



ASp
La revue du GERAS

34 | 2001
Actes de la journée d'étude sur l'évaluation

Sens, forme, ne manque-t-il pas un troisième larron ?

Jean-Paul Narcy-Combes



Édition électronique

URL : <https://journals.openedition.org/asp/1695>
DOI : 10.4000/asp.1695
ISBN : 978-2-8218-0386-2
ISSN : 2108-6354

Éditeur

Groupe d'étude et de recherche en anglais de spécialité

Édition imprimée

Date de publication : 1 décembre 2001
Pagination : 3-12
ISSN : 1246-8185

Référence électronique

Jean-Paul Narcy-Combes, « Sens, forme, ne manque-t-il pas un troisième larron ? », *ASp* [En ligne], 34 | 2001, mis en ligne le 21 septembre 2010, consulté le 23 novembre 2024. URL : <http://journals.openedition.org/asp/1695> ; DOI : <https://doi.org/10.4000/asp.1695>

Ce document a été généré automatiquement le 23 novembre 2024.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

Sens, forme, ne manque-t-il pas un troisième larron ?

Jean-Paul Narcy-Combes

- 1 Parce qu'elle souhaite baliser le terrain, cette contribution sera différente de celles qui la suivront. Son objectif est double : sensibiliser chacun au sens de la démarche qui a justifié l'organisation de cette journée d'étude, et poser une question pour laquelle je cherche une réponse depuis des années comme l'ont montré mes participations au Livre blanc de l'UPLEGESS et au projet DCL. Ces participations m'ont laissé le souvenir de deux entreprises exemplaires de réflexion scientifique, tant au niveau de la rigueur que de l'esprit de travail collectif qui y régnait.
- 2 Cette journée d'étude répond à une demande exprimée lors du dernier colloque du GERAS à Bordeaux. Cette demande est bien sûr liée à la mise en place du CLES, ce qui justifie partiellement la présence de Claude Springer parmi nous, mais nous sommes un groupe de chercheurs et, pour faire face à une demande institutionnelle, il nous a semblé logique de nous réunir avec les chercheurs du domaine qui nous semblaient les plus impliqués dans une réflexion sur l'évaluation, ce qui justifie complètement la présence de Claude Springer, et selon moi, de façon plus légitime que son rôle institutionnel.
- 3 Je partagerai mes propos en deux parties inégales. L'une sera consacrée à cette journée d'étude, sa raison d'être, le jeu des représentations qui conditionnent la mise en place de tout système d'évaluation dans ce pays et la légitimité scientifique des intervenants. L'autre sera consacrée au troisième larron que le titre mentionne et à relancer une recherche sur ce troisième larron.
- 4 Je laisserai à mes collègues le soin de vous apporter les références bibliographiques spécifiques au domaine de l'évaluation et ne mentionnerai que celles qui m'ont aidé à clarifier mon propos. La pauvreté de ce propos est bien le résultat des limites de la bibliographie. La question que je pose ne me semble pas avoir suscité beaucoup de recherche. Il y a là un vaste champ de réflexion ouvert à l'étude scientifique et j'espère que la mise en place du CLES suscitera des vocations.

1. Les objectifs de cette journée d'étude

- 5 Nous verrons dans le point deux ce qui justifie le décalage entre les attentes que la demande formulée à Bordeaux révélait et ce que nous avons décidé de faire. Contentons-nous ici d'aborder les objectifs des uns et des autres. Ce qui nous permettra d'évaluer à chaud du moins, dans quelle mesure ces objectifs ont été atteints.
- 6 Nous travaillons dans un contexte institutionnel français. Une décision descendante impose un système d'évaluation nouveau, national, pour le secteur LANSAD. Il nous faut donc répondre. La réponse habituelle (au sens sociologique du mot habitude, voire du mot *habitus*) consiste à comprendre les instructions et à mettre en place des épreuves en fonction de recettes méthodologiques. Or, ce n'est pas par hasard que la demande a été formulée au GERAS, à des chercheurs. Le secteur LANSAD ne reflète pas exactement la tradition française, il l'interpelle fréquemment. Dans le cas présent, les collègues du terrain se demandent comment appliquer les instructions sans réflexion préalable, d'autant qu'au niveau des épreuves, le texte du CLES, loin d'être aussi explicite que les textes du DCL, laisse une marge d'initiative importante.
- 7 Nous avons donc fait l'hypothèse que, pour répondre à la demande, il fallait, dans un premier temps, faire le point sur l'état de l'art de l'évaluation dans notre domaine et mieux comprendre le fonctionnement du CLES, et ceci explique le contenu de cette journée. Nous n'échangerons pas de recettes, mais personne n'en avait demandé ! D'autres colloques suivront pour compléter la réflexion. Ils aborderont sans doute les tâches d'évaluation, mais nous n'en sommes pas encore à ce stade au début de la journée.

2. Le jeu des représentations

- 8 L'évaluation est une forme de dialogue que nous entretenons avec les apprenants, les institutions et la société, et ceci est d'autant plus vrai lorsque cette évaluation est certificatrice.
- 9 En théorie, les enseignants de langues sont les experts. En fonction d'objectifs déterminés, ils établissent des tests de langues qui permettent de certifier de l'atteinte d'un niveau de performance. Dans notre pays, les chercheurs en interculturel, Hofstede par exemple, avancent que l'expertise est souvent par attribution. Le poste que vous occupez vous donne l'expertise du domaine dont vous êtes en charge. On oppose à cela l'expertise par rétribution qui provient d'une formation attestée dans le domaine. Dans les pays où l'expertise se fait par rétribution, les décideurs font appel à des consultants experts, qui leur donnent toutes les solutions possibles, avec leurs avantages et leurs inconvénients, les décideurs prennent alors leur décision en toute connaissance de cause. Le DCL a été créé de cette façon, mais sa gestion ultérieure n'a pas nécessairement su éviter le retour à un fonctionnement plus typiquement français, que l'organisme interuniversitaire d'évaluation qui avait été proposé aurait évité.
- 10 Le CLES est trop peu connu pour que l'on puisse encore se prononcer, il convient néanmoins, que nous mesurions les avantages et les risques qu'il apporte.
- 11 En l'absence d'expertise attestée, c'est l'opinion qui l'emporte, et nous savons, depuis Bachelard, comme elle est mauvaise conseillère. Le public en général, et le public étudiant en particulier, a souvent des représentations sur l'apprentissage des langues

qui ne correspondent pas aux positions des experts. Par exemple, il existe, en France, une plus grande peur de la faute en expression orale qu'en expression écrite. Ce qui n'est pas logique si l'on se rappelle que, dans les faits, l'interlocuteur étranger sera moins gêné par une erreur à l'oral qu'à l'écrit. L'erreur, à l'écrit, est rémanente et l'énonciateur n'est plus là pour la corriger au moindre signe de son interlocuteur. Le souci de perfection des Français est donc souvent contre-productif.

- 12 Ensuite, en ce qui concerne les épreuves, certaines, moins pertinentes, sont plus sécurisantes ou apparemment plus valorisantes, pour des raisons sociologiques ou psychologiques ; c'est le cas des prestigieux TOEFL et TOEIC. Les intervenants nous parleront plus tard de l'effet de retour. Un test valorisant peut avoir des effets négatifs sur l'apprentissage et donc sur la performance dans la vie réelle. Que signifie un score élevé au TOEFL ou au TOEIC, si l'individu ne peut communiquer oralement et que c'est là son besoin ?
- 13 Aller à l'encontre des représentations est une tâche délicate, seule une expertise assurée permet de le faire. De plus, c'est un travail de longue haleine.
- 14 Les institutions, le Ministère, universités et employeurs véhiculent également des représentations en fonction d'attentes sociales ou économiques qui ne faciliteront pas le dialogue avec les experts. De plus, il existe fréquemment entre ces institutions et les experts des relations hiérarchiques, qui, en France, compliquent la collaboration, puisque le décideur se voit attribuer une expertise par attribution. C'est là où il est important que les enseignants du domaine LANSAD se sensibilisent encore plus à l'importance de la recherche, qui non seulement accroît leur expertise, mais encore leur permet de s'élever dans la hiérarchie universitaire. C'est là une loi culturelle qui a sa sévérité, mais aussi ses récompenses. Rappelons néanmoins que les outils de certification les moins coûteux ne sont pas nécessairement ceux que les experts conseilleraient.
- 15 En ce qui concerne les professionnels, nous et nos collègues, il nous faut admettre que ce milieu également véhicule nombre de représentations et de pratiques que les experts n'approuveraient pas. De plus, l'origine même des enseignants est source de divergences, selon qu'ils sont locuteurs natifs ou non. Un des problèmes majeurs pour moi se situe au niveau des attentes que nous avons. Sont-elles fondées sur des études scientifiques dont la validité devra constamment être vérifiée (interlangue par exemple), ou sont-elles uniquement le fruit de nos représentations ?
- 16 Notre tâche en tant que participant à ce Colloque est énorme, il importe que nous acquérions l'expertise voulue, que nous nous affirmions statutairement pour que notre expertise soit reconnue, que nous négocions des changements avec nos collègues et que nous facilitions un changement de représentations chez les étudiants. Enfin, apparemment, nous devons être en mesure de créer collectivement de nouvelles tâches évaluatives !
- 17 C'est pour apporter un début de réponse que nous avons organisé cette journée d'étude comme nous l'avons fait. En ce qui concerne l'organisation matérielle, il convient de remercier Josiane Hay et Marie-Hélène Fries qui renouvellent avec brio ce qu'elles avaient réalisé pour la Recherche Action.

3. La légitimité des intervenants

- 18 C'est parce qu'il me paraît important de distinguer entre ce qui relève de l'opinion, et ce qui relève d'une réflexion scientifique que je tiens à souligner la légitimité des intervenants à ce colloque.
- 19 Légitimité dans le parcours scientifique : deux ont soutenu une thèse sur l'évaluation en langues (S. Sommer et P. Doucet). Les deux autres sont docteurs en didactique et ont œuvré ou œuvrent encore dans des recherches concernant l'évaluation, et publient leurs résultats dans des revues à comité de lecture.
- 20 Légitimité dans le parcours professionnel : issus du terrain LANSAD, ils ont gravi les échelons et ont atteint un statut qui leur donne une légitimité sociale plus facile à affirmer. Ils ont confirmé leur expertise par un travail de consultant dans divers organismes d'évaluation (DCL, Cambridge entre autres), et ils correspondent donc à ce concept de consultant expert dont notre domaine a besoin pour asseoir sa légitimité sociale.
- 21 Cette expertise, vous le savez, et ils vous le rappelleront, ne leur confère aucune autorité, mais elle nous permettra d'agir et de collaborer en meilleure connaissance de cause. Les différents contextes influent trop sur les décisions à prendre dans notre domaine pour que quiconque puisse proposer des solutions sans un dialogue préalable avec les intervenants. Et espérons que cette journée d'étude, et les colloques qui la suivront nous permettront un débat fructueux.

4 . À la recherche du troisième larron

- 22 C'est en m'inspirant d'écrits antérieurs que j'ai conçu le titre de mon intervention. L'effet d'appel m'a plus intéressé que la justesse de mon propos. Il y a, en effet, plusieurs trinités qui peuvent attirer notre attention :
1. sens, forme et contenu (relève d'une approche globale de la didactique),
 2. *accuracy*, *complexity*, *fluency* (relève plus, selon moi d'une approche linguistique de la didactique),
 3. contenu/savoir, pragmatique, linguistique (relève en partie de la didactique, mais de bien d'autres domaines également, dont la linguistique et la didactique des différentes disciplines).
- 23 Dans son intervention, Claire Bourguignon s'interroge sur la prise en compte de deux axes, un axe pragmatique, et un axe linguistique, et ses réflexions sont d'une extrême pertinence, mais elles ne me semblent pas répondre à toutes mes questions.
- 24 Si, globalement, on peut dire que la pragmatique étudie les effets que nous souhaitons créer, ne pouvons-nous, néanmoins, nous demander si on ne communique que pour créer des effets ?
- 25 Avec quoi créons-nous ces effets et pourquoi faire ? N'y a-t-il que du linguistique, ou ne manipulons-nous pas du savoir (scientifique, professionnel, technique, artistique, etc.), et/ou des idées ou des sentiments ?
- 26 Si nous manipulons du savoir, dans quelle mesure cela influe-t-il sur notre performance ?

- 27 J'ai souvent l'impression que l'opérationnalité (cf. DCL, par exemple) est souvent vécue par notre profession comme :

Utiliser une langue donnée (*fluency*) suivant les conventions d'une communauté donnée (*accuracy* et pertinence pragmatique) pour accomplir des tâches qui sont sans autre conséquence que celle d'être des tâches à des fins d'acquérir une langue et/ou d'en mesurer l'acquisition.

Je serai tenté d'avancer une définition différente :

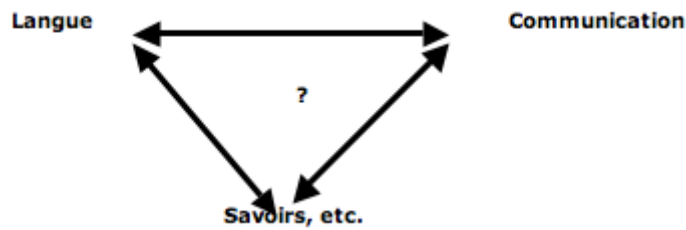
Accomplir des tâches sociales relevant d'un bagage de savoirs (**didactique des disciplines**) en négociant du sens suivant les conventions d'une communauté donnée (**pragmatique**) en utilisant la langue (une des langues) de cette communauté (**linguistique**).

- 28 Cette définition n'est pas sans conséquence sur notre propos. La langue n'y occupe pas la première place, la pragmatique y garde un rôle de médiation, mais c'est bien l'interaction sociale fondée sur des échanges à fins scientifiques, professionnelles ou personnelles qui prime et qui conditionne la production langagière (*output*). Ce renversement des termes ne serait qu'une opération gratuite s'il n'était justifié par certains résultats de recherche.

5. Justification de la recherche d'un troisième larron

- 29 Un des problèmes de cette recherche est qu'il n'y a que des pistes. Le lien entre langue et cognition n'est pas étudié sur le plan des savoirs, mais plutôt de la gestion mentale des concepts.
- 30 Certaines études nous sensibilisent néanmoins aux risques qu'il y aurait à négliger les contenus.
- 31 Pour Dortier (1999), les réseaux sémantiques – *frames* et *scripts* – ne fonctionnent bien que dans un domaine défini. Un changement de domaine perturbera donc sensiblement la production langagière. Des tests comme le DCL peuvent ainsi favoriser certains individus en fonction des thèmes abordés, il faudrait au moins nous assurer de l'acceptabilité de cette variation.
- 32 Selon Alice Eisele Henderson (1997), il existe un seuil de cohésion qui rend la lecture en L2 difficile, même dans un domaine dont on est spécialiste. Certains autres chercheurs ont noté en production ce qu'ils appellent un *proficiency gap* qui handicape l'acquisition du savoir en L2 dans les enseignements bilingues.
- 33 Enfin, Skehan *et al.* ont noté que la familiarité du contenu accroît la fluidité de la performance, mais pas la correction, ni la complexité de celle-ci.
- 34 Toutes ces pistes sont des plus brouillées quand il s'agit d'études sur l'anglais en raison de l'universalité de son emploi et de sa transparence pour des locuteurs européens, en particulier en lecture. Des critères mis en place sur des épreuves en anglais fonctionneront-ils efficacement sur des épreuves dans d'autres langues ?
- 35 Il importe donc que, dans un test, nous sachions qui dessert quoi des éléments du triangle représenté à la figure 1.

Figure 1. Triangle



- 36 Peut-on admettre, par exemple, que l'indigence du contenu scientifique soit compensée par la pertinence pragmatique ou la qualité purement linguistique de la production ?
- 37 Qui doit évaluer quoi dans des épreuves conçues pour des scientifiques de haut niveau ?
Pouvons-nous nous passer d'un chercheur du domaine de ces scientifiques ?
- 38 Mes interrogations ne sont pour le moment que cela, nous verrons en annexe quelques pistes issues des travaux du groupe UPLEGESS et du DCL. Elles confirment néanmoins que des recherches s'imposeraient, tant au niveau du lien entre le contenu et la performance, que de celui de la conception d'épreuves qui respectent plus authentiquement l'interaction sociale selon les conventions du domaine des spécialistes dont la performance en L2 est évaluée. Cette communication est donc également un appel.

6. Quelques pistes pour la prise en compte de ce troisième larron

- 39 Le travail de réflexion du groupe UPLEGESS, qui a conduit à un *Livre blanc* en 1988, se faisait déjà en essayant de prendre en compte ce que les grilles appelaient la « **réalisation** ». Par réalisation, il fallait entendre celle de l'acte social en jeu (ce qui impliquait l'expression langagière d'éléments de savoir, etc.).
- 40 La grille présentée en annexe 1 montre comment, pour l'épreuve considérée, on fonctionnait. Un texte était choisi ; on recherchait le nombre de thèmes qui y étaient traités (dix ici). Les grilles de références nous indiquaient que le niveau 1 était atteint si un apprenant réalisait un score de 75 % environ des points de réalisation (ou un mélange de points de réalisation et de points de correction équivalent à 75 %). Le niveau 2, lui, se justifiait par environ 75 % de l'ensemble des points de réalisation et de correction. Les niveaux 3 et 4 demandaient que l'on prenne en compte d'autres critères moins facilement mesurables que présence/absence d'un thème, ou correction linguistique d'un énoncé/présence d'une seule erreur ou de plusieurs.
- 41 Le lecteur peut ainsi observer que c'est le contenu qui déterminait la mesure de la performance langagière. Ce qui imposait un découpage fin de ce contenu qui a posé problème lorsque, dans une épreuve comme celle du DCL, le contenu est devenu complexe.
- 42 J.-P. Nougier, a essayé de résoudre le problème, en prenant en compte ce que certains avaient appelé la notion d'unité thématique. L'objectif était bien de s'assurer que c'était la présence, ou l'absence, des unités thématiques qui décidaient des autres mesures.
- 43 La grille correspond bien à la grille de références du DCL (évolution de l'interlangue en colonne et niveau d'opérationnalité en ligne). Pour chaque unité thématique, chaque

case avait l'avantage de se remplir en terme de présence/absence, ce qui en termes de Mager (1986) était l'idéal (réduction des différences entre correcteurs).

- 44 Cette grille était néanmoins trop innovatrice et, tout en augmentant la validité du test, exigeait trop des correcteurs, pour être acceptée à un moment où les impératifs d'opérationnalisation du test s'imposaient avec acuité. Il reste à reprendre le travail de J.-P. Nougier et je souhaite vivement que cela soit fait. Lui-même avait repris avec pertinence les résultats des travaux du groupe UPLEGESS.
- 45 Le temps m'a manqué à Grenoble pour aborder ces points qui complètent mon questionnement en montrant comment il peut être pris en compte et c'est bien le rôle des Actes d'un colloque de compléter ce qui n'a pu être qu'effleuré. J'espère qu'à Chambéry nous irons, plus nombreux encore, plus loin, et que nous ébaucherons l'organisation d'un groupe de recherche au sein du GERAS. C'est dans ce type de collaboration que nous trouverons l'heureuse combinaison d'expertise et de créativité qui conduit à la conception d'épreuves innovantes, légitimes scientifiquement et pertinentes dans leur fonction sociale. Merci donc aux collègues de Chambéry qui prendront la relève de l'organisation !

BIBLIOGRAPHIE

Dortier, J. F. 1999. *Le Cerveau et la pensée*. Auxerre : Sciences Humaines Éditions.

Duverger, J. (dir.). 2000. « Actualités de l'enseignement bilingue ». *Le français dans le Monde*, numéro spécial Janvier.

Eisele-Henderson, A. 2000. « La lecture des textes en anglais langue étrangère chez des étudiants en sociologie : implications didactiques d'une perspective socio-cognitive ». Thèse de doctorat, Université de Savoie.

Ellis, R. 1997. *SLA Research and Language Teaching*. Oxford : Oxford University Press.

Gajo, L. et C. Serra. 2000. « L'enseignement bilingue des disciplines : à l'interface de la construction cognitive des savoirs et des langues ». In Actes du 3^e Colloque International de la COFDELA, Avignon : Université d'Avignon.

Hofstede, G. 1994. *Vivre dans un monde multiculturel – Comprendre nos programmations mentales*. Paris : Les Éditions d'Organisation.

Holliday, A. 1999. « Evaluating the discourse: the role of applied linguistics in the management of evaluation and innovation ». In P. Rea-Dickins and K.P. Germaine (dir.), *Managing Evaluation and Innovation in Language teaching*. Londres : Longman, 195-219.

Mager, R. F. 1986. *Comment mesurer les résultats de l'enseignement*. Paris : Bordas,.

Meyer, F. (dir.). 2001. *Language for Special Purposes: Perspectives for the New Millenium*, Tübingen : GNV.

Narcy, J.-P. 1990. *Apprendre une langue étrangère*. Paris : Éditions d'Organisation.

Skehan, P., M. Bygate and M. Swain (dir.). 2001. *Researching Pedagogic Tasks, Second Language learning, Teaching and Testing*. Londres : Longman.

UPLEGESS (dir.). 1988. *L'évaluation des compétences en fin de cursus*. Compiègne : Université de Technologie.

ANNEXES

Annexe 1

Grille d'évaluation d'un condensé de texte

Note affectée d'un coefficient :

Nom :
Prénom :

ÉPREUVE : CONDENSÉ (GRILLE D'ÉVALUATION)

Se référer au texte pour y retrouver les dix points principaux.

APRÈS UNE PREMIÈRE LECTURE DU CONDENSÉ, NOTER GLOBALEMENT SUR 4, CHACUNE DES 3 RUBRIQUES CI-DESSOUS.

Registre :	Richesse lexicale et morphosyntaxique :	
Pertinence de l'argumentation :		
Thèmes	Efficacité de réalisation	Correction
1 - Bell's connection with research.....		
2 - Past pbs with Govt.....		
3 - Consequences of changes in gov't attit.....		
4 - Previous gov't att did not stop innov.....		
5 - 2 aspects in the dev't of tel serv.....		
6 - Products an. serv soon available.....		
7 - Bell's main attitude to computers.....		
8 - Is Bell Ltd as far as innov is conc.....		
9 - New tendencies regarding sizes.....		
10 - Explanation of Bell's successes.....		
TOTAL		

Colonne réalisation : cocher si la réponse vous paraît efficace sur le plan de la communication.
Colonne correction : cocher si la réponse est totalement correcte sur le plan de la langue.
Additionner les points des deux colonnes, additionner toutes les notes pour obtenir le total général.

TOTAL GÉNÉRAL

Annexe 2

Grille d'évaluation du DCL de J.-P. Nougier

Grille de correction Phase IV Production écrite	Unités thématiques sur lesquelles s'appuient les tâches de communication d'un scénario donné	Critères d'é				
		Éléments restitués, classés ou organisés, et présentés selon les attentes à chaque degré	Fortes traces de L 1		Traces de L	
			Lesquels majoritairement restitués du contexte et limité	Syntaxe de base	Appropriation des éléments lexicaux	Syntaxe de L2, sans erreurs systématiques
Niveau d'opérationnalité			Pas de flexion, pas d'éléments cohésifs	Paraphrases pour combler les lacunes	Déficit de flexion, absence d'éléments cohésifs	
Restitution intelligible des éléments essentiels repérés		D 0				
Situation prévisible						
Pas d'autonomie					D 1	
Organisation des éléments fournis dans un format donné						

AUTEUR

JEAN-PAUL NARCY-COMBES

Jean-Paul Narcy-Combes est Professeur à l'Université de La Rochelle. Ses recherches portent sur les stratégies d'apprentissage et la théorisation de la pratique, l'épistémologie et la méthodologie de la recherche, en lien avec les TIC. Il a coordonné les travaux qui ont conduit au Livre blanc sur l'évaluation en Langues (UPLEGESS) et a été consultant scientifique du projet DCL. jean-paul.narcy-combes@wanadoo.fr