilité de réponse correcte pour quelqu'un qui n'a pas lu le texte n'excède pas la chance attribuable au hasard². »

Le test de closure offre l'immense avantage d'explorer la compréhension sans l'intermédiaire de questions.

Comme l'a bien observé J. Bormuth³, une question dont la réponse dépend totalement du texte équivaut à supprimer une partie de ce texte et à la remplacer, par exemple, par un interrogatif neutre, la réponse consistant à restituer le ou les mots originaux. L'exemple le plus simple est :

Marie mange une pomme.

Qui mange une pomme ?

A ce point de vue, la technique de closure et la question procèdent donc du même mécanisme.

Quelles différences existent alors ?

- 1° Dans le test de closure, on ne supprime qu'un seul mot à la fois. Il peut être structural ou lexical. Une question peut porter sur plusieurs mots et ne se limite jamais à des mots structuraux.
- 2° La closure porte exclusivement sur le texte, alors que la question peut le dépasser.
- 3° Pour le test de closure, les sujets n'ont pas lu préalablement le texte intégral.

La supériorité du test de closure réside aussi, avons-nous dit, dans une meilleure validité de contenu. Rappelons que cette validité est acquise quand le test explore un échantillon suffisamment représentatif de la matière considérée.

1 I. LORGE, *Predicting Reading Difficulty of Selection for Children*, in *Elementary English Review*, 16, 1939, 229-233.

2 J. TUINMAN, *Asking Reading-Dependent Questions*, in *Journal of Reading*, 14, 1971, 289-292 et 336.

3 J. BORMUTH, Ed., *Readability in 1968*, Champaign, Ill., National Council of Teachers of English, 1968, 40-47, cité par P. Van Hauwermeiren, o.c.

Pour répondre statistiquement à pareille exigence, l'univers des questions possibles doit être défini et l'échantillon, en proportion à fixer, doit être tiré au hasard.

Il est clair que les questions de compréhension habituellement posées n'offrent guère de garanties à cet égard. Non seulement leur nombre est souvent insuffisant, mais les aspects sur lesquels elles portent sont sélectionnés en fonction de l'intuition de l'interrogateur, et non selon une règle précise. On en arrive ainsi aisément à des problèmes d'application, d'induction, de déduction, d'extrapolation, voire d'évaluation qui constituent autant d'exercices dits d'intelligence à partir du texte, mais non des contrôles de compréhension du donné, c'est-à-dire du décodage du message. En outre, on l'a maintes fois montré expérimentalement, l'intuition ou l'expérience professionnelle sont loin d'être des guides sûrs pour déceler les éléments du texte qui pourraient n'être pas compris.

Anticipant sur nos conclusions, remarquons immédiatement que rien n'empêche un enseignant ou un chercheur d'utiliser simultanément le test de closure et le contrôle de compréhension traditionnel. On observe d'ailleurs une corrélation assez élevée entre les deux. Le tout est de ne pas décider *a priori* que la seconde technique est nécessairement supérieure à la première...

L'opposition compréhension générale - compréhension spécifique est sans doute recouverte par deux autres oppositions bien dégagées par E.F. Rankin¹ :

— *Compréhension des faits* - *Compréhension des relations*.
(signification lexicale) (signification structurale)

Alors que surtout les noms et les verbes pleins sont porteurs des *faits*, du contenu objectif d'un message, les *relations* sont principalement exprimées par les articles, les prépositions, les conjonctions, les pronoms relatifs, les verbes auxiliaires, par la syntaxe et par les morphèmes.

E. Rankin écrit :

« Comme la suppression au hasard de tout n^{me} mot tente d'échantillonner tous ces indices de la signification structurale, et laisse subsister beaucoup de noms et de verbes (très abondants) qui servent de signaux pour retrouver la signification lexicale, la mutilation aléatoire d'un texte devrait plus réduire la signification structurale que la signification lexicale. C'est

¹ E.F. RANKIN, *The Cloze Procedure - Its Validity and Utility*, in R. FARR, *Measurement and Evaluation of Reading*, New York, Harcourt, Brace and World, 1970 ; voir spécialement pp. 247-250.

pourquoi un test de closure construit de la sorte mesurera principalement la compréhension structurale (c'est-à-dire la compréhension des relations entre idées). »

E. Rankin a confirmé cette hypothèse en construisant, parallèlement aux tests proposés par Taylor, des tests de closure où seuls les noms et les verbes pleins étaient supprimés.

Utilisé comme le recommande Taylor, le score de closure dépend donc plus de la connaissance générale de la langue que de la connaissance de la matière sur laquelle le texte porte. Ainsi s'explique que le score de closure s'est révélé bon prédicteur des résultats en langue maternelle dans l'enseignement secondaire.

Compréhension : produit ou processus ?

Dans le test de compréhension habituel, on mesure le *produit* de la lecture en posant des questions de contrôle au terme de celle-ci. Un test de closure peut aussi être administré après lecture.

W. Taylor, puis E. Rankin observent des corrélations d'environ .80 entre les scores de closure et les tests de mémoire immédiate.

Comment le test de closure permet-il de mesurer le *processus* de la lecture, en train de se dérouler ?

Une remarque préliminaire s'impose. Deux sujets qui obtiennent le même score de compréhension d'un texte ne possèdent pas nécessairement la même capacité en lecture. Il se peut, en effet, que l'un des deux ait eu une meilleure connaissance initiale du sujet que l'autre : celui-ci a donc dû compenser un handicap et doit donc être considéré comme meilleur « lecteur ».

Ainsi apparaît le problème du *gain d'information* auquel J. Bor-muth a consacré de pénétrantes analyses.

Le test de closure permet de comparer deux lecteurs en neutralisant la variable « information initiale ». En effet, il suffit de construire deux formes parallèles du test et d'en administrer une avant et une après lecture. Rankin encore signale toutefois que la personnalité, spécialement l'anxiété, joue un rôle susceptible de fausser la comparaison de capacité.

B. Mesure de la lisibilité

Jusqu'à présent, la discussion a surtout porté sur la validité du test de closure utilisé comme instrument de contrôle de la compréhension. Qu'en est-il pour l'évaluation de la lisibilité ?

Nous l'avons déjà signalé, deux types d'utilisation doivent être ici distingués : médiate ou immédiate.

Utilisé immédiatement, le test de closure fournit un indice de lisibilité, unanimement reconnu comme plus valide encore que pour la compréhension. En effet, l'obligation de restituer le mot originalement employé dans le texte, à l'exclusion de tout synonyme, dépasse la simple compréhension. Connaissant bien la langue, le sujet a tendance à combler la lacune d'un texte par le mot le plus fréquemment employé dans l'ensemble considéré. Autrement dit, le sujet tend à suivre la pente de la probabilité la plus élevée. Si le choix de l'auteur s'écarte, de façon plus ou moins accusée, de cette probabilité, le processus de la lecture est ralenti et, rien que de ce fait, la lisibilité baisse¹.

Cette utilisation immédiate pour la lisibilité se différencie des formules de lisibilité en ce sens que ces dernières « prédisent la lisibilité d'un matériel, destiné à certains lecteurs, sur la base de quelques comptages simples (mots, syllabes, ...) alors que la technique de closure mesure la lisibilité en établissant la moyenne des résultats obtenus par un échantillon des lecteurs potentiels »².

Rappelons, par ailleurs, que la mesure de la lisibilité par la technique de closure permet, selon toute vraisemblance, d'explorer aussi la compréhension interphrase, alors que toutes les formules de la première génération (Flesch, Lorge, Dale-Chall, etc.) se cantonnent à des mesures intraphrase.

Dès que la validité du test de closure fut établie, son utilisation médiata pour la construction de formules de lisibilité s'est répandue.

Rappelons que ces formules ont pour objectif de nous informer aussi économiquement que possible, sans encore devoir expérimenter sur des sujets, de la place qu'un texte occupe dans une échelle de difficulté de compréhension, établie pour une population définie.

Grâce aux grands progrès accomplis par la linguistique contemporaine, la recherche sur la lisibilité a trouvé, au cours de la dernière décennie, un regain d'intérêt considérable : au lieu de se contenter d'indicateurs, de variables adventices, les chercheurs contentent maintenant d'appuyer leurs formules sur les véritables variables indépendantes. Tous les travaux importants de ces dernières années procèdent d'une même méthodologie :

- 1° Un choix de textes couvrant une très large bande de contenu et de capacité en lecture est établi.
- 2° Des tests de closure sont construits sur ces textes et fournissent ainsi la mesure de critère.

1 Ce phénomène a trop souvent été démontré à l'aide du tachistoscope pour que nous nous y arrêtions encore.

2 G. KLARE, o.c., p. 8.

- 3° La relation entre un grand nombre de variables linguistiques (J. Bormuth et G. Henry dépassent aujourd'hui la centaine) et le critère est étudiée.
- 4° Par régression multiple, on détermine quelles variables, aussi peu nombreuses et aussi facilement mesurables que possible, fournissent la meilleure prédiction du critère.

II. Aspect expérimental

Plusieurs résultats expérimentaux ont déjà été indiqués au cours de la discussion théorique. Nous regroupons ici des données surtout numériques, choisies parmi une quantité considérable de comptes rendus de recherches (voir bibliographie *in fine*).

A. Mesure de la capacité en lecture

La plupart des tests de compréhension sont constitués de questions à choix multiple, plus rarement de questions ouvertes. Les faiblesses de ces procédures ont été analysées dans la section précédente.

Tirer une conclusion nette d'une corrélation entre les résultats à un test de closure et les résultats à un test de compréhension n'est donc pas possible. Jamais on ne sait avec certitude si l'on explore les mêmes aspects, de part et d'autre ; jamais on ne sait si une corrélation plus basse ou plus élevée que celles obtenues par d'autres auteurs n'est pas due à une moins bonne ou à une meilleure construction des tests de compréhension, plutôt qu'à une faiblesse de l'épreuve de closure¹.

Par conséquent, les chiffres avancés par la suite ne doivent pas être considérés isolément ni pris pour absolus. C'est l'ensemble des résultats expérimentaux qu'il faut considérer. Sous réserve de quelques restrictions dont la pertinence réelle reste à éprouver, une bonne validité paraît indiscutablement établie.

1° *Corrélations avec les tests expérimentaux*

— En 1962, J. Bormuth² rédige neuf textes de complexité linguistique et de contenu très variés. Sur chacun de ces textes, il construit un test de closure et un questionnaire de compréhension à choix multiple.

1 Par exemple, dans une étude factorielle, W. Weaver et A. Kingston concluent que les principaux facteurs saturant un test de compréhension et un test de closure ne sont pas les mêmes. J. Bormuth a pu contester ces résultats en montrant que le test de compréhension utilisé présentait des anomalies et que la population expérimentale était anormalement homogène.

Voir W. WEAVER et A. KINGSTON, A Factor Analysis of the Cloze Procedure..., in *Journal of Communication*, 13, 1963, 252-261. J. BORMUTH, Factor Validity..., o.c.

2 J. BORMUTH, Cloze Tests as a Measure of Comprehension Ability and Readability, in *Dissertation Abstracts*, 23, 1963, 4218-4219.

Ces épreuves, administrées en 4^e, 5^e et 6^e années de l'école primaire, révèlent des corrélations élevées, de .73 à .84, entre les deux types de scores.

On remarquera que, dans cette nouvelle expérience, c'est, en fait, la réaction d'un même élève à des textes de difficulté variée qui est testée. D'une part, on établit ainsi un critère de lisibilité et, d'autre part, on teste la compréhension sous un angle nouveau.

— En 1966, J. Bormuth¹ utilise quatre séries de treize passages et obtient les corrélations suivantes (rô de Spearman).

Tests-critères	Tests de closure			
	A	B	C	D
Compréhension	.91	.96	.91	.90
Reconnaissance des mots	.94	.96	.92	.90

— En 1968, L. Mosberg et al.² contestent la conclusion de Bormuth. Ils estiment que le test de closure n'explore qu'une partie de ce qui est exploré par les tests de compréhension à choix multiple. Ils constatent, en outre, que le test de closure est plus sensible au niveau scolaire (*grade level*) qu'à la difficulté du texte.

— La même année, E. Coleman³ affirme, sans le démontrer rigoureusement, que le test de closure mesure la compréhension de groupes de mots et non la relation entre phrases complexes et paragraphes.

— En réponse, J. Bormuth⁴ reprend, en 1969, son étude de 1962 et l'approfondit. Le test de compréhension auquel il oppose le test de closure comprend des groupes d'*items* portant sur le vocabulaire, le contenu factuel, la chronologie des événements, les inférences, les relations causales, les idées principales et les intentions de l'auteur.

Bormuth conclut que rien ne permet d'affirmer que les tests de closure mesurent autre chose que la compréhension au sens large.

1 J. BORMUTH, Cloze Tests Readability: Criterion Reference Scores, in *Journal of Ed. Measurement*, 1968, 189-196.

2 L. MOSBERG, T. POTTER et R. CURNELL, *The relation between cloze and multiple-choice test scores as a function of relative paragraph difficulty and grade level*, Los Angeles, Southwest Regional Laboratory for Educational Research and Development, 1968.

3 E.B. COLEMAN, Experimental Studies of Readability, in J.R. BORMUTH, Ed., *Readability 1968*, Champaign, Ill., National Conference on Research in English, 1968.

4 J. BORMUTH, Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability, in *Reading Research Quarterly*, IV, 1969, 358-367.

2° Corrélations avec des tests standardisés

— En 1967, J. Bormuth a dressé un tableau récapitulatif de huit recherches relevant de cette rubrique¹. Nous le reproduisons ci-dessous.

Auteur ²	Niveau pédagogique	Tests	Corrélations
Jenkinson (1957)	Enseignement secondaire	<i>Cooperative Reading C2</i>	
		Vocabulaire	.78
Rankin (1957)	Université	Niv. de compréhension	.73
		<i>Diagnostic Survey</i>	
		Compréhension de l'histoire entière	.29
Fletcher (1959)	Université	Vocabulaire	.68
		Compréhension d'un passage	.60
		<i>Cooperative Reading C2</i>	
		Vocabulaire	.63
		Niv. de compréhension	.55
		Vit. de compréhension	.57
Hafner (1963)	Université	<i>Dvorak-Van Wagenen</i>	
		Taux de compréhension	.59
Ruddell (1963) (6 tests de closure)	Ecole primaire	<i>Michigan Vocabulary Profile</i>	.56
		<i>Stanford Achievement Signification</i>	
Weaver et Kingston (1963, 2 tests de closure)	Université	de paragraphes	.61-.74
		<i>Davis Reading test</i>	.25-.51
Green (1964)	Université	<i>Diagnostic Reading Survey</i>	
		Compréhension tot.	.51
Friedman (1964) (20 tests de closure)	Université (étudiants étrangers)	<i>Metropolitan Achievement</i>	
		Vocabulaire	.63-.85
		Score total lecture	.71-.87

¹ J.R. BORMUTH, *Implication and Use of Cloze Procedure in the Evaluation of Instructional Programs*, Los Angeles, UCLA, Center for the Study of Evaluation of Instructional Programs, 1967.

² On trouvera la référence précise de ces huit recherches dans la bibliographie finale.

Dans une recherche faite à l'Université de Liège, M. C. Marlier¹, testant à 14-15 ans, a trouvé une corrélation de .44 avec la version française du test de compréhension 7/c de l'I.E.A. Cette corrélation, relativement basse, s'explique par la grande homogénéité de la population (une centaine de sujets).

— En 1968, J. Bormuth, encore lui², a utilisé le *Test de lecture orale* de Gray pour déterminer les corrélations entre le nombre moyen de mots prononcés correctement par des élèves des 4^e, 5^e et 6^e années de l'école primaire et les scores au test de closure correspondant. Treize passages ont été lus. J. Bormuth obtient des corrélations de .90 à .95.

B. Mesure de la lisibilité

Dans les recherches corrélationnelles de validation, on étudie la relation, soit avec les scores obtenus aux meilleurs tests de lisibilité, soit avec l'avis de groupes d'experts. Dans les deux cas, les expériences sont concluantes.

W. Taylor³ observe des corrélations qui, pour le score de Flesch, s'élèvent jusqu'à .71 et pour le score de Dale-Chall, jusqu'à .94.

On sait que les deux formules reposent sur l'observation que des mots courts utilisés dans des phrases brèves constituent, *en général*, un texte facile à lire⁴. Cette règle peut, cependant, être aisément prise en défaut. Il est, en effet, possible d'exprimer des pensées très difficiles en recourant à des mots fréquents, à des symboles, aux ellipses.

W. Taylor⁵ a sélectionné deux passages anglais qui étaient indubitablement difficiles à lire et que la formule de Flesch classait cependant comme faciles. Il s'agissait d'un texte de Gertrude Stein, composé de mots courts et fréquents et de phrases brèves, et d'un texte de James Joyce fait de petites phrases et de mots brefs, mais très rares (parfois aussi des néologismes). Le test de closure a parfaitement révélé la difficulté réelle de ces deux textes.

On verra, dans la seconde partie, que nous avons fait une expérience similaire en français et qu'elle a été tout aussi concluante.

1 M.C. MARLIER, *Etude factorielle d'épreuves de lecture au niveau 14-15 ans*, Université de Liège, 1970 (mémoire inédit).

2 J.R. BORMUTH, Cloze Test Readability: Criterion Scores, in *Journal of Educational Measurement*, 5, 1968, 189-196.

3 W. TAYLOR, 1953, op. cit.

4 On sait que R. Flesch mesure effectivement la longueur moyenne des mots, tandis que Dale-Chall comptent les mots absents de listes de fréquences élevées, les mots rares étant, en moyenne, plus longs que les autres.

5 W. TAYLOR, 1953, op. cit.

En 1966, J. Bormuth publie une étude qui marque un tournant. S'attachant résolument à la validité de *construct* des formules de lisibilité, il entreprend une vaste recherche corrélacionnelle entre une série de variables linguistiques et la compréhension de textes entiers, de phrases, de mots. Parmi les variables, on note :

- a) La profondeur des mots, mesurée par comptage du nombre de structures commencées et non terminées dans la phrase, avant l'apparition du mot ;
- b) La redondance littéraire ;
- c) La fréquence des propositions indépendantes ;
- d) Le nombre de lettres ;
- e) La nature des mots.

J. Bormuth¹ établit notamment qu'une corrélation de .93 existe entre la combinaison pondérée de quatre variables linguistiques et la compréhension d'un texte entier. Il dispose ainsi de critères solides pour valider l'épreuve de closure.

Pour la corrélation avec l'évaluation de la difficulté des textes par des juges hautement qualifiés, on cite fréquemment l'étude de S. Sukeyori². Ayant fait classer huit textes par trois juges, selon l'ordre de difficulté, il obtient une corrélation de .83 avec les tests de closure correspondants.

Ce type de validation est cependant contestable ; la corrélation est certes très élevée, mais comme elle est obtenue pour un petit nombre de textes, l'erreur standard est, elle aussi, fort grande. Chaque fois que nous avons eu recours à l'évaluation de la lisibilité par des « experts », nous avons observé des divergences importantes entre individus. Nous avons, de nouveau, essayé cette procédure dans la recherche décrite ci-après. On en lira les résultats.

En gros, disons que les jugements individuels ne paraissent valides que pour l'identification de textes spécialement difficiles ou faciles. Ici comme dans les autres évaluations, les juges s'accordent presque toujours sur les cas extrêmes et divergent pour les cas moyens. Surtout les cas extrêmes présentent donc un intérêt réel pour les recherches de validation par la voie des expertises.

Il ne faut toutefois pas sous-estimer la validité des tendances qui se dégagent d'un grand nombre d'évaluations d'un même texte.

1 J.R. BORMUTH, *Readability : A New Approach*, in *Reading Research Quarterly*, Spring 1966.

Matériau : 20 textes de 275 à 300 mots. Sujets : élèves de la 4^e à la 8^e année d'études (*grades*) répartis en cinq groupes d'après le *California Reading Test*.

2 Siba SUKEYORI, *A Study of the Measurement of Readability Application of the Cloze Procedure to the Japanese Language*, in *Japanese Journal of Psychology*, 28, 1957, 67-73, cité par E. RANKIN, *The Cloze Procedure*, o.c., p. 240.

4. EXEMPLES D'APPLICATION

Dans la seconde partie de ce travail figure le compte rendu complet d'un testing de la compréhension de textes non techniques et d'un contrôle de lisibilité.

Il nous a paru utile de compléter cette illustration de l'utilisation la plus courante par d'autres exemples d'application révélant des grandes possibilités de la technique de closure.

I. Mesure de la compréhension de textes scientifiques

Le test de closure permet-il aussi de tester l'intelligibilité et la compréhension de textes scientifiques dans lesquels figurent de nombreux symboles, comme c'est le cas en mathématique ? M. Hater et R. Kane¹ ont démontré cette possibilité.

Cinq textes d'environ 700 mots ou symboles chacun ont été empruntés à des ouvrages de mathématiques. Des tests de closure (cinq formes pour chaque texte) et de compréhension (choix multiple) ont été construits sur ces passages² et administrés à 1 717 élèves des trois années de l'enseignement secondaire inférieur. Les lacunes de closure ont été aussi ménagées pour les mots ou les symboles figurant dans les diagrammes, les tableaux et les figures.

Les textes portaient sur les domaines suivants : matrices (deux textes recouvrant des méthodes d'enseignement différentes : redécouverte ou exposé), système métrique, statistique, logique mathématique.

Les auteurs arrivent aux conclusions suivantes :

1. Pour tester la compréhension d'un passage, une seule forme de closure suffit.
2. La fidélité du test de closure est élevée : elle est supérieure à .90, avec la formule n° 20 de Kuder-Richardson. Cette fidélité est déjà acquise pour une classe de 30 élèves.
3. Le test de closure permet d'ordonner les textes selon leur degré de difficulté.

II. Mesure de la compréhension de poèmes

R. Kamman³ a utilisé la technique de closure pour évaluer le degré de nouveauté et de complexité que représentaient quinze

1 M. HATER et R. KANE, *The Cloze Procedure as a Measure of the Reading Comprehensibility and Difficulty of Mathematical English*, Purdue Univ., texte ronéotypé.

2 Dans cette expérience, les lacunes étaient soulignées en rouge afin d'éviter la confusion avec les barres de fraction qui, elles, étaient en noir.

3 R. KAMMAN, Verbal Complexity and Preferences in Poetry, In *Journal of Verbal Learning and Verbal Behavior*, 5, 1966, 536-540.

poèmes de 100 à 175 mots pour 24 étudiants universitaires. Le but de la recherche était de déterminer, notamment, à quel niveau de difficulté les lecteurs situaient leur choix quand ils sélectionnaient un poème pour leur seul plaisir ou pour en discuter avec un groupe.

Dans ce cas, tous les 4^{es} mots, à partir du deuxième, ont été supprimés et les sujets étaient invités à chercher les solutions qui respectent à la fois le sens et le style du poème.

Les scores moyens obtenus vont de 13,3 % à 65,8 %.

III. Mesure de la compétence en langue étrangère

On doit à J.B. Carroll la première recherche importante dans cette direction.

Une série de textes allant de très facile (6^e primaire) à très difficile (extrait des *Prolegomènes* de Kant) dont il existait des versions anglaises, allemandes et françaises, ont été utilisés. La difficulté a été évaluée avec l'ancienne formule de R. Flesch (1943). Les textes de Kant, trop difficiles, ont finalement dû être abandonnés.

Comme il s'agissait d'une langue étrangère, J.B. Carroll a facilité le test en ne supprimant qu'un mot sur dix. Chaque texte comptait 200 mots ; le score maximum était donc 20.

Voici les résultats globaux, chaque fois pour 18 textes.

Langue maternelle du sujet	Texte anglais		Texte français		Texte allemand	
	Marge de variation	Médian	Marge de variation	Médian	Marge de variation	Médian
Anglais	8,9 — 16 N = 22	11,9	7,3 — 17,2 N = 12	12,2	5,6 — 15,2 N = 10	8,5
Français	6 — 15 N = 4	11,7	9,5 — 17 N = 4	11,7		
Allemand	6,9 — 12,9 N = 12	9,1			8 — 16 N = 12	11,7

Pour des sujets travaillant dans leur langue maternelle, les scores sont proches pour les trois langues.

Un des aspects intéressants de cette recherche était de voir si l'ordre de difficulté des 18 textes serait ressenti de la même façon par des sujets lisant ces textes dans leur langue maternelle

ou dans une langue étrangère. En général, l'ordre de difficulté reste le même :

Difficulté des textes anglais.

Corrélation entre anglophones et germanophones

$$\rho = .63$$

Corrélation entre anglophones et francophones $\rho = .75$

Difficulté des textes français.

Corrélation entre anglophones et francophones $\rho = .74$

Difficulté des textes allemands.

Corrélation entre anglophones et germanophones $\rho = .81$

On constate, en outre, une corrélation élevée entre la capacité d'un sujet à compléter des textes dans sa langue maternelle et dans une langue étrangère qu'il connaît très bien. J.B. Carroll y voit le signe qu'il existe une capacité de compléter des textes, indépendante de la compétence linguistique telle qu'elle est habituellement définie. Pour mesurer la compétence dans une langue étrangère par la technique de closure, il faudrait donc d'abord évaluer la capacité de compléter en langue maternelle et ajuster en fonction le score en langue étrangère. Pratiquement, on ne voit toutefois pas comment la chose serait possible étant donné la trop grande colinéarité entre les deux phénomènes.

IV. Mesure de la difficulté des différentes classes de mots

Miller et Coleman ont sélectionné 36 passages de 150 mots, soit au total 5 400 mots, échelonnés du niveau de la première année primaire aux textes techniques les plus difficiles.

Les scores suivants, extraits d'une liste portant 44 classes de mots (liste de Fries détaillée), fournissent pour l'anglais des indications intéressantes¹.

Classes	Nombre de mots de cette classe dans l'échantillon	Score de closure
Noms communs	566	49 %
Verbes copules	117	76 %
Verbes pleins	539	44 %
Adjectifs	361	26 %
Adverbes	68	28 %
Articles	568	74 %
Conjonctions de coordination	214	65 %
Conjonctions de subordination	175	55 %
Prépositions	595	67 %

La proportion dans laquelle les différentes classes de mots apparaissent dans un texte affecte l'intelligibilité de ce texte ou, plus exactement, constitue un symptôme précieux de phénomènes plus profonds qui la déterminent.

¹ Cf. E.B. COLEMAN, *Developing a Technology of Written Instruction ; some Determiners of the Complexity of Prose*, El Paso, Univ. of Texas, s.d., p. 6.

DEUXIEME PARTIE

EXPÉRIENCE

1. CHOIX DES TEXTES

L'expérience a porté sur six textes, choisis en deux temps.

On a d'abord sélectionné trois textes distants l'un de l'autre d'environ 15 points de difficulté sur l'échelle de lisibilité de R. Flesch, soit :

50 — 35 — 20

D'après nos recherches antérieures, les niveaux pédagogiques correspondant, en Belgique francophone, sont, *grosso modo* :

50 : degré supérieur de l'enseignement primaire

35 : enseignement secondaire inférieur

20 : enseignement secondaire supérieur

Il ne faut toutefois pas perdre de vue que ces indices sont des moyennes ou des médianes, les marges de variation des scores caractéristiques des textes composant, par exemple, un manuel scolaire de lecture étant considérables et les chevauchements entre niveaux pédagogiques aussi. Ainsi, les scores de lisibilité Flesch d'un manuel de français, destiné aux 5^e et 6^e années de l'école primaire¹ se distribuent normalement entre 15 et 85.

Un écart de quinze points entre textes représente donc une faible distance et devait, par conséquent, conduire à une pauvre discrimination entre des populations scolaires hétérogènes, telles qu'on les trouve habituellement dans nos classes de l'enseignement secondaire général et technique. Il ne faut toutefois pas perdre de vue que, contrairement aux tests de lisibilité classiques, le test de closure est une mesure directe de l'interaction lecteur-texte ; la discrimination entre individus ou niveaux scolaires passe ainsi au premier plan.

Une seconde série de trois textes présentant respectivement des scores approximativement égaux aux premiers a été choisie.

On sait que le score Flesch dépend directement de la longueur moyenne des phrases et des mots (nombre de syllabes pour cent mots). Les paires de textes constituées présentent un bon parallélisme au point de vue de la longueur moyenne des mots, élément

¹ L. JEUNEHOMME et G. COLLETTE, *Mon livre de français*, 5^e et 6^e, Liège, Desoer, 1950.

essentiel puisqu'il est aussi censé constituer un indice du niveau d'abstraction du texte. En gros, les trois échelons syllabiques sont :

160 — 180 — 195 syllabes/100 mots

Les textes ne sont pas contrastés au point de vue de la longueur des phrases. Elles sont partout assez longues. A chacun des trois niveaux, on trouve un texte d'une vingtaine de mots par phrase (21-21-20).

Il n'est pas exceptionnel que si l'on a affaire à des textes relativement courts (quelques centaines de mots), la difficulté varie d'une tranche de cent mots à l'autre. Nous n'avons retenu que des textes présentant une homogénéité à peu près parfaite, les six textes se situant tous entre 250 et 300 mots. L'homogénéité a été testée chaque fois sur deux échantillons de cent mots arrondis à la fin de la phrase où se situait le centième mot.

CARACTERISTIQUES DES TEXTES SELECTIONNES

Titre	N. mots	N. phrases	N. mots N. phrases	N. syllabes	N. syllabes 100 mots	Score Flesch		Score arrondi
						1 ^{er} échan- tillon	2 ^e échan- tillon	
Le renne Les primitifs	244	12	20	385	157	51	53	50
	258	12	21	420	160	47	47	50
L'homme et la nature Le cinéma	301	19	16	541	180	35	39	35*
	311	15	21	552	176	38	37	35
L'art contemporain Le pétrole	261	11	24	511	195	20	17	20
	301	15	20	585	194	17	22	20

* Pour l'élégance de la démonstration, nous n'avons pas hésité à arrondir, l'expérience montrant que l'échelle de scores Flesch n'a pas une sensibilité suffisante pour que les trois ou quatre points de différence dans un sens ou dans l'autre exercent une influence significative.

Les textes choisis figurent en annexè.

Une lecture attentive de ces textes fait apparaître un problème d'un grand intérêt. On sait que la méthode d'évaluation proposée par R. Flesch est relativement grossière : sa validité est estimée entre .50 et .70.

On ne songerait pas à évaluer la difficulté de compréhension d'un poème par pareille méthode. Dans bien des cas, elle fournirait des scores tout à fait trompeurs : on peut être très obscur en n'utilisant que des mots courts et des phrases brèves. La technique de Flesch souffre donc de limites graves, d'autant plus qu'elles ne sont pas objectivement prévisibles.

Notre hypothèse était que la technique de closure permettrait d'éviter cet écueil. Le bon sens ne nous dit-il pas que, dans le cas d'un poème difficile, par exemple, nous serons souvent incapable de deviner les mots supprimés, ce qui modifiera le score dans le sens de la difficulté.

Pour vérifier cette hypothèse, un des six textes que nous avons choisis présente un score de difficulté qui ne semble pas correspondre à la réalité. Le texte *Le pétrole* est affecté d'un score de difficulté pratiquement égal à celui du texte *L'art contemporain*, alors que le premier paraît, à la lecture, nettement plus facile que le second.

Or, si l'on se reporte au tableau précédent, on constate que le nombre de syllabes pour 100 mots est le même dans les deux textes (195 et 194) et que la longueur moyenne des phrases varie peu : 24 mots pour *L'art contemporain* et 20 mots pour *Le pétrole*.

L'épreuve de closure révélera-t-elle l'anomalie ?

Enfin, nous l'avons vu au chapitre précédent, la mesure de lisibilité faite par le test de closure devrait aussi trouver une validité supérieure à celle du test de Flesch dans le fait qu'elle n'est pas enserrée dans des phrases isolées les unes des autres, mais porte à peu près certainement sur une appréhension plus globale du texte.

On ne sait toutefois pas encore dans quelle mesure la présentation matérielle du texte, criblé de lacunes, constitue un obstacle au déroulement normal de la lecture. En outre, l'effet de cette présentation diffère probablement selon les individus.

Afin de disposer de points de repère, on a demandé à des juges expérimentés (instituteurs, licenciés en philologie romane, licenciés en sciences de l'éducation, méthodologistes, inspecteurs) d'indiquer à quel niveau scolaire les six textes adoptés conviendraient. Le score de difficulté ne leur avait naturellement pas été révélé. L'ordre de présentation des textes a été chaque fois laissé au hasard. Vingt experts ont été ainsi interrogés. Voici le résultat de leur évaluation.

Indiquez d'une croix le niveau scolaire
à partir duquel chacun des six textes ci-joints conviendrait.

Textes	Primaire						Secondaire (Humanités)					
	2° A	3° A	4° A	5° A	6° A		1° A	2° A	3° A	4° A	5° A	6° A
Le renne			5	8	7							
Les primitifs		1	1	4	5		7				2	
L'homme et la nature				2	3		1	9	4	1		
Le cinéma							6	4	8	1	1	
L'art contemporain									2		6	12
Le pétrole				1	8		5		5	1		

Comme on l'a si souvent observé, les avis sont divergents, ce qui confirme une fois de plus l'utilité d'une méthode d'évaluation objective.

Néanmoins, des tendances nettes se dessinent :

- 1° Tous sont d'accord pour considérer le texte *Le renne* comme convenant à l'école primaire.
- 2° A l'opposé, *L'art contemporain* est généralement reconnu comme le texte le plus difficile, accessible seulement aux élèves qui terminent l'enseignement secondaire supérieur.
- 3° Si l'on admet qu'il est impossible de classer, au jugé, un texte à un an près, la répartition des jugements entre la troisième année primaire et la deuxième année du secondaire indique que tous les textes, sauf *L'art contemporain*, sont considérés, par la majorité des juges, comme accessibles aux élèves de tout l'enseignement secondaire inférieur.

	N. de juges situant les textes		Score Flesch
	entre la 3 ^e primaire et la 2 ^e année du secondaire	à partir de la 3 ^e année du secondaire	
Le renne	20	—	50
Les primitifs	18	2	50
L'homme et la nature	15	5	35
Le cinéma	10	10	35
Le pétrole	14	6	20
L'art contemporain	—	20	20

Les estimations viennent à la fois confirmer et nuancer l'hypothèse déjà formulée à partir des scores : pour les cinq premiers textes, la discrimination sera faible, quelle que soit l'année d'enseignement secondaire. Le sixième texte, par contre, devrait se différencier nettement des autres.

Une expérience conduite à l'école primaire, en 4^e, 5^e et 6^e années, devrait normalement faire apparaître une discrimination plus fine.

Y a-t-il concordance entre le classement établi par les juges et les scores de lisibilité de R. Flesch ?

Pour le savoir, nous avons procédé de la façon suivante : les années scolaires, de la première primaire à la dernière du secondaire, sont respectivement affectées d'un poids allant de 1 à 12. Quand un juge estime qu'un texte peut être étudié dans une année donnée, on attribue donc au texte un nombre de points correspondant à l'année. Finalement, on calcule la moyenne des points.

On obtient :

	Moyenne arrondie	Score Flesch
Le renne	5	50
Les primitifs	6	50
L'homme et la nature	7,5	35
Le cinéma	8	35
L'art contemporain	11	20
Le pétrole	7	20

Pour cinq textes sur six, l'ensemble des juges établit un classement fort proche des scores de Flesch. Le texte *Le pétrole* fait nettement exception.

D'ailleurs, si l'on se base simplement sur les médianes des distributions des jugements, on arrive déjà à une constatation similaire :

Le renne	5 ^e année primaire
Les primitifs	6 ^e année primaire
L'homme et la nature	2 ^e année secondaire
Le cinéma	3 ^e année secondaire
L'art contemporain	6 ^e année secondaire
Le pétrole	1 ^{re} année secondaire

Alors que les six textes ont été utilisés dans l'enseignement secondaire, nous nous sommes limité à trois au niveau primaire et cela pour trois raisons :

- 1° Le sondage dans le primaire a, rappelons-le, été décidé alors que la recherche dans le secondaire était déjà terminée et nous ne voulions pas trop retarder la publication des résultats.
- 2° Nous savions déjà à ce moment que le texte *L'art contemporain* était hors de portée d'élèves d'école primaire.
- 3° Pour l'équilibre du schéma, nous avons gardé un texte de chaque niveau Flesch, — 50, 35, 20, — soit *Le renne*, *L'homme et la nature*, *Le pétrole*.

2. SCHEMA EXPERIMENTAL

I. Enseignement secondaire

Rappel des données du problème.

1. On souhaite étudier les réponses d'élèves appartenant aux six années de l'enseignement secondaire, général et technique.
2. Trois paires de textes ont été choisies : A, B ; C, D ; E, F. Les deux textes de chaque paire sont, selon le score de R. Flesch, du même niveau de lisibilité. Les trois paires sont, en gros, séparées de 15 points Flesch : 20, 35, 50.
3. Pour chaque texte, on a créé deux formes parallèles de test de closure (A₁, A₂ ; B₁, B₂ ; C₁, C₂ ; etc.), en supprimant un mot sur cinq. Dans la forme 1, les lacunes portent sur les mots 1, 6, 11, 16, etc. ; dans la forme 2, elles affectent les mots 3, 8, 13, 18, etc. On dispose donc, en tout, de douze tests.

Expérience.

Tant pour réduire au maximum le temps de testage que pour éviter l'effet de satiété, chaque élève n'a été soumis qu'à trois tests. On ne pouvait descendre en dessous de ce nombre si l'on souhaitait, comme c'était le cas, que chacun s'essaie à des textes de trois niveaux de difficulté différents : 20, 35, 50.

L'ordre de présentation des textes devait évidemment varier afin que l'effet de fatigue porte également sur chacun d'eux.

En se limitant à trois tests par élève, il n'était pas possible que chacun travaille sur des nombres égaux de formes de closure. Ce serait nécessairement deux premières formes et une seconde, ou inversement. La répartition devait donc se faire au hasard.

Pour satisfaire à ces diverses exigences, on a eu recours à un système de distribution relativement complexe.

Vingt-quatre carnets de trois tests ont été constitués par des combinaisons à partir des groupes suivants :

Groupe I	Groupe II	Groupe III	Groupe IV																								
<table border="1"><tr><td>A₁</td><td>B</td></tr><tr><td>C₁</td><td>D</td></tr><tr><td>E₁</td><td>F</td></tr></table>	A ₁	B	C ₁	D	E ₁	F	<table border="1"><tr><td>A</td><td>B₁</td></tr><tr><td>C</td><td>D₁</td></tr><tr><td>E₂</td><td>F</td></tr></table>	A	B ₁	C	D ₁	E ₂	F	<table border="1"><tr><td>A₁</td><td>B</td></tr><tr><td>C</td><td>D₂</td></tr><tr><td>E</td><td>F₂</td></tr></table>	A ₁	B	C	D ₂	E	F ₂	<table border="1"><tr><td>A</td><td>B₂</td></tr><tr><td>C₂</td><td>D</td></tr><tr><td>E</td><td>F₁</td></tr></table>	A	B ₂	C ₂	D	E	F ₁
A ₁	B																										
C ₁	D																										
E ₁	F																										
A	B ₁																										
C	D ₁																										
E ₂	F																										
A ₁	B																										
C	D ₂																										
E	F ₂																										
A	B ₂																										
C ₂	D																										
E	F ₁																										

Chaque groupe conduit à six combinaisons, soit au total 24 combinaisons différentes.

Cahier	Groupe I	Cahier	Groupe II	Cahier	Groupe III	Cahier	Groupe IV
(1)	E ₁ C ₁ A ₂	(7)	D ₁ B ₁ E ₂	(13)	F ₂ D ₂ A ₁	(19)	B ₂ F ₁ C ₂
(2)	E ₁ A ₂ C ₁	(8)	D ₁ E ₂ B ₁	(14)	F ₂ A ₁ D ₂	(20)	B ₂ C ₂ F ₁
(3)	C ₁ A ₂ E ₁	(9)	B ₁ E ₂ D ₁	(15)	D ₂ A ₁ F ₂	(21)	F ₁ C ₂ B ₂
(4)	C ₁ E ₁ A ₂	(10)	B ₁ D ₁ E ₂	(16)	D ₂ F ₂ A ₁	(22)	F ₁ B ₂ C ₂
(5)	A ₂ C ₁ E ₁	(11)	E ₂ B ₁ D ₁	(17)	A ₁ D ₂ F ₂	(23)	C ₂ F ₁ B ₂
(6)	A ₂ E ₁ C ₁	(12)	E ₂ D ₁ B ₁	(18)	A ₁ F ₂ D ₂	(24)	C ₂ B ₂ F ₁

Les trois textes constituant une des vingt-quatre combinaisons ont été agrafés en cahiers. Pour la distribution dans une classe, on constituait un paquet en mélangeant les cahiers dans l'ordre horizontal suivant :

1, 7, 13, 19 ; 2, 8, 14, 20 ; 3, etc.

Une classe de 28 élèves recevait donc la série complète des 24 cahiers différents, puis les cahiers 1, 7, 13 et 19.

II. Enseignement primaire

Pour l'enseignement primaire, les mêmes principes de rotation et de distribution des épreuves ont été suivis. Le schéma a, toutefois, été simplifié puisque l'on ne travaillait plus qu'avec trois textes au lieu de six.

Douze carnets de trois tests ont donc été constitués par des combinaisons à partir des groupes suivants :

Groupe I		Groupe II	
A ₁	A ₂	A ₁	A ₂
C ₁	C ₂	C ₁	C ₂
E ₁	E ₂	E ₁	E ₂

Chaque groupe conduit à six combinaisons, soit au total douze combinaisons différentes.

Cahier	Groupe I	Cahier	Groupe I
(1)	A ₂ C ₁ E ₂	(7)	A ₁ C ₂ E ₁
(2)	A ₂ E ₂ C ₁	(8)	A ₁ E ₁ C ₂
(3)	C ₁ A ₂ E ₂	(9)	C ₂ A ₁ E ₁
(4)	C ₁ E ₂ A ₂	(10)	C ₂ E ₁ A ₁
(5)	E ₂ A ₂ C ₁	(11)	E ₁ A ₁ C ₂
(6)	E ₂ C ₁ A ₂	(12)	E ₁ C ₂ A ₁

Pour la distribution dans une classe, on constituait un paquet en mélangeant les douze cahiers dans l'ordre horizontal (1, 7 ; 2, 8 ; 3, etc.).

3. POPULATION EXPERIMENTALE

L'expérience a porté sur les trois dernières années de l'école primaire (43 classes) et sur les six années de l'enseignement secondaire, général et technique (31 écoles). En gros, une population allant de 9 à 18 ans a donc été examinée.

Certaines écoles primaires de la région liégeoise comptent beaucoup d'enfants d'immigrants. Des familles se sont installées depuis peu dans le pays. On pouvait donc craindre qu'une connaissance trop rudimentaire du français, chez les élèves étrangers, ne fausse gravement les résultats. Nous avons qualifié d'étrangers les enfants dont les parents n'avaient pas encore acquis la nationalité belge.

Nombre d'élèves testés dans l'enseignement primaire

	Belges	Etrangers	Total
4 ^e année	246	33	279
5 ^e année	272	37	309
6 ^e année	230	29	259
	748	99	847

En tout, 847 élèves de l'école primaire ont donc subi chacun trois tests de closure, soit, au total, 2 541 tests.

On constate que, pour chaque niveau, le nombre d'étrangers est peu élevé ; les résultats qui les concernent devront donc être interprétés avec beaucoup de réserves et de prudence.

Nombre d'élèves testés dans l'enseignement secondaire

	Année	Général	Technique	Total
Cycle inférieur	1	360	267	627
	2	292	206	498
	3	208	241	449
		860	714	1 574
Cycle supérieur	4	308	392	700
	5	189	245	434
	6	160	193	353
		657	830	1 487

En tout, 3 061 élèves (1 517 de l'enseignement général et 1 544 de l'enseignement technique) ont donc subi chacun trois tests de closure, soit, au total, 9 183 tests.

4. ADMINISTRATION DES TESTS

Les tests ont été administrés par les instituteurs titulaires et par des professeurs de langue maternelle volontaires, dans des classes où ils enseignaient habituellement.

Les maîtres ont été préalablement mis au courant de la technique et des objectifs poursuivis par la recherche.

Les élèves ont travaillé en temps libre.

Pour l'enseignement secondaire, une première série de tests a été passée pendant l'année scolaire 1969-1970. Le nombre d'élèves examinés étant insuffisant pour l'enseignement secondaire général supérieur et pour tout l'enseignement technique, on a administré une seconde série de tests en 1970-1971.

A l'école primaire, les tests ont été passés au printemps 1972 ; jamais plus d'un test n'a été subi au cours d'une même journée.

Les consignes remises aux instituteurs précisaient :

« Chaque élève de votre classe doit compléter trois textes, — un par jour, si possible au début de la matinée. Des petits cahiers de trois textes sont constitués et doivent être remplis dans l'ordre où ils se trouvent. Cet ordre varie selon le hasard de la distribution.

Veillez à ce que les indications d'identité soient bien portées.

Relevez les cahiers chaque jour immédiatement après l'exercice et rendez à chacun le sien lors des deux séances suivantes. »

5. CORRECTION DES EPREUVES ET TRAITEMENT DES DONNEES

Les travaux ont été corrigés manuellement. Un contrôle de ces corrections a permis d'éliminer les erreurs systématiques et d'évaluer les erreurs accidentelles.

Le pourcentage d'erreurs accidentelles de correction s'élève à environ 1 %, le sigma du pourcentage étant d'environ 0,1 %. Par la suite, les perforations sur cartes IBM ont également fait l'objet d'un contrôle systématique ; les erreurs sont aussi de l'ordre de 1 %.

Les résultats ont été perforés sur cartes IBM selon le code suivant, pour l'enseignement secondaire :

- Ecoles : numérotées de 1 à 31 ;
- Niveaux scolaires : de 1 à 6 ;
- Types d'études :
 - latines : L
 - modernes : M
 - techniques : T

Si une seule classe de même type et de même niveau était testée dans une école, la lettre L, M ou T était suivie de A ; si une deuxième classe de ce type et de ce niveau était testée, on a employé la lettre B, ...

- Enfin, dans chaque classe, les élèves ont été numérotés de 1 à la fin. Exemple : 03/5/MB/07.

Pour l'enseignement primaire, on a simplement fait la distinction entre Belges et étrangers.

Certains instituteurs ont procédé à une correction collective, en collaboration avec les élèves. Un certain nombre de fraudes ont pu ainsi être commises ; les plus importantes ont été aisément décelées lors du contrôle préliminaire au codage. Tout travail suspect a été éliminé.

Le traitement statistique des données a été fait par ordinateur IBM 360/65. Le programme, écrit en FORTRAN IV, a fourni les résultats suivants :

1. Les résultats bruts et les scores (pourcentage de réussite) pour chaque élève.
2. Pour chaque année d'études, pour l'enseignement général, l'enseignement technique, puis les deux réunis :
 - a) La moyenne, la variance, l'écart type ainsi que des étalonnages normalisés en cinq et en neuf classes, pour les résultats bruts ;
 - b) Idem pour les pourcentages.
3. En outre :
 - a) Le pourcentage de réussite par *item* ;
 - b) Le coefficient de corrélation r_{bis} ;
 - c) Deux estimations de la fidélité (formule 20 de Kuder-Richardson et *split-half method*).

Toutes les analyses statistiques inférentielles ont aussi été faites automatiquement.

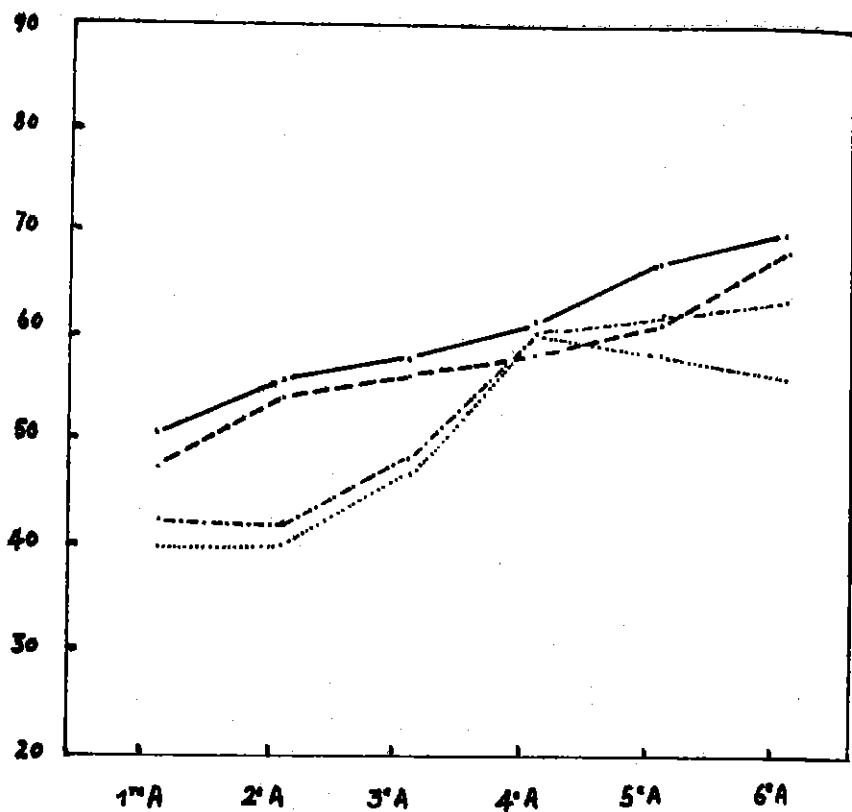
6. RESULTATS - ENSEIGNEMENT SECONDAIRE

Afin de ne pas alourdir la présentation, le détail des résultats a été omis. Nous ne reprenons ici que les tableaux récapitulatifs servant directement de base à l'analyse.

I. Enseignements général et technique séparés, formes 1 et 2 séparées

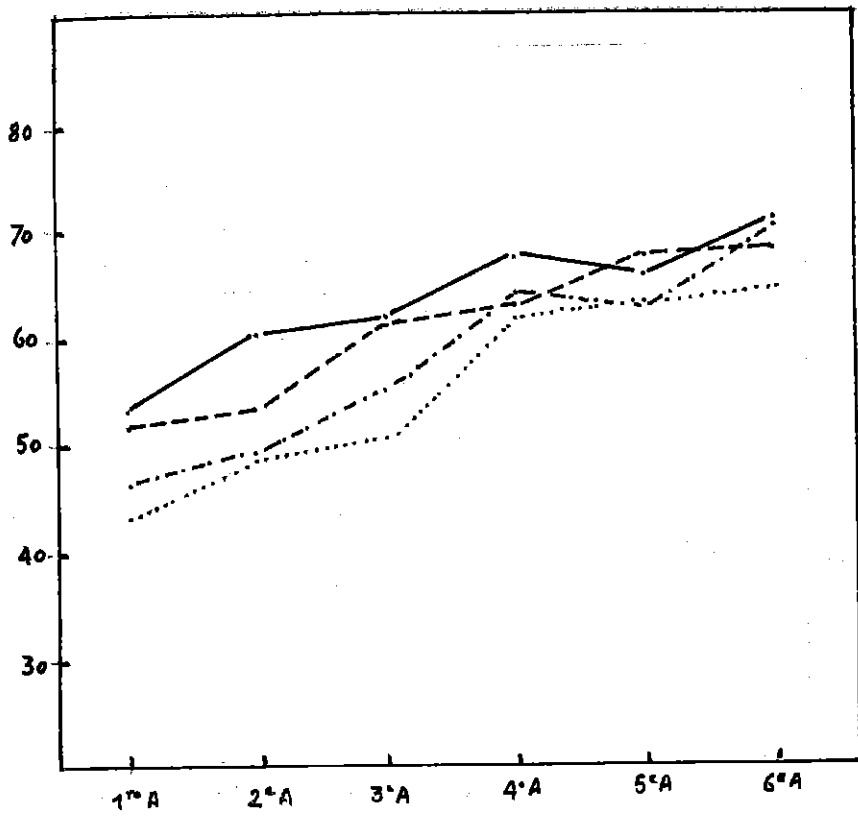
A. Scores moyens et graphiques

	1 ^{er} A		2 ^e A		3 ^e A		4 ^e A		5 ^e A		6 ^e A		
	Général	Technique	Général	Technique	Général	Technique	Général	Technique	Général	Technique	Général	Technique	
Renne	Forme 1	51,13	43,13	56,05	42,73	58,03	48,43	61,85	61,43	67,73	62,51	70,39	64,12
	Forme 2	48,02	40,19	54,65	40,96	56,67	47,87	58,80	61,40	62,04	58,83	68,81	56,87
Primitifs	Forme 1	53,02	46,10	59,99	49,37	61,70	55,19	67,41	63,81	65,60	63,28	70,85	70,70
	Forme 2	51,44	43,21	53,58	48,73	61,23	50,46	63,05	61,87	67,34	63,17	68,15	64,50
Cinéma	Forme 1	46,76	36,48	57,34	40,37	60,12	46,37	63,70	60,10	66,81	62,89	72,38	63,72
	Forme 2	54,57	49,99	60,78	46,76	63,63	53,57	69,67	68,14	69,11	68,37	76,23	69,01
Homme et nature	Forme 1	51,44	39,11	57,61	44,50	60,39	45,59	66,28	63,88	70,79	68,10	81,44	68,49
	Forme 2	48,03	38,54	57,37	39,96	57,99	48,79	65,38	65,31	69,07	65,60	67,97	72,23
Art contemporain	Forme 1	31,71	25,87	40,93	29,01	42,16	31,51	57,11	58,70	51,92	51,08	60,01	54,15
	Forme 2	27,43	23,65	38,25	25,74	34,40	26,10	49,86	48,19	51,03	47,42	56,44	48,70
Pétrole	Forme 1	50,15	40,54	56,76	48,09	59,12	48,07	65,21	68,77	68,19	66,22	72,77	69,07
	Forme 2	46,29	38,98	54,38	44,75	59,04	49,76	64,50	64,02	65,96	65,45	75,19	67,79



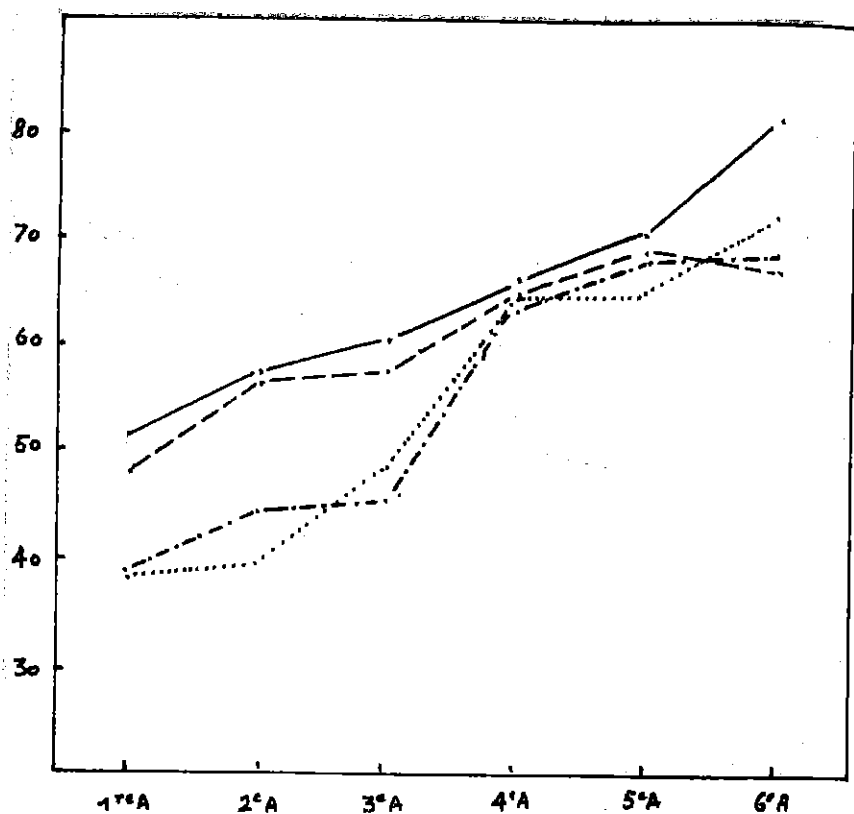
LE RENNE.

- Forme 1-6: Enseignement secondaire général.
- - - Forme 3-8: Enseignement secondaire général.
- · - · Forme 1-6: Enseignement secondaire technique.
- · · · Forme 3-8: Enseignement secondaire technique.



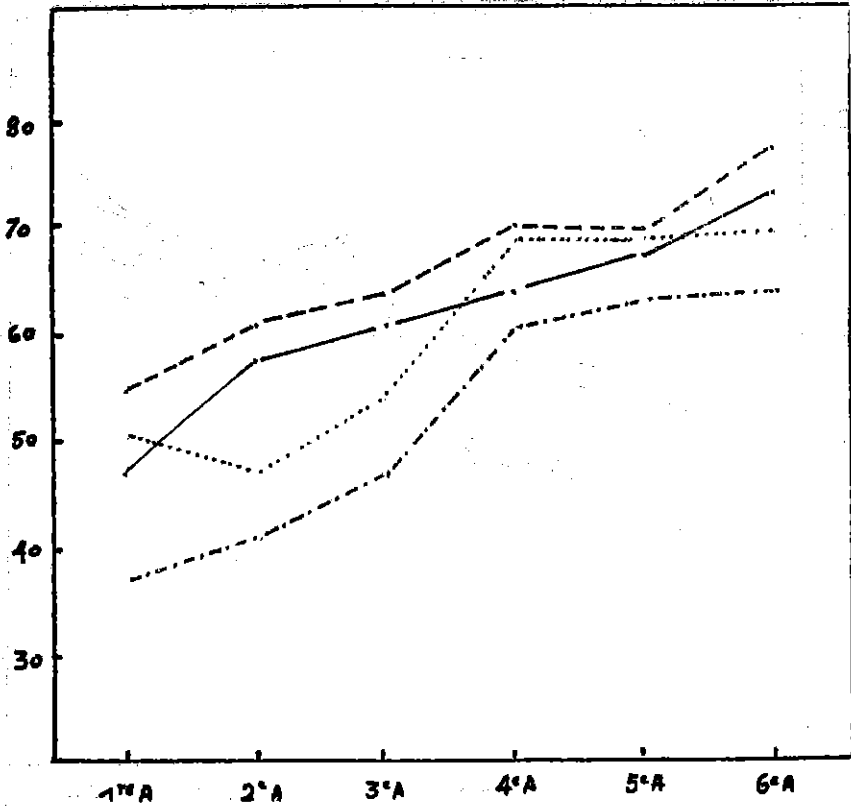
LES PRIMITIFS.

- Forme 1-6 : Enseignement secondaire général.
- - - - - Forme 3-8 : Enseignement secondaire général.
- . - . - Forme 1-6 : Enseignement secondaire technique.
- Forme 3-8 : Enseignement secondaire technique.



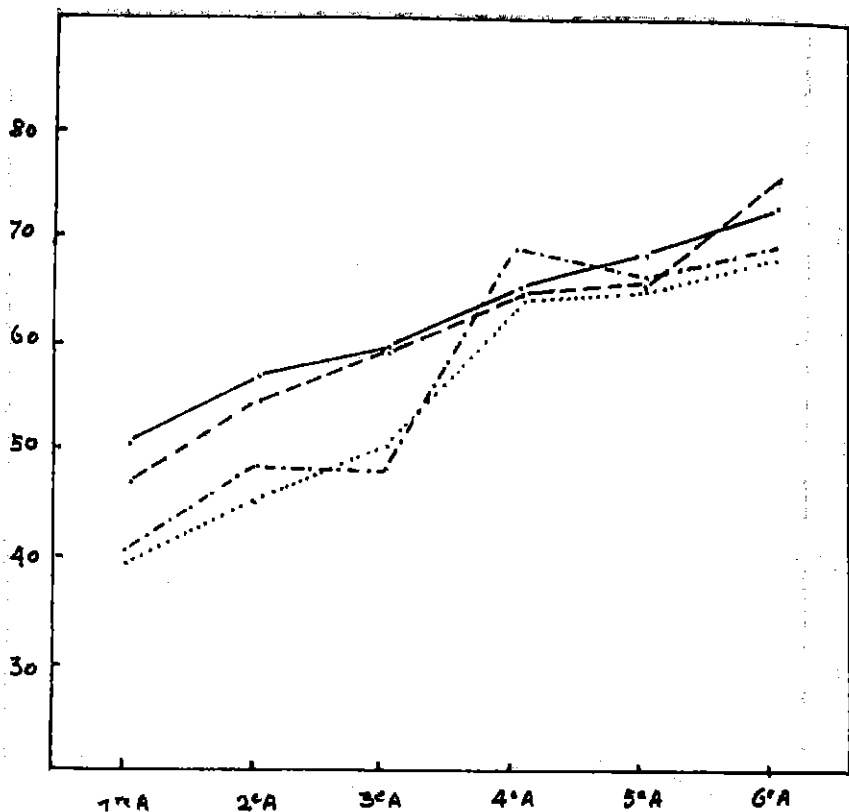
L'HOMME ET LA NATURE.

- Forme 1-6 : Enseignement secondaire général.
- Forme 3-8 : Enseignement secondaire général.
- . - . - . - . - . Forme 1-6 : Enseignement secondaire technique.
- Forme 3-8 : Enseignement secondaire technique.



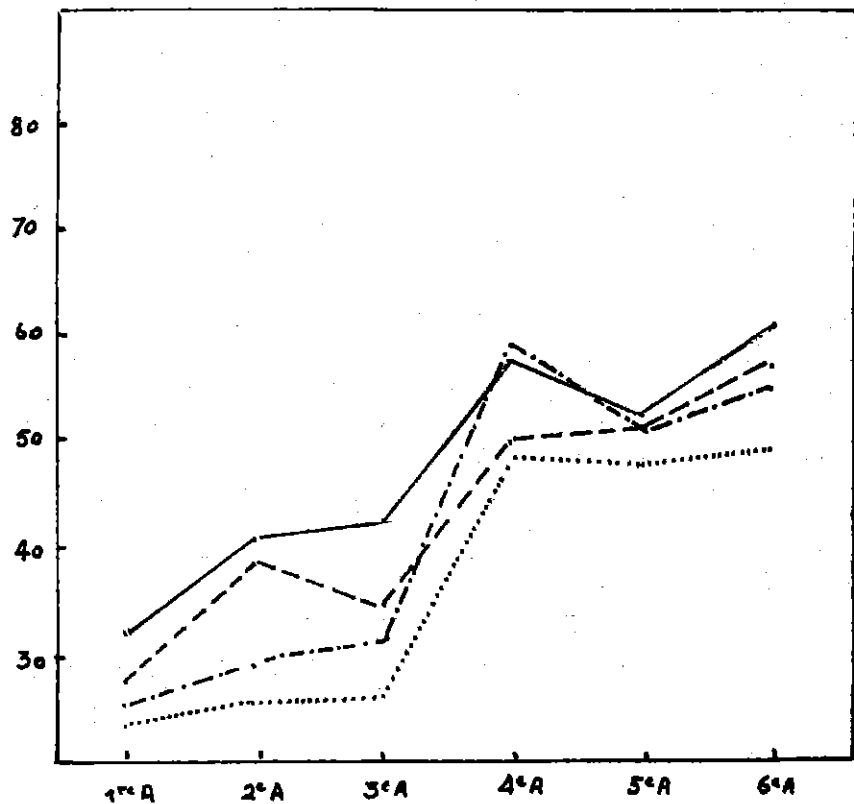
LE CINEMA.

- Forme 1-6 : Enseignement secondaire général.
- - - - - Forme 3-8 : Enseignement secondaire général.
- . - . - . - Forme 1-6 : Enseignement secondaire technique.
- Forme 3-8 : Enseignement secondaire technique.



LE PETROLE.

- Forme 1-6 : Enseignement secondaire général.
- - - - - Forme 3-8 : Enseignement secondaire général.
- . - . - . Forme 1-6 : Enseignement secondaire technique.
- Forme 3-8 : Enseignement secondaire technique.



L'ART CONTEMPORAIN.

- Forme 1-6 : Enseignement secondaire général.
- - - - - Forme 1-6 : Enseignement secondaire général.
- . - . - . Forme 1-6 : Enseignement secondaire technique.
- Forme 3-8 : Enseignement secondaire technique.

B. Constatations générales

L'examen des scores moyens et surtout de leur représentation graphique révèle d'abord un fait frappant : le *pouvoir de discrimination de la technique de closure entre années scolaires*. On observe certes, de-ci de-là, quelques irrégularités, mais, en général, on trouve l'esquisse d'une ligne droite ascendante correspondant à l'avancement de la scolarité.

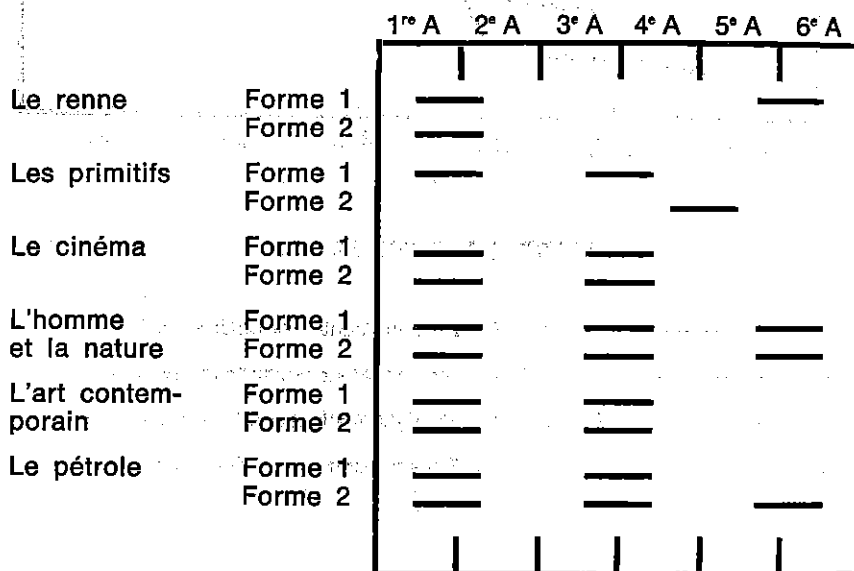
Ainsi se trouve confirmée l'observation des Américains L. Mosberg, T. Potter et R. Curnell qui ont souligné la sensibilité du test de closure au niveau pédagogique (*grade level*).

Ce pouvoir de la technique est naturellement de grande importance pour l'étalonnage des manuels scolaires.

Il serait toutefois illusoire d'espérer calibrer pour une seule année d'études, dans un système d'enseignement encore axé sur des classes rigides et non sur des groupes homogènes par branche : l'hétérogénéité reste trop grande.

Il est d'autant plus frappant de voir que, pour l'enseignement général, le test de closure réussit à discriminer significativement la première de la deuxième année et la troisième de la quatrième dans 20 cas sur 24.

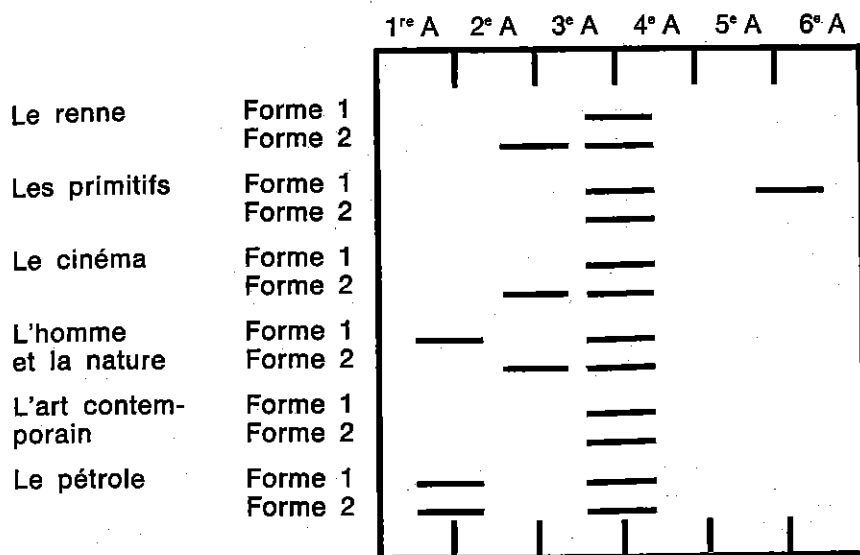
Enseignement général



— Différence significative à au moins P.05 (t de Student).

Dans l'enseignement technique, la discrimination se fait avec une parfaite netteté entre la troisième et la quatrième année. On observe ici une véritable rupture (toutes les différences, sauf une, sont statistiquement significatives à P.001) entre le secondaire inférieur et le secondaire supérieur.

Enseignement technique



— Différence significative à au moins P.05

Dans le système belge en vigueur partout jusqu'à ces dernières années, l'enseignement secondaire inférieur à orientation technique regroupe des populations d'aptitudes fort diverses : alors que certains élèves suivent des cours professionnels devant leur permettre d'exercer un métier dès la fin de l'obligation scolaire (14 ans), d'autres désirent, au contraire, acquérir une formation technique de haut niveau, pouvant d'ailleurs donner accès à l'enseignement supérieur.

A partir de la 4^e année, la compréhension des textes, telle qu'elle est mesurée par le test de closure, devient pratiquement aussi bonne dans l'enseignement technique que dans l'enseignement général, alors que dans les trois premières années, on observait, par contre, une nette supériorité de l'enseignement général.

Ce phénomène est admirablement mis en lumière par le calcul de la signification statistique des différences observées (t de student). Alors qu'elles sont pratiquement significatives partout à au moins P.05 (souvent à P.001), pour le cycle inférieur, on ne trouve, au contraire, presque aucune différence significative pour le supérieur.

Différences significatives à au moins P.05 entre enseignements général et technique.

		1 ^{re} A	2 ^e A	3 ^e A	4 ^e A	5 ^e A	6 ^e A
Le renne	Forme 1	—	—	—			
	Forme 2	—	—	—			—
Les primitifs	Forme 1	—	—	—			
	Forme 2	—		—			
Le cinéma	Forme 1	—	—	—			—
	Forme 2	—	—	—			
L'homme et la nature	Forme 1	—	—	—			—
	Forme 2	—	—	—			—
L'art contemporain	Forme 1	—	—	—			
	Forme 2		—	—			
Le pétrole	Forme 1	—	—	—			
	Forme 2	—	—	—			—

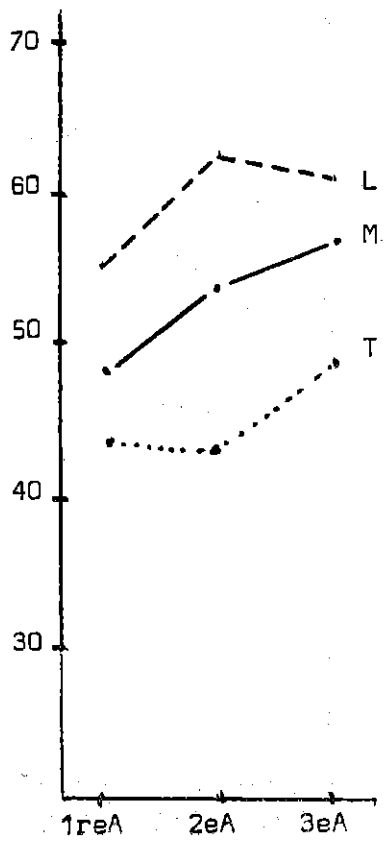
Les réactions à *L'art contemporain* méritent une attention particulière. Comme on le verra par la suite, la closure confirme la difficulté exceptionnelle de ce texte, déjà détectée par les experts. Ici non plus, les scores de l'enseignement technique ne sont pas significativement différents de ceux de l'enseignement général, alors que le facteur verbal (v) doit pourtant jouer à plein. On se souviendra que l'échantillon testé compte 1 544 élèves de technique, dont 630 du cycle supérieur.

Ces résultats devraient faire réfléchir ceux qui continuent à croire à l'infériorité du technique, voire à l'impossibilité des « humanités techniques », et regrettent encore l'omnivale des certificats de fin d'études secondaires supérieures.

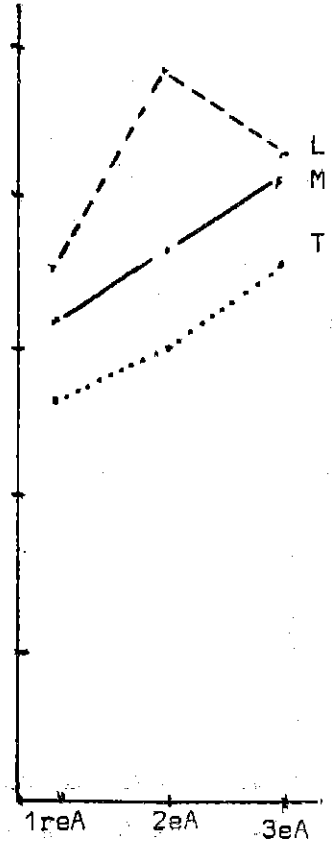
Frappé par l'extraordinaire pouvoir de discrimination de la technique de closure entre niveaux scolaires et entre types d'enseignement, dans le secondaire inférieur, nous avons alors fait éclater l'enseignement général en sections latine et moderne. Les graphiques montrent mieux que tout commentaire la netteté des résultats, calculés uniquement pour la Forme 1 des textes¹. Une hiérarchie Latine - Moderne - Technique apparaît partout, sans exception aucune.

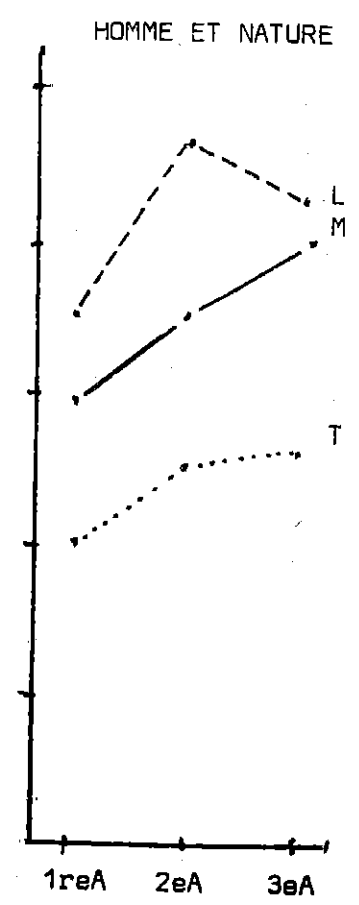
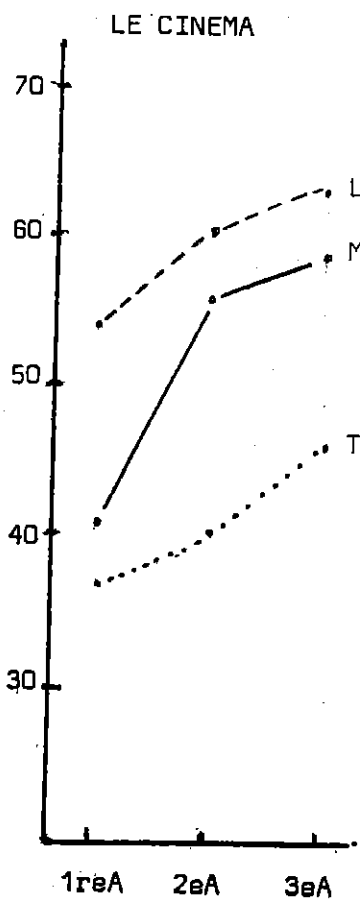
¹ On n'accordera pas trop d'importance à l'allure des profils des sections latines, car, dans certains cas, elles comptaient nettement moins d'élèves que les autres.

LE RENNE



LES PRIMITIFS

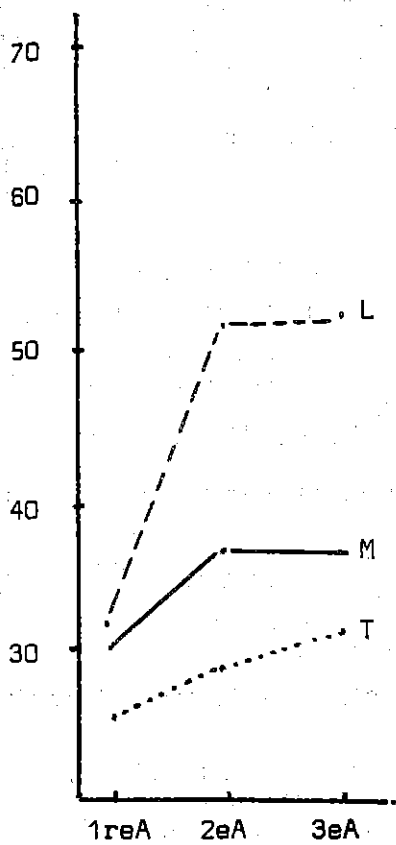




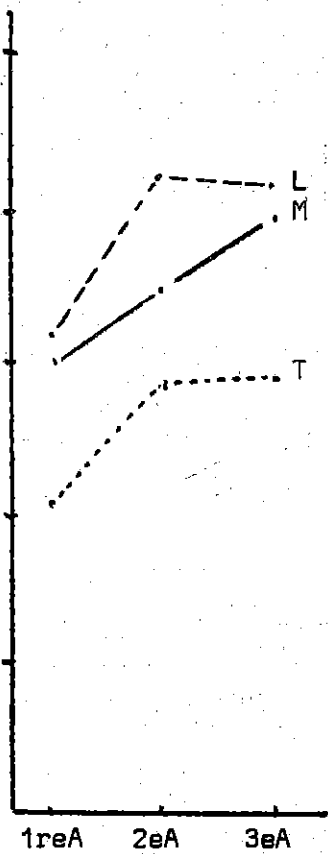
L = Latine M = Moderne T = Technique

10
8
6
4
2
0

ART CONTEMPORAIN



PETROLE



L = Latine

M = Moderne

T = Technique

Autre observation, assez frappante : si l'on se reporte au classement des textes opéré par les vingt juges, on constate que le texte *Le cinéma* est estimé plus difficile que *L'homme et la nature* (alors que les scores Flesch sont égaux). Les graphiques font aussi apparaître une différence, spécialement pour la première année du secondaire.

C. Comparaison des formes 1 et 2

La possibilité de créer aisément cinq formes parallèles de test de closure sur un même texte, en commençant respectivement les lacunes au 1^{er}, 2^e, 3^e, 4^e, 5^e mot, constitue certainement un des grands avantages potentiels de la technique.

Il importe donc de vérifier si des différences significatives de rendement se manifestent entre les deux formes expérimentales et, plus généralement, quel est le comportement des populations examinées.

Première constatation : *l'homogénéité des comportements.*

Observons, par exemple, le graphique concernant *Le renne* :

- a) Pour l'enseignement général comme pour le technique, le rendement de la forme 2 (closure 3, 8, 13, ...) est inférieur à celui de la forme 1 (closure 1, 6, 11, ...).
- b) Pour chaque année scolaire considérée isolément, cette observation se répète.
- c) La distance qui sépare les deux formes reste à peu près constante.

Ces tendances se vérifient dans la plupart des cas avec, naturellement, quelques irrégularités (8 % environ). L'extraordinaire pouvoir discriminatif de l'instrument pour les différents niveaux scolaires est ainsi une nouvelle fois confirmée.

Qu'en est-il maintenant de la *signification statistique des différences entre formes* ?

Les deux tableaux récapitulatifs suivants montrent que, dans plus de 80 % des cas, les deux formes peuvent être considérées comme équivalentes. Même si l'on choisissait le seuil de P.10 au lieu de P.05, l'équivalence subsisterait dans environ 70 % des cas.

Le texte *Le cinéma* seul rompt le parallélisme pour l'enseignement technique et participe aux anomalies pour l'enseignement général.

Quoi qu'il en soit, l'équivalence des formes semble établie dans une large mesure. Par précaution, on devrait néanmoins administrer au moins deux formes différentes chaque fois que le nombre de sujets le permet.

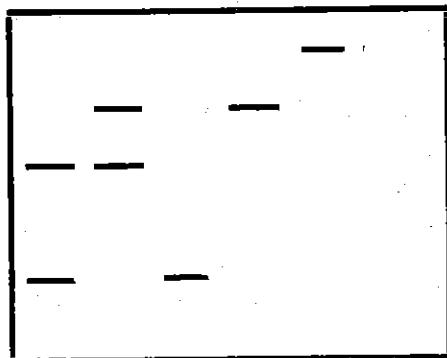
Une façon plus rigoureuse de procéder serait d'administrer les cinq formes à des échantillons tirés de la même population que celle dont on souhaite tester la compréhension et de retenir comme épreuve la forme qui donne le score moyen médian.

COMPARAISON FORME 1 - FORME 2

Enseignement général

1^{re} A 2^e A 3^e A 4^e A 5^e A 6^e A

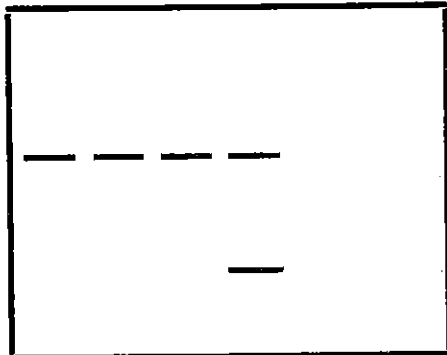
Le renne
 Les primitifs
 Le cinéma
 L'homme et la nature
 L'art contemporain
 Le pétrole



Enseignement technique

1^{re} A 2^e A 3^e A 4^e A 5^e A 6^e A

Le renne
 Les primitifs
 Le cinéma
 L'homme et la nature
 L'art contemporain
 Le pétrole



— indique une différence statistiquement significative à au moins P.05 (t de Student)

D. Les dispersions

Il ne faut évidemment pas perdre de vue que les différences que nous étudions sont des différences de *moyennes*.

Dans une forme d'enseignement peu individualisé, où les classes restent rigides, on observe des marges de variation considérables à chaque niveau scolaire, dans le domaine des aptitudes, des *skills* ou des connaissances.

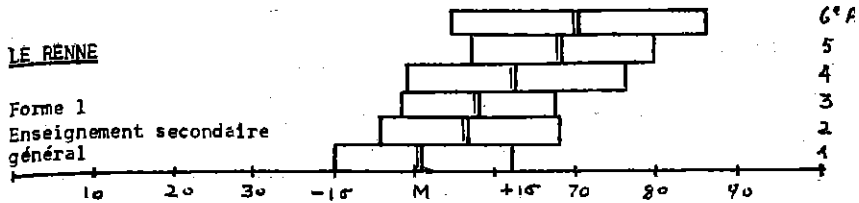
La capacité en lecture n'échappe pas à ce phénomène. Des auteurs comme Lindquist, Cook, Cornell évaluent la marge de variation à 7 ou 8 ans, à la fin de l'école primaire.

L'observation se renouvelle dans notre recherche : une fraction considérable des élèves des années précédentes ou suivantes obtiennent des résultats égaux ou supérieurs aux résultats moyens d'une année particulière. La superposition des courbes de distribution montre clairement la situation. De nouveau, la cassure entre enseignement secondaire inférieur et supérieur apparaît nettement.

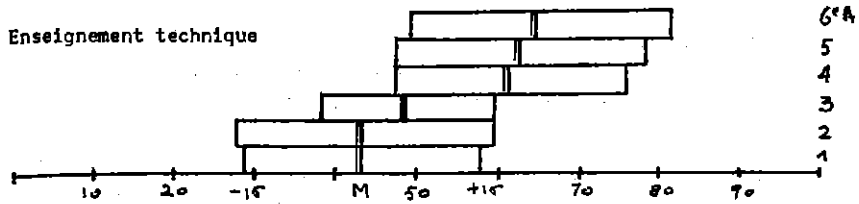
Les graphiques suivants montrent, pour la Forme 1 de chacun des textes, les moyennes et, de part et d'autre, un écart type.

LE RENNE

Forme 1
Enseignement secondaire
général

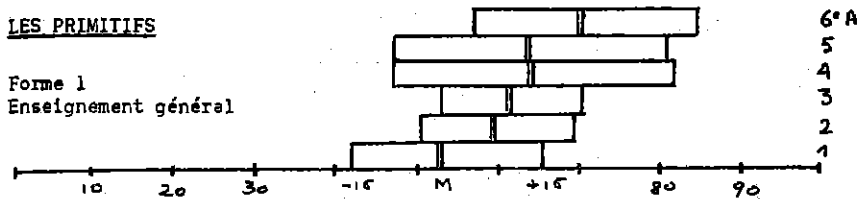


Enseignement technique

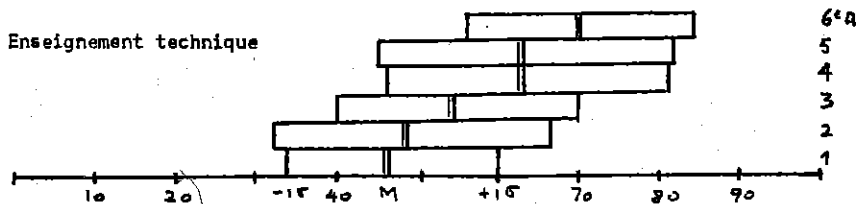


LES PRIMITIFS

Forme 1
Enseignement général

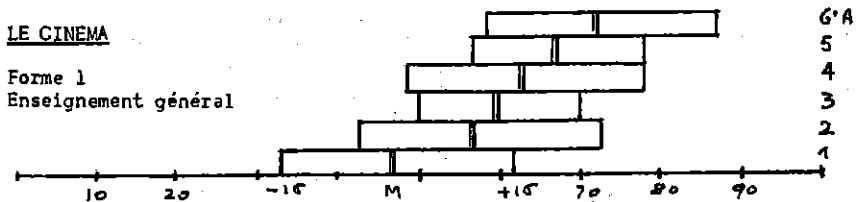


Enseignement technique

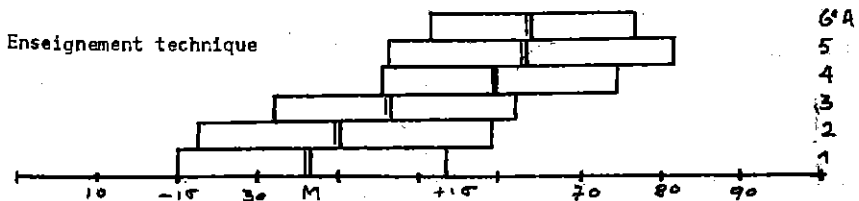


LE CINEMA

Forme 1
Enseignement général

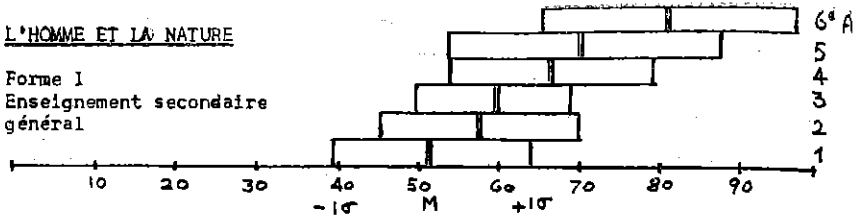


Enseignement technique

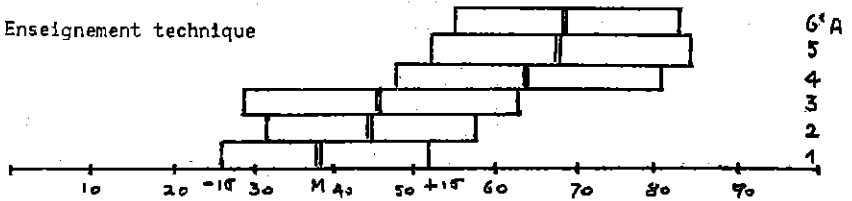


L'HOMME ET LA NATURE

Forme 1
Enseignement secondaire
général

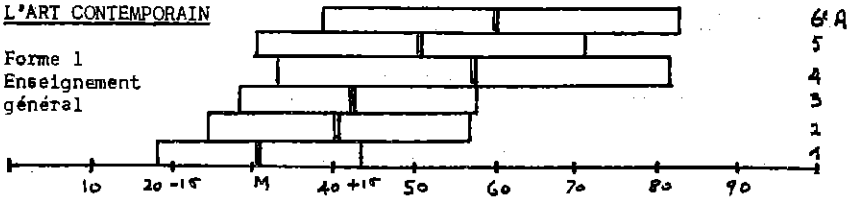


Enseignement technique

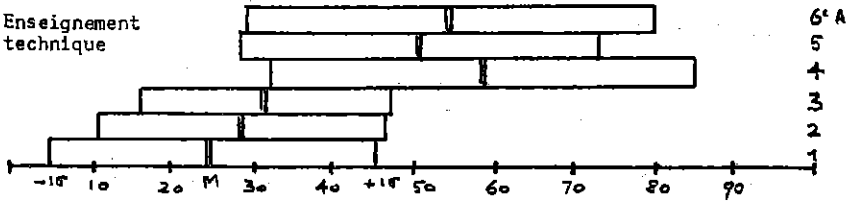


L'ART CONTEMPORAIN

Forme 1
Enseignement
général

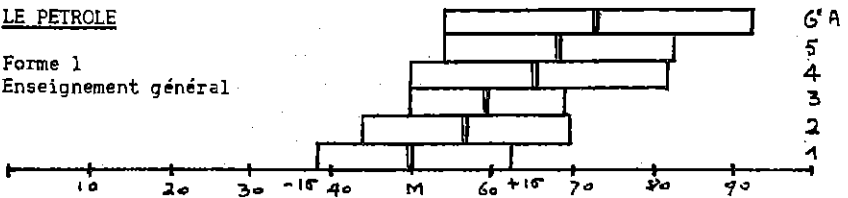


Enseignement
technique

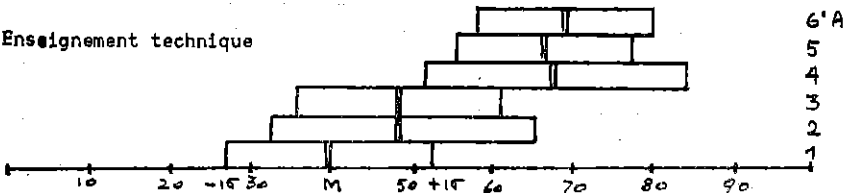


LE PETROLE

Forme 1
Enseignement général



Enseignement technique



II. Enseignements général et technique séparés, formes 1 et 2 réunies

La constatation de l'élévation soudaine du rendement dans l'enseignement technique, à partir de la 4^e année, est suffisamment importante pour que nous y revenions.

Pour mieux faire apparaître l'évolution des résultats, nous avons calculé les scores moyens, pour les formes 1 et 2, dans les deux types d'enseignement.

Les graphiques sont, de nouveau, très parlants et confirment ce que nous avons observé antérieurement.

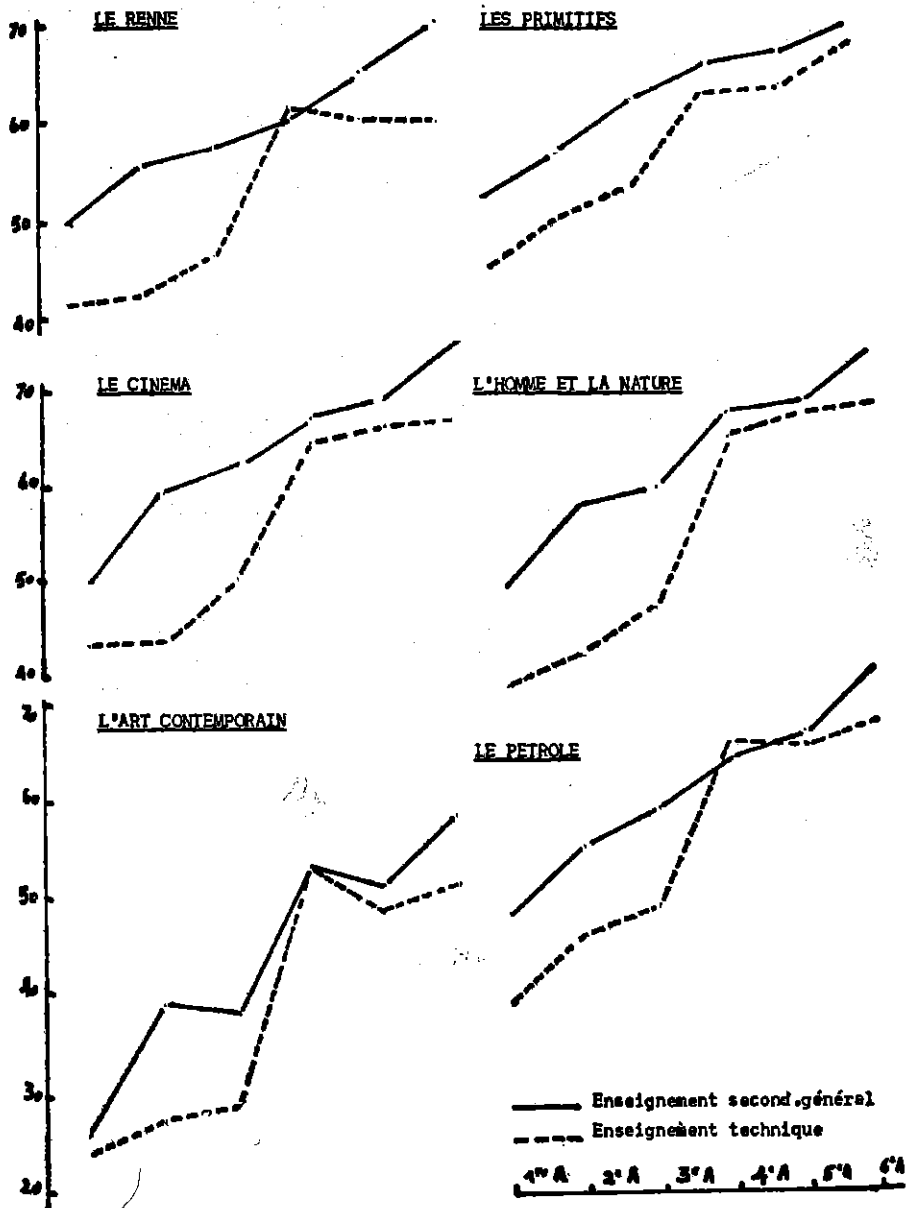
Au cycle inférieur, une remarque complémentaire peut être faite : l'écart moyen entre le général et le technique est plus marqué en 2^e et en 3^e année qu'en 1^{re}. L'enseignement technique semble subir à ce moment le frein d'une partie de sa population qui s'es-souffle dans ses études.

SCORES MOYENS

Formes 1 et 2 réunies

	1° A		2° A		3° A		4° A		5° A		6° A	
	G	T	G	T	G	T	G	T	G	T	G	T
Le renne	49,60	41,65	55,35	41,85	57,35	48,15	60,30	61,40	64,85	60,70	69,60	60,50
Les primitifs	52,25	44,50	56,25	49,00	61,50	52,75	65,25	62,85	66,50	63,22	69,50	67,50
Le cinéma	50,15	43,25	59,00	43,50	61,85	50,00	66,70	64,12	68,00	65,50	74,30	66,30
L'homme et la nature	49,80	38,80	57,50	42,25	59,20	47,10	66,78	64,50	69,90	66,80	74,75	70,25
L'art contemporain	29,60	24,75	39,60	27,40	38,30	28,80	53,50	53,50	51,50	49,25	58,25	51,50
Le pétrole	48,20	39,75	55,55	46,40	59,08	48,80	64,85	66,40	67,10	65,80	74,00	68,40

SCORES MOYENS - FORMES 1 ET 2 REUNIES



III. Les deux formes et les deux enseignements réunis

C'est surtout en pensant à la mesure de la lisibilité pour des populations assez hétérogènes, telles qu'elles dominent dans le monde adulte, qu'il nous a paru intéressant de fondre toutes les moyennes.

Le diagramme montre que, pour cinq textes :

- a) La progression des pourcentages de réussite se marque admirablement en fonction du niveau pédagogique.
- b) Les moyennes sont assez proches ; seul le texte *L'art contemporain* fait exception.

Nous voyons ainsi se confirmer une hypothèse émise lors de l'analyse des évaluations de la difficulté des textes par un groupe de juges. On se souviendra que la grande majorité estimait que tous les textes, sauf précisément *L'art contemporain*, étaient compréhensibles pour tout l'enseignement secondaire inférieur.

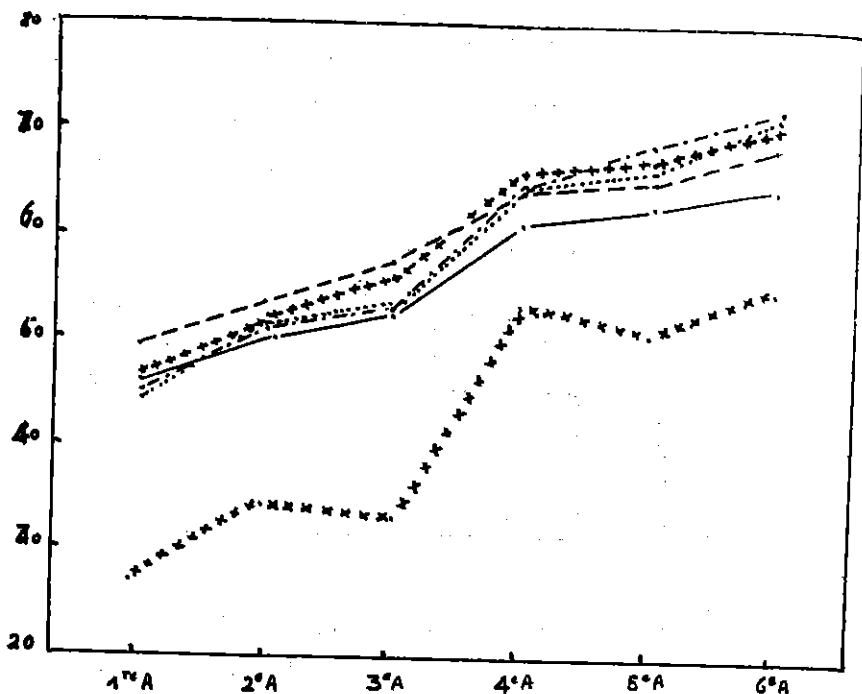
Nous avons vu que 44 % de réussite constitue un repère grossier permettant de dire à partir de quel moment un texte est utilisable avec un groupe d'élèves.

Le diagramme confirme cette observation. Dès la première année, cinq textes répondent au critère d'admissibilité alors que le sixième, toujours *L'art contemporain*, reste près de la limite de l'acceptable dans le secondaire supérieur.

SCORES MOYENS

Les deux formes et les deux enseignements réunis

	1 ^{er} A	2 ^e A	3 ^e A	4 ^e A	5 ^e A	6 ^e A
Le renne	46,10	49,50	52,00	60,90	62,50	64,00
Les primitifs	48,90	53,30	56,80	63,90	64,60	68,40
Le cinéma	46,10	52,10	55,20	65,30	66,60	69,70
L'homme et la nature	44,80	50,90	52,70	65,10	68,20	72,40
L'art contemporain	27,50	34,00	33,00	53,40	50,20	54,40
Le pétrole	44,80	51,40	53,80	65,70	66,30	70,70



LES DEUX FORMES ET LES DEUX ENSEIGNEMENTS REUNIS

- Renne.
- - - - - Primitifs.
- + + + + + Cinéma.
- . - . - . Homme et nature.
- Pétrole.
- x x x x x Art contemporain.

IV. Comparaison avec les scores de lisibilité de R. Flesch

Intuitivement, nous sentions que si la technique de closure de Taylor est, comme on l'affirme, nettement supérieure au test de lisibilité de R. Flesch, une différence nette devait apparaître entre les textes *Le Pétrole* et *L'art contemporain* affectés de scores Flesch fort proches, respectivement 22 et 17.

Les chiffres sont absolument concluants : sans exception aucune, le calcul des *t* de Student révèle des différences très significatives, à au moins P.001, à tous les niveaux pédagogiques, que l'on compare les formes 1 ou les formes 2.

t de Student

L'art contemporain - Le pétrole

Année	Forme 1 - Forme 1			Forme 2 - Forme 2		
	Degré de liberté	t	Signification	Degré de liberté	t	Signification
1	316	9,69	P.001	310	9,81	P.001
2	249	8,08	P.001	259	7,84	P.001
3	226	8,70	P.001	208	13,20	P.001
4	347	4,00	P.001	350	6,56	P.001
5	218	6,03	P.001	208	5,37	P.001
6	155	4,48	P.001	150	5,82	P.001

Pour les autres paires de textes affectés des mêmes scores de Flesch, on ne prévoyait pas de différences marquées.

Différences significatives à P.05

Le renne - Les primitifs
(scores Flesch : 51,5 - 47)

	1 ^{re} A	2 ^e A	3 ^e A	4 ^e A	5 ^e A	6 ^e A
Formes 1		—	—	—		
Formes 2	—		—		—	

Le cinéma - L'homme et la nature
(scores Flesch: 38,5 - 36,5)

	1 ^{re} A	2 ^e A	3 ^e A	4 ^e A	5 ^e A	6 ^e A
Formes 1					—	—
Formes 2	—	—	—			

Le nombre de cas où les différences sont significatives à au moins P.05 n'est pas négligeable, surtout si l'on considère l'ensemble des deux formes.

Il est intéressant de voir comment la nouvelle formule élaborée par G. Henry pour la langue française classe ces textes¹.

Comme on pouvait s'y attendre les scores Flesch ne seront pas ici gravement démentis.

¹ G. HENRY, *Construction de trois formules de lisibilité spécifiques à la langue française*, Université de Liège, 1973 (thèse de doctorat).

Dans la recherche de G. Henry, les épreuves-critères ont été constituées par des tests de closure. 116 variables linguistiques ont été prises en considération. En appliquant sa formule la plus complexe, G. Henry obtient les scores suivants pour nos textes, ces scores de lisibilité étant exprimés en pourcentage de réussite probable en closure, en 5^e-6^e primaires.

L'art contemporain	17 %
L'homme et la nature	22 %
Le pétrole	24 %
Le cinéma	26 %
Les primitifs	30 %
Le renne	30 %

L'erreur standard étant d'environ 4 %, on voit apparaître une répartition des textes qui vient intégralement confirmer nos observations. On remarquera, en particulier, la situation du texte *Le pétrole*.

Il importe néanmoins d'insister sur la faible différence entre les pourcentages de réussite.

En tenant compte de l'erreur standard et de la marge laissée par les seuils de probabilité (signification des différences), on doit admettre une fourchette de quelque 8-10 %.

Ainsi s'expliquera, par la suite, la situation relative des textes pour l'enseignement primaire.

V. Etude de la fidélité des épreuves

Les observations des chercheurs américains sont largement confirmées : que l'on calcule le coefficient *split-half* ou le Kuder-Richardson formule 20, les chiffres sont élevés : ils oscillent entre .82 et .97 ; le tiers des coefficients se regroupent entre .90 et .91 (mode de la distribution).

		1° A	2° A	3° A	4° A	5° A	6° A
Le renne	1	.855 .880	.901 .887	.825 .790	.868 .863	.880 .869	.910 .906
	2	.879 .890	.900 .887	.886 .859	.918 .876	.871 .846	.934 .919
Les primitifs	1	.847 .860	.894 .869	.856 .848	.908 .886	.938 .906	.917 .884
	2	.837 .849	.914 .872	.899 .851	.922 .879	.905 .876	.941 .917
Le cinéma	1	.915 .937	.939 .919	.939 .897	.944 .916	.905 .907	.911 .902
	2	.906 .930	.932 .906	.913 .869	.953 .919	.941 .901	.961 .921
L'homme et la nature	1	.879 .894	.913 .878	.911 .904	.931 .903	.941 .918	.946 .915
	2	.886 .916	.945 .915	.922 .887	.949 .919	.940 .912	.971 .935
L'art con- temporain	1	.915 .938	.957 .921	.939 .912	.971 .949	.970 .938	.969 .942
	2	.903 .935	.959 .925	.931 .888	.976 .948	.977 .951	.966 .946
Le pétrole	1	.865 .894	.900 .897	.917 .862	.931 .911	.920 .890	.928 .901
	2	.890 .929	.934 .902	.857 .824	.914 .899	.950 .913	.908 .873

* Le premier des deux chiffres est le coefficient *split-half*, le second, le Kuder-Richardson.

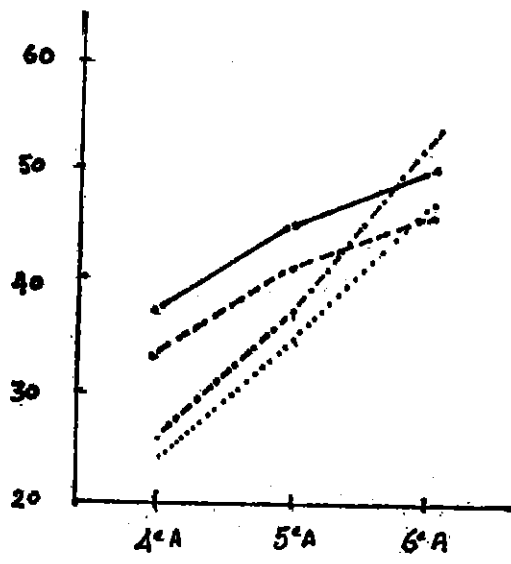
7. RESULTATS : ENSEIGNEMENT PRIMAIRE

1. Belges et étrangers séparés Formes 1 et 2 séparées

A. Scores moyens et graphiques

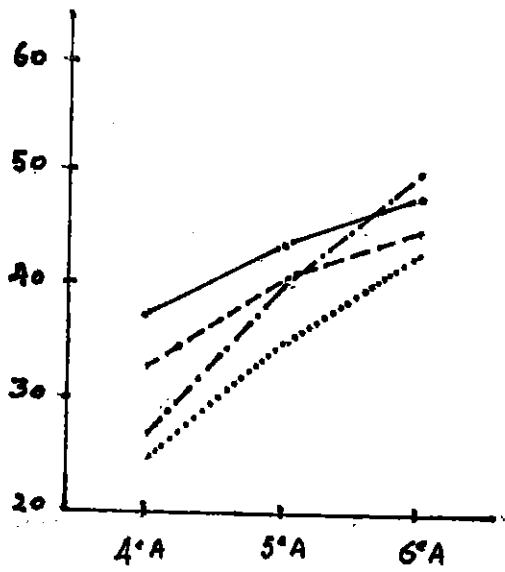
SCORES MOYENS

	4 ^e année		5 ^e année		6 ^e année		
	Belges	Etrangers	Belges	Etrangers	Belges	Etrangers	
Le renne	Forme 1	25,52	45,05	37,05	49,74	53,53	
	Forme 2	32,88	24,37	41,10	35,17	45,73	46,24
L'homme et la nature	Forme 1	37,63	27,01	43,59	39,63	47,73	49,61
	Forme 2	32,82	25,47	40,71	35,16	44,61	42,63
Le pétrole	Forme 1	29,51	21,44	36,49	29,19	43,95	47,17
	Forme 2	27,31	16,56	33,45	27,87	42,80	41,17



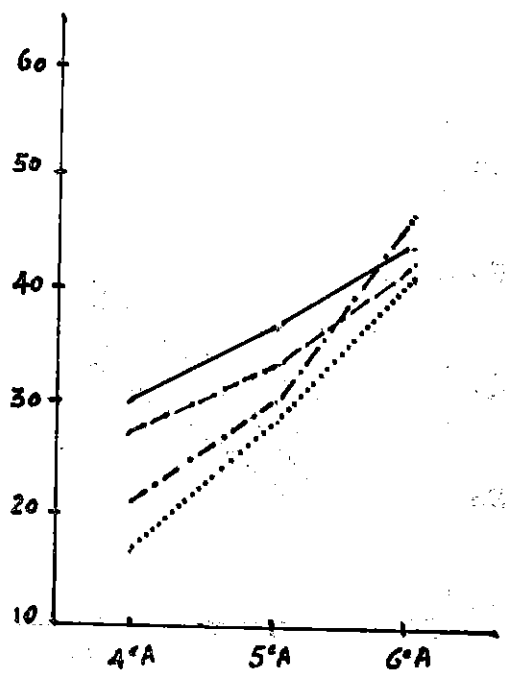
LE RENNE

- Forme 1-6 : Enseignement primaire : Belges.
- - - - - Forme 3-8 : Enseignement primaire : Belges.
- .-.-.-.- Forme 1-6 : Enseignement primaire : Etrangers.
- Forme 3-8 : Enseignement primaire : Etrangers.



L'HOMME ET LA NATURE

- Forme 1-6 : Enseignement primaire : Belges.
- Forme 3-8 : Enseignement primaire : Belges.
- . - . - . - . Forme 1-6 : Enseignement primaire : Etrangers.
- Forme 3-8 : Enseignement primaire : Etrangers.



LE PETROLE

- Forme 1-6 : Enseignement primaire : Belges.
- - - - - Forme 3-8 : Enseignement primaire : Belges.
- . - . - . Forme 1-6 : Enseignement primaire : Etrangers.
- Forme 3-8 : Enseignement primaire : Etrangers.

B. Constatations générales

Il saute d'abord aux yeux que, quels que soient le texte ou la forme, qu'il s'agisse de Belges ou d'étrangers, l'épreuve discrimine remarquablement entre années scolaires. Sans exception aucune, on observe une croissance progressive des scores moyens, selon une ligne presque droite.

Le contrôle statistique confirme cette impression directe.

Pour les élèves belges, on obtient les chiffres suivants :

Texte		Seuil de signification de la différence entre	
Le renne	Forme 1	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .001 P = .01
	Forme 2	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .001 P = .02
L'homme et la nature	Forme 1	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .001 P = .02
	Forme 2	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .001 P = .02
Le pétrole	Forme 1	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .001 P = .001
	Forme 2	4 ^e et 5 ^e année 5 ^e et 6 ^e année	P = .01 P = .001

Bien que le nombre d'élèves étrangers soit trop peu élevé pour permettre des conclusions fermes, une tendance générale ne peut cependant manquer de frapper. Alors qu'en 4^e année, les étrangers se situent, dans tous les cas, nettement en dessous des Belges, ils remontent peu à peu pour obtenir, en 6^e année, des résultats proches ou supérieurs à ceux des Belges.

Des différences d'âge nous mettront-elles sur le chemin de l'explication ?

Nous avons tiré, au hasard, dans chaque année scolaire, un nombre d'élèves belges approximativement égal à celui des étrangers.

Comme le montrent les polygones de fréquence ci-dessous :

1° Les populations belges sont mieux groupées et plus jeunes que les étrangères en 4^e et 5^e années. En 6^e année, la différence n'apparaît plus très nettement.

Les médians calculés au mois de juin sont les suivants :

4° année Belges 10 ans 4 mois
Etrangers 11 ans
Différence significative à $P = .05$

5° année Belges 11 ans
Etrangers 11 ans 9 mois
Différence significative à $P = .02$

6° année Belges 12 ans 3 mois
Etrangers 12 ans 9 mois
Différence non significative

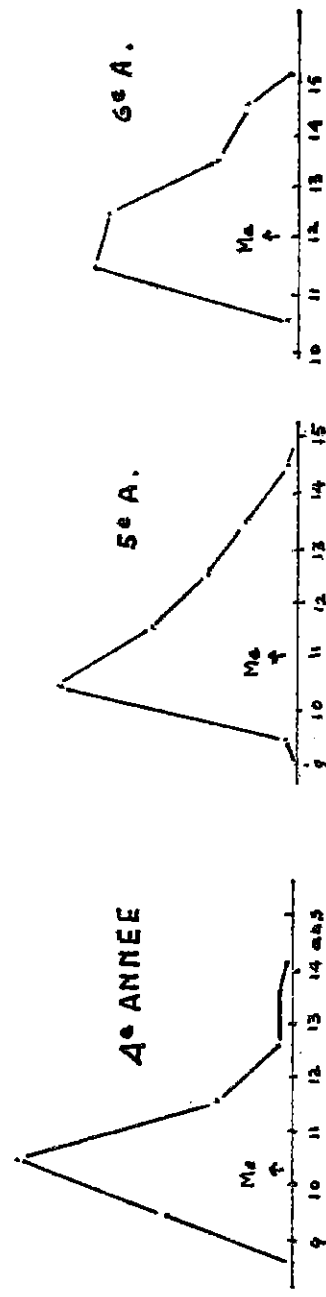
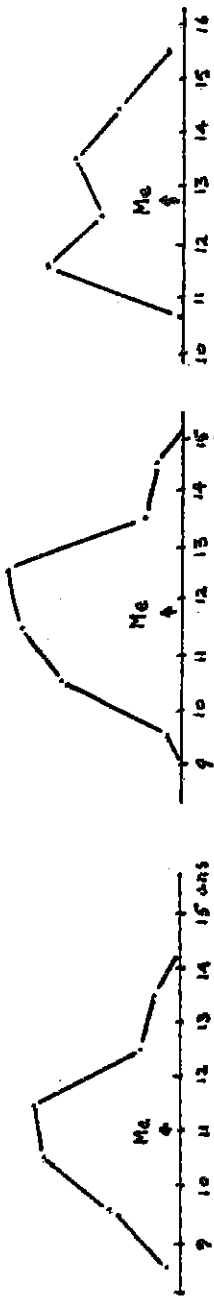
2° Les marges de variation sont partout considérables (nombres arrondis) :

4° année Belges de 8;6 à 14 : 5 ans 6 mois
Etrangers de 8;6 à 14;3 : 5 ans 9 mois

5° année Belges de 9;6 à 15 : 5 ans 4 mois
Etrangers de 9 à 15;3 : 6 ans 3 mois

6° année Belges de 10;6 à 15;3 : 4 ans 9 mois
Etrangers de 10;6 à 15;6 : 5 ans

ETRANGERS



BELGES

Tant les polygones de fréquences que les chiffres montrent qu'en 6^e année, il n'existe plus de grande différence d'âge entre les deux populations.

Comment cette égalisation d'âges et de rendement s'est-elle produite ? Par une sélection opérée spontanément par les parents étrangers surtout et, dans une mesure moindre, par les Belges. En effet, las de voir leur enfant répéter classe après classe et se trouver encore à l'école primaire à 14 ou 15 ans, alors qu'il devrait normalement en sortir à 12 ans, les parents se tournent vers l'enseignement secondaire technique. La majorité des élèves très retardés à l'école primaire viennent y chercher une simple formation professionnelle leur permettant d'exercer une activité rémunérée le plus rapidement possible (les cas les plus accusés se retrouvent dans les sections A4, en Belgique).

Cette entrée relativement massive d'élèves faibles explique l'espèce de régression que nous allons observer, au début du technique (voir ci-après, point IV : *L'enseignement primaire face au secondaire général et au technique*).

C. Comparaison des formes 1 et 2

Comme on l'avait déjà observé dans l'enseignement secondaire, la forme 1 se révèle souvent plus facile que la forme 2. La différence n'est pas très élevée ; le phénomène pourrait cependant ne pas être fortuit. La cause éventuelle devra être recherchée et seule l'analyse des *items* peut normalement la fournir.

Existe-t-il une différence statistiquement significative entre les deux formes d'un même texte ? Oui, dans six cas sur neuf.

	4 ^e année	5 ^e année	6 ^e année
Le renne	P = .02	P = .02	P = .01
L'homme et la nature	P = .01	P = .05	P = .02
Le pétrole	NS	NS	NS

Pour obtenir une évaluation sûre, il semble donc opportun, ici aussi, de faire subir au moins deux épreuves à un même élève.

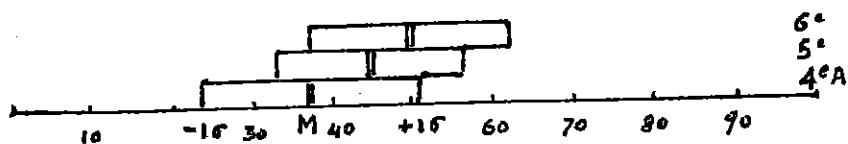
D. Les dispersions

Les scores des élèves belges se distribuent normalement. Le nombre d'étrangers n'était pas assez élevé pour permettre une épreuve de normalité.

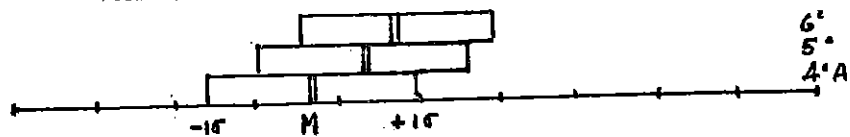
Les graphiques montrent clairement un harmonieux étagement des moyennes et une grande unité dans les dispersions.

ENSEIGNEMENT PRIMAIRE : ELEVES BELGES
Moyennes et dispersions

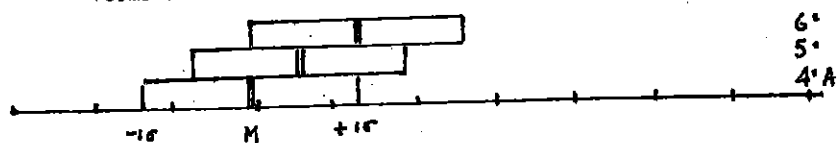
Le Renne
Forme 1



L'Homme et la Nature
Forme 1



Le Pétrole
Forme 1



II. Groupes et formes réunis

Cette rubrique n'est reprise que pour mémoire.

Calculer les moyennes pour les formes 1 et 2 réunies ne nous apprendrait rien d'important.

Par ailleurs, fondre les résultats des élèves belges et des étrangers n'est guère indiqué puisque le rapport entre les deux groupes est, en gros, de 1 à 10.

III. Comparaison avec les scores de lisibilité de R. Flesch

On se souviendra que les vingt juges consultés ont considéré que le texte *Le renne* (score Flesch 50) convenait à l'école primaire. Cinq juges ont estimé que *L'homme et la nature* (score Flesch 35) pourrait être utilisé avant le secondaire et neuf juges ont émis un avis similaire pour *Le pétrole* (score Flesch 20).

Les résultats obtenus sont pratiquement les mêmes pour *Le renne* et pour *L'homme et la nature*. Une différence apparaît pour *Le pétrole*.

Voici les données pour la Forme 1, élèves belges (nombres arrondis) :

	4 ^e année		5 ^e année		6 ^e année	
	M	σ	M	σ	M	σ
Le renne	37	13	45	12,5	50	13
L'homme et la nature	38	14	44	13	48	13
Le pétrole	30	13	36	13,5	44	14

Le contrôle statistique de la signification des différences entre textes confirme de nouveau la première impression. Le texte *Le pétrole* est plus difficile que les deux autres, entre lesquels on ne note aucune différence significative. Les élèves de l'enseignement primaire semblent avoir été plus sensibles à la différence de longueur moyenne des mots (souvent révélatrice du niveau d'abstraction) que les élèves du secondaire.

Le renne : 156 syllabes pour 100 mots

L'homme et la nature : 180 syllabes pour 100 mots

Le pétrole : 194 syllabes pour 100 mots

<i>Le renne - L'homme et la nature</i>	4 ^e année	5 ^e année	6 ^e année
Forme 1 - Forme 1	NS	NS	NS
Forme 2 - Forme 2	NS	NS	NS

Le renne - Le pétrole

Forme 1 - Forme 1	P = .001	P = .001	P = .001
Forme 2 - Forme 2	P = .001	P = .001	NS

L'homme et la nature - Le pétrole

Forme 1 - Forme 1	P = .001	P = .001	P = .01
Forme 2 - Forme 2	P = .01	P = .001	NS

Si l'on admet qu'un texte convient à un groupe de lecteurs s'ils obtiennent au moins 44 % de réussite à l'épreuve de closure, *Le renne* et *L'homme et la nature* iraient à la 5^e primaire tandis que *Le pétrole* serait conseillé à partir de la 6^e. Ces indications paraissent raisonnables. Nous ne possédons pas de données suffisantes pour nous référer aux nouveaux scores-critères proposés par Bormuth en 1971.

On note, encore une fois, que le score Flesch du texte *Le pétrole* se révèle trompeur.

IV. L'enseignement primaire face au secondaire général et au technique

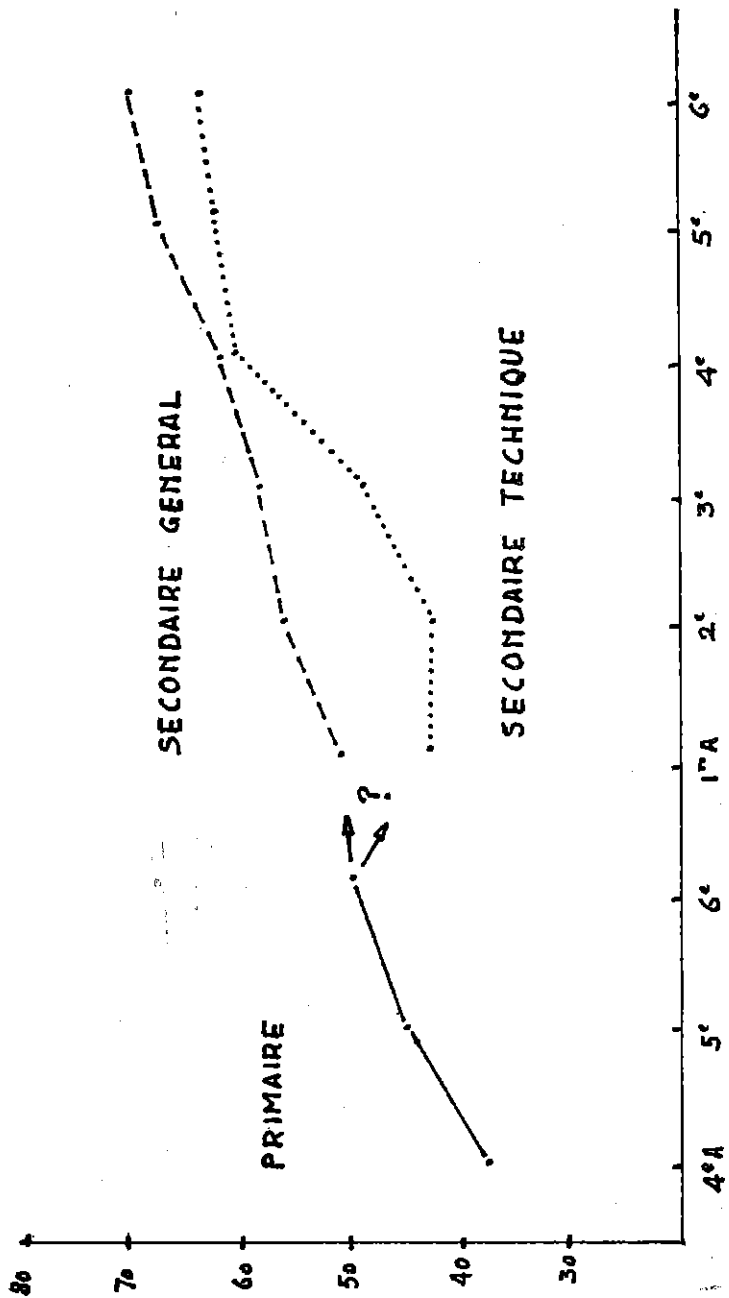
Lors de l'examen des résultats de l'enseignement secondaire, nous avons observé, au cycle inférieur, un net décalage entre l'enseignement général et l'enseignement technique. Il confirmait une sélection, au moins dans le domaine verbal, au profit de l'enseignement général.

Pour préciser la situation, nous avons représenté graphiquement l'évolution complète de la 4^e primaire à la fin du secondaire pour le texte *Le renne* (Forme 1).

L'écart entre le général et le technique est net. En 1^{re} et en 2^e année de l'enseignement technique, la moyenne se situe au niveau d'une 5^e primaire non sélectionnée, et la 3^e du technique se place au niveau de la 6^e primaire. Nous avons proposé plus haut une hypothèse explicative de ce phénomène.

Comme nous l'avons déjà vu aussi, la situation se redresse complètement à partir de la 4^e année technique (A2 en Belgique), précisément parce que, à partir de ce niveau, les élèves faibles directement destinés à la profession, ont quitté l'enseignement.

Il faut aussi tenir compte d'un apport non négligeable d'étudiants quittant l'enseignement secondaire général pour le technique.



Le renne - Forme 1-6

CONCLUSION

De façon générale, les observations et les résultats fournis par la présente étude confirment les conclusions des recherches américaines ; le test de closure est un instrument très économique fournissant une mesure directe de la lisibilité, de validité élevée, et une mesure fort satisfaisante de la compréhension.

Comme on pouvait s'y attendre, le français se prête aussi bien que l'anglais à la technique.

Sondant en premier lieu la connaissance générale de la langue, le test de closure constitue donc un outil de jaugeage et de diagnostic tant pour les professeurs de langue maternelle que de langues étrangères. Il est, croyons-nous, susceptible de rendre d'importants services dans la pratique scolaire.

Le test de closure est aussi un instrument de recherche à usages multiples.

Utilisé comme mesure de critère, il permet d'élaborer des formules de lisibilité d'une validité jamais atteinte jusqu'à ces dernières années. A travers ces formules, il peut donc aider les auteurs ou les éditeurs de manuels scolaires et, plus généralement, tous ceux qui désirent opérer un contrôle systématique de l'intelligibilité de l'information.

Le test de closure peut-il être accepté comme instrument *universel* de mesure de la compréhension ? Nous n'oserions pas l'affirmer, bien que nous ayons, dès à présent, acquis la conviction qu'il mesure nettement plus que de courtes portions des messages.

En attendant des données complémentaires sur le champ de validité, un testage fin de la compréhension peut aisément être obtenu en combinant un test de vocabulaire classique avec un test de closure et un test de compréhension générale à réponses fermées ou ouvertes.

Pour le langage écrit, on va enfin pouvoir aborder, sur une large échelle, l'étude de l'intelligibilité des textes littéraires, — surtout de la poésie, — et scientifiques à laquelle on avait jusqu'à présent dû renoncer faute d'outils adéquats.

Après de longs tâtonnements, nous entrevoyons aussi la façon d'utiliser une technique parallèle à celle de Taylor pour mesurer les messages audio-visuels.

ANNEXES

ANNEXE I

Les six textes choisis

On a imprimé en gras les mots qui ont été supprimés dans la première forme du test de closure et en italique les mots supprimés dans la seconde forme.

LE RENNE

De la *même* famille que le cerf, le renne est **plus** bas et en même **temps** plus *gros* que lui. **Ses** jambes *sont* plus courtes et plus *massives* ; ses pieds **sont** plus *larges* et son **poil** plus *fourni*. Sa tête **est** ornée de bois superbes **qui** lui *servent* à se **défendre**.

Le *renne* vit par **troupe**s nombreuses, *dans* les régions **glaciales** des *deux* continents. Il **est** surtout très commun en **Amérique** où *on* lui donne le nom *de* caribou.

Cet **utile** ruminant *constitue* l'unique **richesse** des *contrées* de l'**extrême** Nord *qui*, sans lui, **seraient** inhabitées.

Son lait procure **une** boisson *saine* et fortifiante ; **sa** chair, *une* nourriture substantielle ; **sa** peau, *d'excellentes* fourrures et des *chaussures* très souples. **Ainsi**, le *renne* donne tout **seul**, *ce* que nous tirons **du** cheval, *de* la vache et du *mouton*.

Il n'est pas *jusqu'à* ses **excréments** que *l'on* ne sèche **pour** brûler.

Quand la terre **est** couverte de neige, on **attelle** les *rennes* à des **traîneaux** avec *lesquels* on parcourt **parfois** vingt-cinq *lieues* en un **jour**. Ces *quadrupèdes* sont, en **effet**, très *agiles* et leurs **sabots** sont *bien* conformés pour **courir** sur le sol neigeux **sans** s'y enfoncer.

Aussi **doux** que *laborieux*, soumis à **son** maître, *dur* à la **peine**, le *renne* ne refuse **jamais** ses *services* et ne **s'arrête** que vaincu par la fatigue.

Sa sobriété est **bien** connue. // se nourrit **de** toutes *espèces* de plantes et n'est pas difficile **sur** le *choix* de ses **mets**.

LES PRIMITIFS

Je crois avoir rencontré au cours de mes voyages les êtres les plus primitifs de notre planète dans certaines régions inexplo-
rées de l'Australie.

Je les vis par petits groupes d'une dizaine d'hommes, de quatre ou cinq femmes et de quelques enfants. A l'écart, assis dans la pierraille, un indigène, à l'aide d'un gros galet rond, taillait de longs silex destinés à armer des lances. Il en avait des douzaines qui serviraient de monnaie d'échange à l'intérieur du pays.

Toutes les femmes avaient un gros sac près d'elles. Se déplaçant sans cesse à la recherche de nourriture, ces gens sont bien obligés d'emporter tout ce qu'ils possèdent et ce sont les femmes qui servent de bêtes de somme.

Ils sont lamentablement bornés. Quelques-uns savaient compter jusqu'à cinq, mais la plupart jusqu'à trois seulement.

Ils mangent ce qu'ils trouvent ou parviennent à prendre : oiseaux, œufs, grenouilles, lézards, des feuilles, des racines, de grosses fourmis, du miel. Tout leur semble bon, même la pourriture.

Dans un des sacs, je vis des pierres taillées, des tendons de kangourou : c'est leur fil ; des os pointus et percés : ce sont leurs aiguilles, sans compter des racines comestibles, de la terre rouge, blanche ou jaune pour se décorer le corps, une série de peaux de bêtes et une grosse pierre plate pour la fabrication du pain et de la cire.

Seuls, deux ou trois hommes avaient un boomerang, morceau de bois recourbé au centre, aux extrémités pointues et qui, bien jeté, frappe le gibier, tue des oiseaux surtout, pour retomber non loin du lanceur.

L'HOMME ET LA NATURE

L'homme, *grâce* à son **intelligence** et sa ténacité, a **profondément** modifié *la* terre.

Sur *le* sol *ingrat* et sur **les** éléments, *il* a remporté **et** *il remporte* journallement des **victoires**. L'homme a transformé **le** tapis *végétal*. Des terres **infertiles** sont *devenues* fécondes. **Il** a créé *des* paysages ; **il** a construit ; *il* a détruit ; **il** a *modifié* la nature.

Vous constaterez *que* moins un **peuple** est *civilisé*, plus **il** *subit* l'**action** de la **nature** parce *qu'il* n'a pas *découvert* les moyens **de** *lutter* contre celle-ci **et** de *la* transformer.

Au *contraire*, les *peuples* civilisés subissent **moins** l'**action** de la **nature**. **Ils** *peuvent* transformer et **améliorer** le **sol**, conquérir des **terres** sur *la* mer, construire **de** grands *barrages* pour éviter **les** inondations *et* produire de l'électricité, *transporter* des produits **dans** des *régions* moins riches **ou** dépourvues de ces produits, **grâce** aux *chemins* de fer, **aux** routes, à l'aviation **et** aux *canaux*.

Un pays **qui** manque de fer importe **son** minéral. **Vous** pouvez manger **du** poisson *de* mer, des **oranges**, dans *tout* le pays. **On** peut *même* provoquer des **chutes** de *pluie* en cas de sécheresse. **Aussi**, l'homme **civilisé** peut-*il* souvent, grâce à ses *inventions*, modifier la **nature** et se défendre contre **elle**.

L'homme, mes enfants, **ne** *subit* pas toujours la **nature** : souvent, *il* la contraint. **La** géographie *étudiée* intelligemment est une science **du** plus haut **intérêt** parce *qu'elle* est, **en** quelque *sorte*, un résumé de toutes les sciences.

Mais **vous** voyez *qu'elle* est **plus** encore. *N'est-elle* pas **l'illustration** vivante de **ce** que *peuvent* la volonté **et** l'**énergie** humaines ? *N'est-elle* pas, dans ses **grandes** lignes *et* ses petits **détails**, une *espèce* d'épopée **du** travail *humain* ? *N'est-elle* pas **le** déroulement de l'effort de l'homme, **de** sa *bataille* incessante **et** **inlassable** vers *de* belles **et** **précieuses** conquêtes ? *C'est* une **grande** **et** **vivante** leçon.

LE CINEMA

Le cinéma a, vous le savez, un peu plus de cinquante ans.

Depuis longtemps déjà, il est devenu parlant. Il a rendu de grands services pour la recherche scientifique et dans l'enseignement : films scientifiques, films documentaires, emploi du microfilm dans les bibliothèques et les dépôts d'archives, etc., mais c'est parce qu'il amuse qu'il est devenu, pour presque tout le monde, une habitude.

La radio lui fait concurrence, car, avec elle, on n'a même plus besoin de se déplacer pour en jouir. Elle s'est développée avec une rapidité inouïe. Le poste de radio est installé maintenant dans presque chaque foyer. Il a triomphé si vite parce qu'il offre à la fois les avantages du théâtre, du conservatoire et du journal.

Avec la télévision, il faut ajouter les agréments du cinéma.

La radio rend de grands services. Elle a augmenté la sécurité des marins et des aviateurs ; elle a permis à la météorologie de réaliser de grands progrès ; elle aide les armées en campagne, les explorateurs isolés dans la brousse ou sur la banquise. C'est ainsi que l'amiral Byrd, installé seul au pôle Sud, restait, par son poste de radio, en contact avec le monde civilisé.

La radio et la télévision renforcent, d'autre part, l'influence qu'exercent sur l'opinion publique le journal et le cinéma. C'est pourquoi l'action de ces moyens de suggestion modernes est énorme sur les masses.

Pour juger de son importance, il suffit de penser aux sommes énormes englouties chaque jour pour la publicité faite à la radio, à la télévision, au cinéma et dans les journaux. Si les hommes d'affaires n'hésitent pas à dépenser ainsi des millions, c'est qu'ils savent bien que c'est là de l'argent bien placé.

Il y a donc un danger de voir l'homme moderne devenir un type uniforme, qui s'habille, mange, vit et pense de la même façon sur toute la terre, perdant ainsi toute personnalité.

LE PETROLE

Le pétrole *brut*, tel qu'il est *extrait* de la *terre*, se *présente* sous la *forme* d'un liquide trouble, épais, de *couleur* généralement foncée, d'odeur *plutôt* désagréable.

On le découvre *dans* le sol, à des *profondeurs* diverses qui, *actuellement*, dépassent *parfois* les six mille mètres, alors que les premiers *forages*, *n'atteignaient* pas trente mètres.

Le *pétrole* brut provenant *des* puits est d'abord *emmagasiné* provisoirement *dans* d'immenses *réservoirs* métalliques *appelés* tanks. Ensuite, *de* puissantes *pompes* permettent l'*écoulement* du *pétrole* dans *de* longues conduites, *dans* de longs *tubes* en *acier*, que les *Anglais* nomment *pipe-lines*. Le liquide *noirâtre* est *ainsi* transporté vers *des* usines où se fera le raffinage du pétrole.

Cette *opération* industrielle *permet* d'obtenir *de* nombreux *produits*. Les uns sont gazeux, tels le butane livré en petites bonbonnes, bien connu des *ménagères* qui l'*utilisent* pour *préparer* les repas, et le *propane* qui est *fourni* aux *usines* pour chauffer les fours des brûleurs ou *faire* tourner *des* moteurs. *D'autres* sont *liquides*, comme l'*essence*, le *pétrole* du commerce, ou certaines *huiles* de graissage. *D'autres enfin* sont solides : la vaseline, la paraffine.

Pour *obtenir* une essence de haute *qualité* qui *permettra* aux avions *de* voler *plus* vite et *aux* autos *de* parcourir plus *de* kilomètres, les chimistes ont *ajouté* à l'essence un *produit* à base de plomb. *Cette* essence *contenant* du plomb *est* vendue *actuellement* sous le *nom* de *super*.

Un deuxième *procédé* consiste à chauffer l'*essence* ordinaire *sous* une forte *pression* et à haute température. *Cette* opération est appelée craquage *thermique*.

Il a fallu des *dizaines* d'*années* de patientes *recherches* et *de* travail acharné *de* la *part* des chercheurs *de* pétrole, *des* ingénieurs, *des* *chimistes*, *des* *marins*, *des* ouvriers, *pour* arriver au développement prodigieux que connaît l'industrie pétrolière *mondiale*.

Ce *liquide* est devenu si précieux *pour* le bien-être de *chacun* qu'on l'a surnommé l'*or* noir.

L'ART CONTEMPORAIN

On a dit parfois, on le dit encore, que l'art contemporain s'est heureusement libéré de toutes entraves.

Jugement sommaire à plusieurs égards.

L'art vit de contraintes, c'est un lieu commun de l'esthétique, Valéry, entre autres, l'a rappelé.

Il est évident que l'insignifiance du **sujet**, a fortiori l'abandon des modes traditionnels de la **figuration**, n'est en soi d'aucun secours pour le **peintre**, pas plus que l'accroissement des facilités techniques ne **garantit** une meilleure architecture. Par **contre**, que le désarroi provoqué par le **rejet** brutal des **règles** fondamentales de l'art issu de la Renaissance, règles **maintenues** jusqu'au seuil de **notre** époque, que ce désarroi ait permis à de nombreux **pseudo-artistes** de donner le **change**, c'est l'évidence même.

Loin donc d'être enviable, la situation de l'artiste contemporain est, en elle-même, peu confortable. Contraint d'inventer les premiers linéaments d'un art neuf, l'artiste doit multiplier les expériences, non sans utiliser ce que les **psychologues** appellent la méthode des **essais** et **erreurs**.

On se gardera cependant de n'attribuer qu'une valeur expérimentale et **prospective** à l'art, souvent appelé en bloc, d'avant-garde. S'il est vrai que les œuvres modernes ne se comprennent bien que par **groupes**, par *relation* d'accompagnement ou d'opposition, il n'en reste pas moins que la notion d'œuvre aboutie n'est nullement périmée.

Une autre attitude *superficielle* est également à rejeter : celle qui tend à ne voir au mieux dans l'art d'aujourd'hui, et particulièrement, dans la **peinture** dite *abstraite* que jeux de lignes agréables ou saisissants.

Beaucoup d'œuvres assurément ne dépassent pas le niveau du **chatouillement** optique, mais il en est assez dont la signification est autrement *profonde*.

ANNEXE II

Spécimen de test

CONSIGNES

Dans le texte suivant, on a supprimé un grand nombre de mots. Devinez-les.

A chaque lacune correspond un SEUL mot (exceptionnellement un nombre écrit en chiffres comme 1972 ou une abréviation comme S.V.P.)

Attention ! Tous les blancs ont la même longueur. Le mot manquant peut donc être ou très long (par exemple : *malheureusement*) ou très court (par exemple : *d'*).

Lisez d'abord l'ensemble du texte avant de commencer à écrire les mots que vous trouvez.

Ecrivez ensuite les mots dont vous êtes sûr.

Ecrivez pour finir ceux qui vous paraissent simplement probables.

Ne renoncez pas à écrire un mot parce que vous n'êtes pas certain de l'orthographe.

*
**

Nous allons faire un essai ensemble et le corriger immédiatement.

Après la correction de l'essai, attendez mon signal pour vous mettre au travail.

Remplissez d'abord le cadre « Identification ».

IDENTIFICATION

Commune :

Ecole :

Année d'études :

Nom :

Prénom :

Date de naissance :

Nationalité :

TEXTE D'ESSAI

La gare

Qu'elle jolie, la petite gare, ses volets vert clair, pots de fleurs aux vives, et sa façade !

Quand on pénétrait dans salle d'attente, on frappé par sa propreté ses vitres claires qu' rayon de soleil caressait.

CORRECTION DU TEXTE D'ESSAI

Qu'elle était jolie, la petite gare, avec ses volets vert clair, ses pots de fleurs aux couleurs vives, et sa façade blanche !

Quand on pénétrait dans la salle d'attente, on était frappé par sa propreté et ses vitres claires qu'un rayon de soleil caressait.

L'HOMME ET LA NATURE

Réservé
à la
correction

.....¹ homme, grâce à son² et sa
ténacité, a³ modifié la terre.

Sur⁴ sol ingrat et sur⁵ élé-
ments, il a remporté⁶ il remporte journalle-
ment des⁷. L'homme a transformé⁸
tapis végétal. Des terres⁹ sont devenues
fécondes. Il¹⁰ créé des paysages ; il
.....¹¹ construit ; il a détruit ;¹² a modi-
fié la nature.

.....¹³ constaterez que moins un¹⁴
est civilisé, plus il¹⁵ l'action de la¹⁶
parce qu'il n'¹⁷ pas découvert les moyens
.....¹⁸ lutter contre celle-ci¹⁹ de la
transformer.

Au²⁰, les peuples civilisés subissent
.....²¹ l'action de la²². Ils peuvent
transformer et²³ le sol, conquérir des
.....²⁴ sur la mer, construire²⁵ grands
barrages pour éviter²⁶ inondations et produire
de²⁷ électricité, transporter des produits
.....²⁸ des régions moins riches²⁹
dépourvues de ces produits,³⁰ aux chemins
de fer,³¹ routes, à l'aviation³² aux
canaux.

Un pays³³ manque de fer importe
.....³⁴ minéral. Vous pouvez manger³⁵
poisson de mer, des³⁶, dans tout le pays.
.....³⁷ peut même provoquer des³⁸ de
pluie en cas³⁹ sécheresse. Aussi, l'homme
.....⁴⁰ peut-il souvent, grâce⁴¹ ses
inventions, modifier la⁴² et se défendre contre
.....⁴³.

L'homme, mes enfants,⁴⁴ subit pas tou-
jours la⁴⁵ : souvent, il la contraint.⁴⁶
géographie étudiée intelligemment est⁴⁷

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42

science du plus haut ⁴⁸ parce qu'elle est,	43
..... ⁴⁹ quelque sorte, un résumé ⁵⁰	44
toutes les sciences.	45
	46
Mais ⁵¹ voyez qu'elle est ⁵²	47
encore. N'est-elle ⁵³ l'illustration vivante de	48
..... ⁵⁴ que peuvent la volonté ⁵⁵ l'éner-	49
gie humaines ? N' ⁵⁶ -elle pas, dans ses	50
..... ⁵⁷ lignes et ses petits ⁵⁸ , une espèce	51
d'épopée ⁵⁹ travail humain ? N'est- ⁶⁰	52
pas le déroulement de ⁶¹ effort de l'homme,	53
..... ⁶² sa bataille incessante et ⁶³ vers	54
de belles et ⁶⁴ conquêtes ? C'est une	55
..... ⁶⁵ et vivante leçon.	56
	57
	58
	59
	60
	61
	62
	63
	64
	65

LE RENNE

De la¹ famille que le cerf,²
 renne est plus bas³ en même temps plus
⁴ que lui. Ses jambes⁵ plus courtes
 et plus⁶, ses pieds sont plus⁷ et
 son poil plus⁸. Sa tête est ornée⁹
 bois superbes qui lui¹⁰ à se défendre.

Le¹¹ vit par troupes nombreuses,
¹² les régions glaciales des¹³ conti-
 nents. Il est surtout¹⁴ commun en Amérique
 où¹⁵ on lui donne le nom¹⁶ caribou.

Cet utile ruminant¹⁷ l'unique richesse
 des¹⁸ de l'extrême Nord¹⁹, sans
 lui, seraient inhabitées.

.....²⁰ lait procure une boisson²¹
 et fortifiante ; sa chair,²² nourriture substan-
 tielle ; sa peau,²³ excellentes fourrures et des
²⁴ très souples. Ainsi, le²⁵ donne,
 tout seul, ce²⁶ nous tirons du cheval,
²⁷ la vache et du²⁸.

Il n'est pas²⁹ à ses excréments
³⁰ on ne sèche pour brûler.

.....³¹ la terre est couverte³² neige,
 on attelle les³³ à des traîneaux avec
³⁴ on parcourt parfois vingt-cinq³⁵
 en un jour. Ces³⁶ sont, en effet, très
³⁷, et leurs sabots sont³⁸ conformés
 pour courir sur³⁹ sol neigeux sans
 s'.....⁴⁰ enfoncer.

Aussi doux que⁴¹, soumis à son maître,
⁴² à la peine, le⁴³ ne refuse jamais
 ses⁴⁴ et ne s'arrête⁴⁵ vaincu par
 la fatigue.

.....⁴⁶ sobriété est bien connue.⁴⁷
 se nourrit de toutes⁴⁸ de plantes et
 n'.....⁴⁹ pas difficile sur le⁵⁰ de ses
 mets.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50

LE PETROLE

Le pétrole¹, tel qu'il est² de la terre, se³ sous la forme d'⁴ liquide trouble, épais, de⁵ généralement foncée, d'odeur⁶ désagréable.

On le découvre⁷ le sol, à des⁸ diverses qui, actuellement, dépassent⁹ les six mille mètres, alors¹⁰ les premiers forages n'¹¹ pas trente mètres.

Le¹² brut provenant des puits¹³ d'abord emmagasiné provisoirement¹⁴ d'immenses réservoirs métalliques¹⁵ tanks. Ensuite, de puissantes¹⁶ permettent l'écoulement du¹⁷ dans de longues conduites,¹⁸ de longs tubes en¹⁹, que les Anglais nomment²⁰. Le liquide noirâtre est²¹ transporté vers les usines²² se fera le raffinage²³ pétrole.

Cette opération industrielle²⁴ d'obtenir de nombreux²⁵. Les uns sont gazeux,²⁶ le butane livré en²⁷ bonbonnes, bien connu des²⁸ qui l'utilisent pour²⁹ les repas, et le³⁰ qui est fourni aux³¹ pour chauffer les fours³² brûleurs ou faire tourner des moteurs. D'autres sont³³, comme l'essence, le³⁴ du commerce, ou certaines³⁵. D'autres³⁶ sont solides : la vaseline,³⁷ paraffine.

Pour obtenir une³⁸ de haute qualité qui³⁹ aux avions de voler⁴⁰ vite et aux autos⁴¹ parcourir plus de kilomètres,⁴² chimistes ont ajouté à⁴³ essence un produit à⁴⁴ de plomb. Cette essence⁴⁵ du plomb est vendue⁴⁶ sous le nom de⁴⁷.

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47

Un deuxième procédé consiste⁴⁸ chauffer l'essence ordinaire⁴⁹ une forte pression et⁵⁰ haute température. Cette opération⁵¹ appelée craquage thermique.

Il⁵² fallu des dizaines d'⁵³ de patientes recherches et⁵⁴ travail acharné de la⁵⁵ des chercheurs de pétrole,⁵⁶ ingénieurs, des chimistes, des⁵⁷, des ouvriers, pour arriver⁵⁸ développement prodigieux que connaît⁵⁹ industrie pétrolière mondiale.

Ce⁶⁰ est devenu si précieux⁶¹ le bien-être de⁶² qu'on l'a⁶³ l'or noir.

48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63

BIBLIOGRAPHIE

- BEST, B.J., *Use of Cloze Procedures in Remedial Work with Children Reading Below Grade Level*, Paper presented at the American Educational Research Association Annual Meeting, New York, Feb. 4-7, 1971.
- BICKLEY, A.C. et al., The Cloze Procedure : a Conspectus, in *Journal of Reading Behavior*, 2, 1970, 232-249.
- BLOOMER, R.J., The Cloze Procedure as a Remedial Reading Exercise, in *Journal of Developmental Reading*, 1962, 5, 173-181.
- BLOOMER, R.H., LOUTHAN, V., et HEITZMAN, A.J., *Non-Overt Reinforced Cloze Procedure*, U.S. Office of Education Project Report n° 2245, Storrs, University of Connecticut, 1966.
- BLOOMER, R.H., The Effects of Non-Overt Reinforced Cloze Procedure Upon Reading Comprehension, in G.B. SCHICK and M.M. MAY (Eds.), *Fifteenth Yearbook of the National Reading Conference*, Milwaukee, National Reading Conference, 1966, 31-40.
- BLOOMER, R.H., Cloze symposium, in G. SCHICK et M. MAY, *Multidisciplinary aspects of college-adult reading*, 17th Yearbook of the National Reading Conference, 1968.
- BORMUTH, J.R., Cloze as a Measure of Readability, in J.A. FIGUREL, Ed., *Reading as an Intellectual Activity*, New York, Scholastic Magazines, 1963.
- BORMUTH, J.R., Cloze Readability Procedures, in *Elementary English*, 45, 1968, 429-436.
- BORMUTH, J.R., et McDONALD, O.L., Cloze Tests as a Measure of Ability to Detect Literary Style, in J.A. FIGUREL (Ed.), *Reading and Inquiry*, Newark, Delaware, International Reading Association, 1965, 287-290.
- BORMUTH, J.R., Cloze Test Readability : Criterion Reference Scores, in *Journal of Educational Measurement*, 5, 1968, 189-196.
- BORMUTH, J.R., Comparable Cloze and Multiple-Choice Comprehension Test Scores, in *Journal of Reading*, 11, 1967, 291-299.
- BORMUTH, J.R., Comparisons Among Cloze Test Scoring Methods, in J.A. FIGUREL, Ed., *Reading and Inquiry*, Newark, Delaware, International Reading Association, 1965.
- BORMUTH, J.R., Experimental Applications of Cloze Tests, in *International Reading Association Proceedings*, 9, 1964, 303-306.
- BORMUTH, J.R., Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability, in *Reading Research Quarterly*, IV, n° 3, Spring 1969, 358-367.
- BORMUTH, J.R., *The Implications and Use of the Cloze Procedure in the Evaluation of International Programs*, Los Angeles, UCLA, Center of the Evaluation of Instructional Programs, Report n° 3, 1967.

- BORMUTH, J.R., Optimum Sample Size and Cloze Test Length in Readability Measurement, in *Journal of Educational Measurement*, 2, 1965, 111-116.
- BORMUTH, J.R., Validities of Grammatical and Semantic Classifications of Cloze Test Scores, in *Proceedings of the International Reading Association*, 10, 1965, 283-286.
- BORMUTH, J.R., *Development of Standards of Readability: Toward a Rational Criterion of Passage Performance*, Univ. of Chicago, 1971.
- BYRNE, M.A., FELDHUSEN, J.F., KANE, R.B., *The Relationships among Two Cloze Measurement Procedures and Divergent Thinking Abilities*, Rapport à la conférence annuelle de l'American Educational Research Association, Minneapolis, mars 1970.
- CARROLL, J.B. et al., *An Investigation of Cloze Items in the Measurement of Achievement in Foreign Language*, Cambridge, Mass., Harvard University, Laboratory for Research in Instruction, 1959, 138 p.
- CARVER, R.P., Analysis of « Chunked » Test Items as Measures of Reading and Listening Comprehension, in *Journal of Educational Measurement*, 1970, 7 (3), 141-150.
- COLEMAN, E.B. et BLUMENFIELD, J.P., Cloze Scores of Nominalizations and Their Grammatical Transformations Using Active Verbs, in *Psychological Reports*, 13, 1963, 651-654.
- COLEMAN, E.B., *Developing a Technology of Written Instruction: Some Determiners of the Complexity of Prose*, document ronéotypé, 56 p.
- CULHANE, J.W., Cloze Procedure and Comprehension, in *Reading Teacher*, 23, 1970, 410-413.
- DE LANDSHEERE, G., *Le test de closure*, Université de Liège, Laboratoire de Pédagogie expérimentale, 1968 (document ronéotypé).
- DE LANDSHEERE, G., Application du cloze test de W. Taylor à la langue française, in *Paedagogica Experimentalis*, II, 1972, 10-60
- DE LANDSHEERE, G., Du texte lacunaire au test de closure, in *Le moniteur des instituteurs*, 1972, 137-139.
- DICKENS, M. et WILLIAMS, F., An Experimental Application of Cloze Procedure and Attitude Measures to Listening Comprehension, *Speech Monographs*, 1964, 31, 103-108.
- ETIENNE, A., *Analyse des démarches et des types d'erreurs dans le test de closure*, Université de Liège, 1972 (mémoire inédit).
- FILLENBAUM, S. et JONES, L.V., An Application of Cloze Technique to the Study of Aphasic Speech, in *Journal of Abnormal and Social Psychology*, 1962, 65, 183-189.

- FILLENBAUM, S. *et al.*, The predictability of words and their grammatical classes as a function of rate of deletion from a speech transcript, in *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 186-194.
- FLETCHER, J.E., *A Study of the Relationships Between Ability to Use Context as an Aid in Reading*, University of Washington, 1959 (thèse de doctorat inédite).
- FRIEDMAN, M., The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at the University of Florida, in *Dissertation Abstracts*, 25, 1964, 3420-3421.
- GALLANT, R., Use of Cloze Test as a Measure of Readability in the Primary Grades, in J.A. FIGUREL, Ed., *Reading and Inquiry*, Newark, Delaware, International Reading Association, 1965, 286-287.
- GREENE, F.P., Modification of the Cloze Procedure and Changes in Reading Test Performances, in *Journal of Educational Measurement*, 1965, 2 (2), 213-217.
- GUICE, B.M., The Use of the Cloze Procedure for Improving Reading Comprehension of College Students, in *Journal of Reading Behavior*, 1969, 1 (3), 81-90.
- HAFNER, L.E., A One-Month Experiment in Teaching Context Aids in Fifth Grade, in *Journal of Educational Research*, 1965, 58, 472-474.
- HAFNER, L.E., Cloze Procedure, in *Journal of Reading*, 1966, 9, 415-421.
- HAFNER, L.E., Relationships of Various Measures to the Cloze, in E. THURSTONE and L. HAFNER, *Thirteenth Yearbook of the National Reading Conference*, Milwaukee, National Reading Conference, Inc., 1963, pp. 135-145.
- HAFNER, L.E., Using Context to Determine Meaning in High School and College, in *Journal of Reading*, 1966, 10, 491-498.
- HATER, M.A. et KANE, R.B., *The Cloze Procedure as a Measure of the Reading Comprehensibility and Difficulty of Mathematical English*, Purdue University (document ronéotypé), 13 p.
- HEITZMAN, A. et BLOOMER, R., The Effect of Non-Overt Reinforced Cloze Procedure Upon Reading Comprehension, in *Journal of Reading*, 1967, 11, 213-223.
- HENRY, G., *Construction de trois formules de lisibilité spécifiques à la langue française*, Université de Liège, 1973 (thèse de doctorat inédite).
- JENKINSON, M.E., *Selected Processes and Difficulties in Reading Comprehension*, Univ. of Chicago, 1957 (thèse de doctorat inédite).
- JONGSMA, E., *The Cloze Procedure as a Teaching Technique*, Newark, Del., Intern. Reading Assoc., 1971.
- KAMMANN, R., Verbal Complexity and Preferences in Poetry, in *Journal of Verbal Learning and Verbal Behavior*, 5, 1966, 536-540.
- KINGSTON, A.J. et WEAVER, W.W., *Feasibility of Cloze Technique for Teaching and Evaluating Culturally Disadvantaged Beginning Readers*, Research and Development Center, University of Georgia, 1969.

- KLARE, G.R., ROWE P.P., St JOHN, M.C. et STOLUROW, L.M. CLIP : A Computer Program for Cloze Instruction Procedure, Paper presented at New England Educational Research Organization Convention, Boston College, June 1969.
- KLARE, G.R., SINAIKO, H.W. et STOLUROW, L.M., *The Cloze Procedure : A Convenient Readability Test for Training Materials and Translations*, Arlington, Virginia, Institute for Defense Analyses, 1971.
- KOHLER, E.T., *A Factor Structure Study of Cloze Tests*, New York, National Council on Measurement in Education, 1967.
- KULM, G., *Measuring the Readability of Elementary Algebra Materials Using the Cloze Technique*. Paper presented at the American Educational Research Association Annual Meeting, New York, Feb. 4-7, 1971.
- LOUTHAN, V., Some systematic grammatical deletions and their effects on reading comprehension, in *English Journal*, 54, 1965, 295-299.
- Mc LEOD, J., *The GAP Reading Comprehension Tests*, London, Heinemann, 1965.
- MARLIER, M.C., *Etude factorielle d'épreuves de lecture au niveau 14-15 ans*, Université de Liège, 1970 (mémoire de licence inédit).
- MARTIN, R., *Transformational grammar, cloze and performance in college freshmen*, Syracuse Univ., 1968 (thèse doctorale inédite).
- MILLER, G.R. et COLEMAN, E.B., *A Set of 36 Passages Calibrated for Comprehensibility*, cité par E.B. COLEMAN, *Developing a Technology of Written Instruction : Some Determiners of the Complexity of Prose*.
- MOSBERG, L. et al., *The relation Between Cloze and Multiple-Choice Test Scores as a Function of Relative Paragraph Difficulty and Grade Level*, Inglewood, Calif., Southwest Regional Laboratory for Educational Research and Development, Technical report 6, 1968.
- OHNMACHT, F.W., WEAVER, W.W. et KOHLER, E.T., Cloze and Closure : A Factorial Study, in *The Journal of Psychology*, 1970, 74, 205-217.
- POTTER, C., *A Taxonomy of Cloze Research*, Inglewood, Calif., Southwest Regional Laboratory for Educational Research and Development, Technical Report 1, 1968.
- RAMANAUSKAS, S., Oral reading errors, and cloze, in J. CAWLEY et al., *The slow learner and the reading problem*, Springfield, Ill., Chas. Thomas, 1972.
- RAMANAUSKAS, S., The responsiveness of cloze readability measures to linguistic variables operating over segments of text longer than a sentence, in *Reading Research Quarterly*, VIII, 1, 1972, 72-91.
- RANKIN, E.F., Cloze Procedure. A Survey of Research, in E.L. THURSTON and L.E. HAFNER (Eds.), *The Philosophical and Sociological Bases of Reading, Fourteenth Yearbook of the National Reading Conference*, Milwaukee, Wisconsin : The National Reading Conference, Inc., 1965, pp. 133-150.

- RANKIN, E.F. et DALE, L., Cloze Residual Gain, A Technique for Measuring Learning Through Reading, in G.B. SHICK and M.M. MAY (Eds.), *The Psychology of Reading Behavior, Eighteenth Yearbook of the National Reading Conference*, Milwaukee, National Reading Conference, Inc., 1969.
- RANKIN, E.F. et CULHANE, J.W., Comparable Cloze and Multiple-Choice Comprehension Test Score, in *Journal of Reading*, 13 (3), 1969, 193-198.
- RANKIN, E.F., Uses of the Cloze Procedure in the Reading Clinic, in J.A. FIGUREL (Ed.), *Reading in a Changing Society*, New York, Scholastic Magazines, 1959 (4), 228-232.
- RANSOM, P.E., Determining Reading Levels of Elementary School Children by Cloze Testing, in J.A. FIGUREL (Ed.), *Forging Ahead in Reading*, Newark, Delaware : International Reading Association, 1968, 477-482.
- ROSENSHINE, B., New Correlates of Readability and Listenability, in J.A. FIGUREL (Ed.), *Reading and Realism*, Newark, Delaware : International Reading Association, 1969, 13 (1), 710-716.
- SCHNEYER, J.W., Use of the Cloze Procedure for Improving Reading Comprehension, in *The Reading Teacher*, 1966, 19 (3), 174-179.
- TAYLOR, W.L., Cloze Procedure : A New Tool for Measuring Readability, in *Journalism Quarterly*, 30, 1953, 415-433.
- TAYLOR, W.L., Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude, in *Journal of Applied Psychology*, 41, 1957, 19-26.
- TAYLOR, W.L., Recent Developments in the Use of Cloze Procedure, in *Journalism Quarterly*, 33, 1956, 42-46.
- WEAVER, W. et KINGSTON, A., A Factor Analysis of the Cloze Procedure and Other Measures of Reading and Language Ability, in *Journal of Communication*, 13, 1963, 252-261.
- WEINTRAUB, S., Cloze Procedure, in *Reading Teacher*, 21, 1968, 567-571.

TABLE DES MATIERES

INTRODUCTION	7
Première partie : LE TEST DE CLOSURE	9
1. Définition et histoire	11
2. Description du test	13
I. Les lacunes	13
II. Règles pour la suppression des mots	14
III. Consignes	15
IV. Correction	15
V. Scores	15
VI. Etalonnage	18
VII. Echantillonnage des textes	18
VIII. Fidélité	18
IX. Champ d'application	21
A. Mesure de l'intelligibilité	21
Lisibilité	21
Audibilité	21
Audio-visibilité	21
B. Mesure de la capacité en lecture	21
Compréhension générale	21
Compréhension spécifique	21
C. Instrument pronostique	22
D. Instrument diagnostique	23
E. Recherche linguistique	23
F. Test d'intelligence verbale	24
G. Technique d'enseignement	25
3. Validité de la technique	29
I. Aspect théorique	29
A. Mesure de la capacité en lecture	30
B. Mesure de la lisibilité	36
II. Aspect expérimental	38
A. Mesure de la capacité en lecture	38
1° Corrélation avec des tests expérimentaux	38
2° Corrélation avec des tests standardisés	40
Mesure de la lisibilité	41

4. Exemples d'application	43
I. Mesure de la compréhension de textes scientifiques	43
II. Mesure de la compréhension de poèmes	43
III. Mesure de la compétence en langue étrangère	44
IV. Mesure de la difficulté de différentes classes de mots	45
Deuxième partie : EXPERIENCE	47
1. Choix des textes	49
2. Schéma expérimental	56
3. Population expérimentale	58
4. Administration des tests	59
5. Correction des épreuves et traitement des données	60
6. Résultats : enseignement secondaire	62
I. Enseignements général et technique séparés, formes 1 et 2 séparées	62
A. Scores moyens et graphiques	62
B. Constatations générales	70
C. Comparaison des formes 1 et 2	76
D. Les dispersions	78
II. Enseignements général et technique séparés, formes 1 et 2 réunies	81
III. Les deux enseignements et les deux formes réunis	84
IV. Comparaison avec les scores de lisibilité de R. Flesch	87
V. Fidélité	90
7. Résultats : enseignement primaire	91
I. Belges et étrangers séparés, formes 1 et 2 séparées	91
A. Scores moyens et graphiques	91
B. Constatations générales	95
C. Comparaison des formes 1 et 2	98
D. Les dispersions	98
II. Groupes et formes réunis	100
III. Comparaison avec les scores de lisibilité de R. Flesch	100
IV. L'enseignement primaire face au secondaire et au technique	102
CONCLUSION	104
ANNEXES	105
I. Les six textes choisis	107
II. Spécimen de test	113
Consignes - Identification - Essai - Test	113
BIBLIOGRAPHIE	121

ACHEVE D'IMPRIMER
LE 5 JUIN 1978

Imprimé en Belgique

D/1978/258/19

ISBN 2-8259-0063-X.