

Le discours électronique médié : bilan et perspectives.....	2
Résumé.....	2
1. Contexte et terminologie.....	2
1.1. Médiation vs. médiatisation.....	3
1.2. Choix de recherche.....	3
2. Caractéristiques du discours électronique médié.....	4
2.1. Les marqueurs linguistiques et extra-linguistiques de l'oral.....	4
2.2. Les marques écrites omniprésentes.....	5
2.3. Les spécificités.....	5
3. Étude de cas.....	7
3.1. Présentation du corpus.....	7
3.2. Commentaires & résultats.....	8
3.2.1. Le forum vs. le chat.....	8
3.2.2 Traits spécifiques du DEM.....	8
3.2.3 Relations DEM/écrit/oral.....	9
3.2.4. Comparaison évolutive et nouveautés.....	10
3.2.5. Usage du Texto.....	10
4. Conclusion.....	11
Références :.....	12

Le discours électronique médié : bilan et perspectives.

Rachel Panckhurst
Praxiling ICAR UMR 5191 CNRS
Université Montpellier 3,
Route de Mende
341799 Montpellier cedex 5
France
+33 4 67 14 58 61
rachel.panckhurst@univ-montp3.fr

Résumé

Quand l'ordinateur est utilisé pour le courriel, les forums de discussion et les chats, en tant qu'outil permettant la communication entre individus, il devient un véritable *médiateur* ; son utilisation modifie notre discours et ainsi notre façon de communiquer avec autrui. Émerge alors un nouveau « genre de discours », le « discours électronique médié (DEM) ». Le DEM contient des marques linguistiques et extra-linguistiques qui lui sont propres et il entre dans le cadre plus global de la « communication médiée par ordinateur » (CMO). Nous continuons à préférer cette appellation — qui se rapproche à CMC (*Computer-Mediated Communication*) en anglais (Herring, 1996) — plutôt que « communication (écrite) médiatisée par ordinateur » (Marcoccia, 2000) ou « Netspeak » (Crystal, 2001), ou encore « nouvelles formes de communication écrite » (Guimier de Neef & Véronis, 2004), « communication électronique scripturale » (Anis, 2002), « communication électronique » (Anis, de Fornel, Fraenkel, 2004), etc.

Suite à différentes recherches portant sur le courriel, le forum et le chat (Panckhurst, 1999a in Anis, 1999, Panckhurst, 1999b, 2001, 2003b, Panckhurst & Bouguerra, 2003a), nous proposons ici une étude du discours contenu dans les forums de groupes d'étudiants dans le cadre de deux enseignements : (1) une initiation à l'informatique, qui constitue une préparation au certificat national C2i (*Certificat Informatique et Internet*), délivré en L1, L2 et L3, et, (2) un cours portant sur l'évaluation de logiciels en formation ouverte et à distance (FOAD) en Master, niveau 2 (Gestion des apprentissages en FOAD) à l'université Paul-Valéry, Montpellier. Les étudiants devaient échanger sur des forums limités à une communication entre 3 à 5 participants dans le cadre d'un travail à effectuer en groupe ; l'enseignant était inclus comme participant au forum, mais n'est intervenu que rarement. Notre analyse porte donc sur l'appropriation de l'outil et l'évolution du discours dans un cadre pédagogique dans lequel les étudiants se sentaient en situation quasi-privée.

Dans ce chapitre, nous exposons en premier lieu le contexte de notre recherche ainsi que nous choisissons des terminologies. Les marques linguistiques et extra-linguistiques que nous avons repérées depuis le début de nos recherches sont ensuite énumérées. Les résultats d'un corpus d'étude élaboré en 2005, dans un contexte d'enseignement supérieur français, sont également présentés, avant une comparaison finale avec d'autres corpus antérieurs.

1. Contexte et terminologie

Étudier l'évolution du langage au sein d'espaces communicationnels contemporains pourrait sembler banal aujourd'hui. Cependant, Herring (1996) fait remarquer que ce choix n'avait rien d'évident dix ans en arrière et elle estime, qu'à l'époque, les linguistes réagissaient tardivement à l'idée de poursuivre des recherches dans ce domaine : « Bien que la communication médiée par ordinateur sous forme textuelle soit quasi exclusivement élaborée à partir de signes linguistiques, il est surprenant que les linguistes aient réagi aussi tardivement à l'idée que le langage médié par ordinateur soit un objet d'étude légitime. » (p.3, notre traduction). Elle ajoute que « la recherche sur la communication médiée par ordinateur en est encore à ses premiers balbutiements. » (p.5, notre traduction).

Nous avons fait ce choix en 1996 : nous intéresser à l'évolution du langage au sein des messages électroniques écrits, à travers différents moyens informatisés qui sont utilisés dans l'enseignement supérieur français. Plus précisément, le sujet d'étude concerne le discours qui apparaît au sein de trois moyens de communication asynchrone et synchrone : le courriel, le forum et le chat. Les analyses effectuées entrent dans le cadre disciplinaire de la linguistique-informatique et, à cet effet, des outils informatisés sont mis à contribution afin d'extraire des informations de type linguistique et extra-linguistique.

1.1. Médiation vs. médiatisation

Avant d'évoquer de manière plus détaillée notre choix de recherche, définissons rapidement les termes que nous utilisons : communication médiée par ordinateur et discours électronique médié. L'expression néologique que nous avons proposée en 1997 pour le français « la communication médiée par ordinateur » (CMO), ou, plus précisément dans un contexte linguistique, le « discours électronique médié » (DEM), a été adoptée essentiellement par référence à Vygotsky (1933, in Schneuwly & Bronckart, 1997), pour qui la « médiation sémiotique » était théoriquement cruciale : « Le fait central de notre psychologie est le fait de la médiation. » (p. 139). Partant d'une réflexion sur le rôle de l'outil dans le travail, Vygotsky pressent que le psychisme humain, dans ce qu'il a de spécifique, s'est développé grâce au rôle des médiateurs analogues, d'« outils psychologiques ». Il prend l'exemple des moyens mnémotechniques, à commencer par le simple nœud au mouchoir. Pour lui, ces outils sont d'abord des réalités sociales externes avant de s'intérioriser en catégories intramentales. Le langage retenait son attention du point de vue du signe, de l'outil psychologique, en tant que processus médiatisé, ou médié. Selon Vygotsky (1985), « Toutes les fonctions psychiques supérieures sont unies par une caractéristique commune, celle d'être des processus médiatisés, c'est-à-dire d'inclure dans leur structure, en tant que partie centrale et essentielle du processus dans son ensemble, l'emploi du signe comme moyen fondamental d'orientation et de maîtrise des processus psychiques. Dans la formation des concepts ce signe est le mot, qui sert de moyen de formation des concepts et devient par la suite leur symbole. » (p. 151).

Si Vygotsky évoque effectivement la médiation, en revanche, comme le soulignent Schneuwly & Bronckart (1997) : « [II] n'a cependant que peu abordé les aspects discursifs et syntaxiques de la langue. En ce domaine, l'œuvre complémentaire de Bakhtine, très proche dans ses options épistémologiques, constitue un complément indispensable. » (p. 18). Bakhtine (1979/1984), quant à lui, propose les « genres du discours ». Ces genres s'étendent de la « réplique brève du dialogue quotidien » à « l'univers du discours des publicistes » en passant par « le récit familial, la lettre, l'exposé scientifique » etc. Selon lui, « Tout énoncé pris isolément est, bien entendu, individuel, mais chaque sphère d'utilisation de la langue élabore ses types relativement stables d'énoncés, et c'est ce que nous appelons les genres du discours. » (p. 265). Il évoque ces genres du discours, et les énoncés qui en découlent sont selon lui en lien étroit avec la médiation. « Le contexte de la proposition, c'est le contexte du discours d'un seul et même sujet parlant (du locuteur) ; la relation qui existe entre la proposition et le contexte trans-verbal de la réalité (la situation, les circonstances, la pré-histoire), et les énoncés d'autres locuteurs, cette relation n'est pas immédiate ou personnelle, elle est *médiatisée* par tout le contexte environnant, autrement dit, par l'énoncé dans son tout. » (p. 279-280, nous soulignons).

L'ordinateur (entre autres facteurs) est un « outil de médiation » ou tout au moins un « support » de médiation. Il ne s'agit sans doute pas exactement d'un outil de médiation psychologique au sens de Vygotsky, mais l'ordinateur serait un outil de médiation grâce auquel le discours se voit modifié ; une autre forme de discours émerge.

Nous préférons au verbe « médiatiser » (diffuser par les médias) ou au nom « médiatisation », le verbe « médier », même s'il n'est pas officiellement reconnu en français. Le verbe *médiatiser* est, en effet, trop connoté en direction des médias pour convenir à notre cadre. De plus, *médier* existe en italien (*mediare*) et en anglais (*mediate*) et sa morphologie dérivationnelle et flexionnelle en français à partir du nom *médiation* (qui existe) permet ce néologisme.

1.2. Choix de recherche

Les corpus d'étude que nous élaborons depuis une décennie émanent exclusivement du contexte précis de l'enseignement supérieur français. Ceux-ci contiennent des messages écrits

uniquement, et ce, pour deux raisons. Tout d'abord, nous avons facilement à notre disposition les trois outils suivants : le courriel, le forum et le chat ; nous n'étudions donc pas les aspects visuels et/ou oraux, apparaissant dans les visio-conférences et/ou dans les chats audio. De plus, notre spécialisation, en linguistique-informatique, englobe l'étude de l'écrit en syntaxe et en sémantique lexicale.

Dans un premier temps, ces corpus contenaient des messages émanant d'étudiants et de collègues, majoritairement dans des échanges de personne à personne, via l'utilisation du courrier électronique (cf. Panckhurst, 1997, 1998a, 1998b, 1999a, 1999b) ; les personnes sont systématiquement prévenues que leurs messages nourrissent nos recherches, et ce, à titre anonyme. Les discours apparaissant au sein de forums asynchrones (d'une personne à un groupe de personnes) (cf. Panckhurst, 2001), puis de chats synchrones (cf. Panckhurst, 2003b) ont ensuite fait l'objet de nos recherches.

Par ailleurs, les raisons du choix de cette recherche sont triples :

1. Au début de ces recherches, en 1996, les étudiants ne disposaient pas d'adresses électroniques dans le cadre de leurs études ; nous leur avons donc fourni cette possibilité et nous avons par la suite rapidement constaté qu'ils se servaient de l'outil pour communiquer avec leur enseignante dans un contexte périphérique à celui du cours :

J'aurais préféré m'adresser à vous de vive voix pour vous présenter mes excuses, mais constatant que vous étiez très demandée par les élèves en fin de cours, je n'ai donc pas osé insister. Aussi, je profite de ce mode de communication pour vous informer de ma situation actuelle...J'élève seule ma fille de deux ans et demi et j'ai beaucoup de difficultés à la réveiller tôt en ce moment, ceci expliquant mes retards systématiques aux T.D. du mardi matin. (étudiant, corpus 1996).

Si je vous écris, c'est parce que je préfère ne pas être en face de vous pour vous entendre me répondre définitivement « non » ; au moins le courrier électronique m'aura donné « cet avantage »...je ne sais même pas pourquoi je continue à espérer que pour 0,5 points les choses pourraient changer. (étudiant, corpus 1997).

2. En parallèle, nous assurions des formations de formateurs enseignants, pour les TICE ; certains collègues étaient réticents et nous voulions comprendre pourquoi :

Je n'arrive pas à m'habituer au courrier électronique. C'est un moyen de communication qui me paraît distant, froid. (enseignant, corpus 1997).

J'espère que ce moyen de communication ne remplacera jamais la communication inter-personnelle. (enseignant, corpus 1997).

3. L'accès aux documents électroniques était quasi-immédiat ; même si les recherches en traitement automatique des langues sur des vastes corpus foisonnaient déjà, l'idée de pouvoir disposer de manière quasi-immédiate de corpus de taille raisonnable était très séduisante.

Enfin, le choix du contexte de l'enseignement supérieur français, n'était pas uniquement dû à l'accès simplifié aux données textuelles. Nous voulions étudier l'évolution du langage, et nous pensions — à tort, à cette époque — que les discours contenus dans ces messages électroniques ne seraient pas fondamentalement différents d'autres moyens écrits plus classiques. Nous avons très vite observé que ce n'était pas le cas, et, à travers ce chapitre nous nous efforcerons de démontrer en quoi.

2. Caractéristiques du discours électronique médié (DEM)

Dans un premier temps, nous comparons le DEM à l'oral, d'une part, et à l'écrit, d'autre part, puis nous montrons comment le DEM contient, dans certains cas, des marques spécifiques.

2.1. Les marqueurs linguistiques et extra-linguistiques de l'oral

Le message écrit est quasi dépourvu de marqueurs para-verbaux et non-verbaux (intonation, regard, proximité, posture, gestualité, etc.) qui permettent, dans une situation d'échange verbal conversationnel, une certaine régulation. Tout naturellement, en situation de CMO, les internautes essaient de reproduire ces fonctionnalités. Certaines utilisations rapprochent, en effet, le DEM d'une situation de communication orale :

- les binettes (emoticons, smileys) permettent d'introduire des aspects sémiologiques non-verbaux :-)

- l'utilisation de mots saisis entièrement en majuscules exprimant colère ou agacement ou encore angoisse : *C'est pourquoi je vous lance un S.O.S. HELP ME !!!!!* ;
- l'allongement ou la répétition de caractères simule l'intonation : *ssssuuuuuppppeerrrr !!!!!*
- certaines indications sont spécifiquement linguistiques ; verbes de clôture : *reparler, laisser* — *Un cours va débiter dans cette salle, je vous laisse* ; *Salutations amicales et au plaisir de te reparler* ;
- certaines catégories d'erreurs sont typiques d'une situation de communication orale : *Voici mon nouvelle adresse électronique* ;
- simulation de tours de parole ; marques typographiques ainsi que l'usage des chevrons '>' ou des barres verticales '|' qui indiquent une répétition du discours entre l'énonciateur et son destinataire.

2.2. Les marques écrites omniprésentes

De manière générale, le DEM semble ressembler à d'autres formes de l'écrit, car certaines marques sont omniprésentes :

- l'usage des noms est plus important que l'usage des verbes ;
- les formes interrogatives sont majoritairement des formes classiques qui excluent l'intonation ;
- la négation 'ne...pas' est tenace (pour plus de détails, cf. Panckhurst 1998b, 1999a, 1999b).

Véronis (in Guimier de Neef & Véronis, 2004) constate que les fréquences (ou les marques) dépendent davantage de la situation, du registre, etc., que de l'opposition écrit/oral. Par exemple, le niveau de langue contenu dans un message de courriel sera différent en fonction du destinataire. S'il est vrai qu'une certaine souplesse est de facto admise quand on se sert d'un ordinateur, nous avons constaté, dans nos recherches, des aspects inattendus, à la fois sur le plan linguistique et extra-linguistique :

1. Erreurs : L'analyse des courriels d'étudiants (corpus 1996-1997) a montré que ceux-ci comportaient des erreurs. Les enseignants trouvaient cela normal : « les étudiants ne savent plus écrire correctement ». À titre comparatif, nous avons analysé également des messages en provenance de collègues. Ceux-ci contenaient également certaines erreurs grammaticales « graves » qui ne seraient sans doute pas apparues dans un autre contexte, par exemple dans des lettres manuscrites.
2. Appropriation de l'outil : Nous avons fait l'hypothèse que les étudiants, se servant d'un forum électronique pour la première fois en situation pédagogique (corpus 2001), resteraient dans un cadre formel circonscrit, en respectant, par exemple, des règles fournies par l'enseignant. Nous avons ainsi donné quelques indications concernant les ouvertures et les clôtures de message. Cependant, les étudiants se sont approprié extrêmement rapidement l'outil en fonction de leurs propres usages et besoins, même en situation pédagogique et même après avoir reçu des consignes strictes.
3. Situation de communication : Au début de ces recherches, nous nous attendions à ce que les étudiants, précisément, parce qu'ils étaient en situation pédagogique, utilisent les outils électroniques (le courriel uniquement à cette époque) pour des échanges relevant du contexte pédagogique/didactique. En fait, les étudiants se sont servis aussi de l'outil pour faire part d'information hors contexte stricte pédagogique (cf. § 1.2.)

2.3. Les spécificités

Quelques marques linguistiques et extra-linguistiques s'avèrent spécifiques au DEM. Aussi, même s'il existe des ressemblances avec certains aspects appartenant à l'oral (ou des oraux) et/ou à l'écrit (des écrits), le DEM est effectivement un genre de discours à part entière.

D'un point de vue linguistique, nous avons remarqué :

- une utilisation prédominante du présent de l'indicatif ;
- une utilisation importante de pronoms à la 1^{re} et à la 2^e personne ;
- un pourcentage diminué de verbes par rapport à d'autres formes de l'écrit (plus de 25 % dans des écrits plus classiques).

Dans les différents corpus étudiés, le taux d'utilisation des verbes est en moyenne de 9 à 20 %. Par ailleurs, ces verbes sont souvent des auxiliaires ou des semi-modaux (*être, avoir, pouvoir, devoir, sembler, paraître, croire, savoir, vouloir, etc.*)

D'un point de vue extra-linguistique, les caractéristiques sont diversifiées selon les supports. Certaines marques sont plus présentes lors de l'utilisation du courriel, d'autres concernent davantage le forum de discussion ou le chat, et d'autres encore sont présentes au sein des trois moyens de communication.

Les échanges par courriel sont souvent rapides (*Je dois lui rendre ma réponse demain. J'attends donc ta réponse.*) et informels. L'émetteur s'attend souvent à une réponse le jour même ; si le destinataire n'accuse pas réception, il est relancé (*Je ne sais pas si vous avez bien reçu mon courrier.*) Comme les messages sont concis, et les marques non verbales et/ou paraverbales réduites, une certaine montée d'angoisse, d'impulsivité, d'agressivité peut être notée (*Quand on a été élu [...], il faut être un minimum responsable ; Tu me pompes le peu d'énergie qui me reste.*). Parfois un réajustement ultérieur est opéré (*J'espère que je ne t'ai pas blessée avec certaines de mes paroles. Avec l'email c'est toujours le danger. On n'a pas la patience d'écrire des pages, une bonne discussion c'est quand même mieux.*). Ces relations sont typiques de ce mode de communication.

Herring (1996) a, par ailleurs, fait état de véritables problèmes d'accoutumance et d'incapacité relationnelle. Selon elle, « des visions 'utopiques' de sociétés virtuelles sans distinction de classe ou de genre émergent en même temps que des visions 'dystopiques' incluant un surplus d'information, une accoutumance au courriel, une agressivité sans bornes, et une incapacité grandissante à assumer des rencontres en face-à-face » (p. 1, notre traduction).

La question est de savoir alors s'il s'agit bien dans ces cas extrêmes d'une situation de communication, et si oui de quel type. Breton (1997) répond, pour sa part, de la façon suivante : « L'élément essentiel d'un nouvel individualisme que provoque le recours systématique aux médias est sans doute la façon dont ceux-ci privilégient la communication indirecte, abstraite, sans véritable rencontre avec l'autre. [...] Le risque que nous prenons est celui d'une société certes fortement communicante, mais faiblement rencontrante. » (p. 160).

Pour ce qui concerne les forums de discussion et les chats, l'absence d'ouverture est en général de mise. L'internaute se lance directement dans la discussion. En revanche, s'il est en retard au rendez-vous virtuel pour le chat, une ouverture est présente, accompagnée d'explications concernant le retard. Les formules de clôture sont souvent absentes dans les forums car le nom est automatiquement posté avec le message. En revanche, dans tous les chats pédagogiques que nous avons analysés, les étudiants sans exception saluent avant de quitter la salle virtuelle.

Les ellipses (*et l'exercice de la 2^e séance ? Lequel ? Dernière page*) et les abréviations foisonnent dans les chats. Par ailleurs, à cause des problèmes de séquentialité et parce que les chats n'exploitent pas certaines propriétés fondamentales d'une conversation classique (Crystal, 2001) des malentendus peuvent émerger, et les messages doivent être « démêlés ». Entre le moment de l'écriture du message et son affichage à l'écran, le décalage s'instaure. Ces problèmes (turn-taking, floor-taking — tours de parole ; éviter de parler en même temps, adjacency pairing — ordre, séquentialité : questions suivies de réponses, plaintes suivies d'excuses, etc.), peuvent, selon le romancier Gibson (2003), décourager certaines personnes : « La salle de chat lui est peu reconfortante. C'est étrange, même avec des amis, c'est comme si tous étaient assis dans une cave sombre en train de se parler à 5 mètres les uns des autres. La vitesse frénétique et la brièveté des fils de discussion couplées avec le sentiment que tout le monde parle en même temps, à contre-courant, la découragent » (p. 4, notre traduction).

Pour toutes les raisons qui viennent d'être évoquées dans ce paragraphe, le discours médié électronique constitue un genre présentant des marques spécifiques. À ce sujet, Crystal (2001) rappelle que « Netspeak (parler Net) correspond davantage à un agrégat de traits relevant de l'écrit et de l'oral. [...] Ce genre permet des choses que les autres media ne permettent pas, et doit être perçu comme une nouvelle espèce de communication. [...] Les textes électroniques, quelle que soit leur nature, ne sont tout simplement pas semblables à d'autres formes textuelles. [Leurs propriétés] se combinent avec celles associées au discours et à l'écriture afin de produire le parler Net, qui constitue alors un véritable 'medium d'un troisième genre' » (p. 47-48, notre traduction).

3. Étude de cas

Depuis 2001 nous avons élargi nos études initiales, portant sur le discours contenu dans les messages électroniques de type courriel, aux forums (asynchrones) et aux chats (synchrones) dans le cadre de l'utilisation de l'outil WebCT, la plateforme d'enseignement à distance choisie par l'université Paul-Valéry (<http://ead.univ-montp3.fr:8900>), à la fois pour des cours délivrés aux étudiants en présentiel (en complément de cours) et en formation ouverte et à distance (FOAD).

Nous avons constitué un corpus constitué à partir de deux cours : le premier a été recueilli lors d'une initiation à l'informatique, dans le cadre d'un cours de préparation au certificat national C2i (*Certificat Informatique et Internet*), délivré en présentiel en L1, L2 et L3 — corpus C2i. Le second a été constitué à partir d'un cours portant sur l'évaluation de logiciels en FOAD en Master, niveau 2 (Gestion des apprentissages en FOAD), cours mené entièrement à distance, à l'université Paul-Valéry, Montpellier — corpus M2. Les étudiants devaient échanger sur des forums limités à une communication entre 3 à 5 participants dans le cadre d'un travail à effectuer en groupe. L'enseignant était inclus comme participant au forum, mais n'est intervenu que rarement.

3.1. Présentation du corpus

L'analyse que nous proposons dans ce chapitre porte donc sur l'appropriation de l'outil et l'évolution du discours dans un cadre pédagogique dans lequel les étudiants se sentaient en situation quasi-privée. Pour mener cette étude, et afin de dégager des traits lexicaux, syntaxiques, voire pragmatiques, nous avons eu recours à deux logiciels de traitement automatique des langues : Cordial, version 8.13 (Synapse : <http://www.synapse-fr.com/>) et Nomino, version 4.2.22 (<http://www.ling.uqam.ca/nomino/>).

Le corpus C2i a été élaboré à partir de forums étudiants entre février et mars 2005 et le corpus M2 a été constitué de forums et de chats entre mars et avril 2005. Le tableau 1 ci-dessous récapitule ces données :

Tableau 1 : Statistiques des corpus C2i et M2 : date, participants, messages

	Forum-C2i	Forum-M2	Chat-M2
Date et participants	février-mars 2005 ; 7 groupes de 3 à 4 personnes ; 26 participants.	mars-avril 2005 forums : 5 groupes de 2 à 4 personnes ; 14 participants.	3 séances (mars-avril 2005) ; 15 participants.
Nombre de messages	295 au total (de 17 à 78 messages par groupe).	225 au total (de 4 à 150 messages par groupe).	1020 messages.
Volume (n° de mots)	13 896	24 980	10 820
Volume (moyenne de mots par message)	47,9	97	10,7

La moyenne des mots par message est nettement plus réduite dans les chats que dans les forums. Ce résultat est concordant avec d'autres recherches sur les chats (cf. Panckhurst, 2003b, pour une étude portant sur un cours de L1, mené entièrement à distance : 11 mots par message au sein du chat).

Une différence de longueur des messages est à noter. Cette longueur dépend de la nature du travail demandé. En Master les exigences sont plus fortes, car une participation plus impliquée de l'étudiant est demandée (97 mots en Forum-M2 vs. 47,9 mots en Forum-C2i). De plus, ces étudiants sont habitués à communiquer avec l'environnement de FOAD, depuis octobre 2004. Enfin, la longueur moyenne plus importante d'un message peut également traduire certaines pratiques de production (par exemple, messages copiés-collés en importation à partir d'autres logiciels de chat).

3.2. Commentaires & résultats

3.2.1. Le forum vs. le chat

Les observations qui viennent d'être faites montrent que le forum convient mieux pour donner des informations, alors que le chat permet d'échanger des éléments de *valeur sociale*. À ce sujet, Crystal (2001) note que : « Les chats permettent une interaction personne-à-personne qui est essentiellement sociale. Les avantages sociaux dépassent les inconvénients sémantiques. » (p. 168, notre traduction). Cette activité sociale devrait jouer un rôle développemental comme le suggère Vygotsky (1985).

Dans le cadre de la situation pédagogique analysée, il s'agit aussi de rassurer et motiver l'étudiant en enseignement à distance (EàD). Ce dernier est, en effet, toujours en danger potentiel d'abandon, comme l'indique l'extrait qui suit : « Au revoir et merci. Ce contact m'a motivée pour la suite. » (étudiante, L1, EàD, à l'étranger).

Le chat contient une proportion beaucoup plus importante de phrases interrogatives que le forum ou le courriel. D'ailleurs les étudiants en sont conscients : « C'est rigolo ; puis j'ai l'impression qu'on communique plus, on pose plus de questions. » (étudiant L3, présentiel).

Tableau 2 : Taux en (%) des formes interrogatives présentes dans les messages (chats, forums, courriels), entre 2003 et 2005.

	2003 L1 EàD (forum)	2003 L1 EàD (chat)	2003 prés. (courriel)	2005 C2i (forum)	2005 M2 (forum)	2005 M2 (chat)
formes interrogatives	7,2 %	36,4 %	1,1 %	9,4 %	4,4 %	15,3 %

Dans le tableau 2, on constate une différence entre le forum et le chat en situation pédagogique quant à l'usage des formes interrogatives. Le seul corpus qui correspond à une situation non-pédagogique ne comporte que 1,1 % seulement de formes interrogatives. Dans cette situation, ont été analysés des messages envoyés par la présidente sur les listes électroniques des personnels. Il s'agit uniquement de questions rhétoriques (pour plus d'éléments, cf. Panckhurst & Bouguerra, 2003).

Pour les forums, le pourcentage d'usage de tournures interrogatives varie de 4,4 % (corpus M2, dans lequel les étudiants sont plus autonomes) à 9,4 % (corpus C2i, dans lequel certains sont peu sûrs d'eux et comptent sur les camarades pour les aider). L'utilisation accrue des formes interrogatives est donc nette pour le chat. Elle est toujours plus importante que celle faite dans les forums. Dans ce cas, l'utilisation varie entre 15,3 % (M2) et 36,4 % (corpus 2003 L1). Il s'agit sans doute d'une situation pédagogique différente. En effet, pour certains étudiants le cours constituait une toute première pratique de l'EàD.

3.2.2 Traits spécifiques du DEM

Plusieurs traits du DEM caractérisent notre corpus 2005 :

- Dans ce contexte, les étudiants utilisent fréquemment les auxiliaires (*avoir, être*) et les semi-auxiliaires de modalité (*devoir, pouvoir, savoir, croire, vouloir, falloir, etc.*). À eux seuls, ces verbes représentent plus de 40 % du total des verbes utilisés dans le corpus. Il semblerait, d'après Yates (in Herring 1996), que l'utilisation de modaux en DEM soit bien plus importante qu'à l'oral ou qu'à l'écrit (l'écrit ayant le pourcentage le plus bas des trois) : Voici quelques exemples : obligations/contraintes : *je dois vous quitter* ; possibilités, accomplissements : *je pourrai bientôt vous l'envoyer* ; combinés avec une connaissance de la situation : *je sais, mais comment on le code ?* ; au contraire, hésitations/questions : *oui, mais je ne sais si c'est juste, est-ce qu'on doit vous envoyer les exercices ?*
- Le présent de l'indicatif constitue le temps verbal le plus utilisé (C2i : 67,5%, forum-M2 : 70,7% ; chat-M2 : 70%).
- Les pronoms sont massivement employés : de 11 à 16 % des occurrences totales ici (contre seulement 7 % dans l'étude du courriel (Panckhurst, 1999)). Les pronoms sujet sont très importants afin d'ancrer le discours entre partenaires de la situation de

communication. Il s'agit surtout des pronoms de 1^{re} et de 2^e personne qui sont nettement plus élevés que les pronoms sujet de 3^e personne.

- L'usage des formes interrogatives et négatives était resté relativement classique jusqu'au corpus 2005. L'intonation était rare, sauf dans des situations elliptique et/ou abrégative dans les chats et la négation figurait avec les deux particules « ne » et « pas ». Dans les corpus de 2005, seules 10 occurrences de « pas » figurent sans la particule « ne » (sur 285 au total dans le chat-M2). Cependant, ces usages vont souvent de pair avec une situation particulière (niveaux de langue, ellipses, utilisation combinée avec le texto). Par exemple : *Ouai ! on peut pas parler en audio ailleurs ? ; Peux pas non plus ; C pas grave*. 37 % des formes interrogatives sont intonatives, apparaissant avec l'unique point d'interrogation, mais ce procédé reste circonscrit aux cas contenant essentiellement des tournures elliptiques ou des tournures consacrées : *Simple ou simpliste ? ; Ça va ?*, etc.

3.2.3 Relations DEM/écrit/oral

Nous avons constaté que le DEM contient des marques spécifiques. La question est de savoir si le DEM se rapproche de l'oral, de l'écrit. Plusieurs auteurs (Gadet, 1996, pour le français ; Haliday, 1989, pour l'anglais) ont indiqué des différences notables entre des formes orales et écrites. Pour sa part, Herring (1996) a repéré des différences entre l'oral, l'écrit et le DEM. Haliday (1989) a spécifié, pour l'anglais, que l'oral est « grammaticalement dense » (autrement dit, il comporte beaucoup d'articles, de pronoms, de prépositions, de conjonctions, d'adverbes, de verbes), et que l'écrit est « lexicalement dense » (les noms sont fréquemment utilisés) : « Il existe une différence caractéristique entre la langue parlée et la langue écrite. L'écrit contient un taux très élevé d'items lexicaux par rapport au nombre total de mots existant dans un texte » (p. 61, notre traduction).

Ces observations sont valables pour le courriel également (Panckhurst, 1999). Nous avons observé un faible quota de verbes et un pourcentage très important de noms dans le DEM. À partir de ce constat, il est possible d'inférer que le DEM ressemble à d'autres formes de l'écrit, mais, pour cet aspect seulement. En effet, d'autres aspects le rapprochent davantage de l'oral.

Selon Gadet (1996), les traits pour l'oral sont les suivants :

- hétérogénéité, variabilité de l'oral comparé à l'homogénéité de l'écrit, ayant des codes et des normes stables ;
- très nombreuses pauses, hésitations, mots incomplets, anticipations, interruptions ;
- marques intonatives, etc.

Précédemment (§ 2.1.), nous avons mentionné les aspects du DEM qui le rapprochent de la communication orale (tours de parole, ratages, binettes etc.). Selon Gadet (1996), « L'oral relève d'une complexité d'ordre grammatical (alors que l'on entend souvent dire que l'oral n'a pas de grammaire) et l'écrit d'une complexité d'ordre lexical (qui va de pair avec une certaine monotonie grammaticale). Ceci s'accompagne d'une préférence de l'oral pour les verbes, et de l'écrit pour les noms. » (p. 23).

Dans Panckhurst (1999a), la richesse lexicale était encore plus importante : 75 % de formes nominales et moins de 10% de formes verbales. Dans des écrits plus classiques, les verbes sont censés représenter plus de 25 % des formes utilisées.

Les moyennes observées dans le tableau 3 indiquent que le pourcentage d'utilisation des noms est effectivement très élevé. Il faut toutefois noter que les moyennes du corpus C2i sont particulières (cf. § 3.2.5.).

Tableau 3 : Catégories syntaxiques (formes/types) en fonction des différents moyens de communication.

	C2i (forum)	M2 (forum)	M2 (chat)
Noms	70,90%	60,20%	59,40%
Verbes	14,40%	17,40%	19,40%
Adjectifs	9,60%	15,50%	12,90%
Adverbes	5,20%	7%	8,40%

Le calcul a été réalisé de la façon suivante. Dans un corpus donné, différentes catégories syntaxiques apparaissent : verbes, noms, etc. Prenons l'exemple des verbes ; chaque verbe

distinct compte pour une et une seule *forme* (ou type). Si le même verbe apparaît à plusieurs reprises dans le corpus, des *occurrences* sont alors comptabilisées. Par exemple, dans le corpus C2i, le verbe *pouvoir* est utilisé. Cela compte pour une *forme* et ce verbe *pouvoir* a été utilisé 113 fois, soit 113 occurrences. Le tableau 3 récapitule les statistiques pour les formes et le tableau 4, pour les occurrences.

3.2.4. Comparaison évolutive et nouveautés

Les occurrences observées jusqu'à présent indiquaient une préférence nette pour l'usage des noms dans le DEM (courriel, forum ou chat). Grâce à l'analyse d'un plus grand nombre de corpus produits avec différents types de support de communication, nous avons pu constater un phénomène nouveau.

Nos observations de 1999 (Panckhurst, 1999 — corpus constitué de courriels datant de 1996 à 1999, contenant plus de 342 000 mots) montraient que le DEM se rapprochait d'autres formes écrites (avec une utilisation verbale encore plus réduite : 18% d'occurrences verbales). Or, dans le tableau 4, entre 1999 et 2005, nous constatons la réduction de la fréquence nominale, l'augmentation de l'utilisation verbale, la stabilité adjectivale, et une fréquence croissante d'adverbes modalisateurs. Du courriel au chat en passant par les forums de discussion, on voit que les usages linguistiques tendent à s'inverser en fonction de la pratique des usagers. Tout se passe comme si pour certaines marques, les catégories grammaticales du DEM se rapprochent de celles de l'oral.

Tableau 4 : Catégories syntaxiques (occurrences) en fonction des différents moyens de communication (comparaison 1999/2005).

	Courriel (1999)	M2 (forum) (2005)	M2 (chat) (2005)
Noms	69%	53%	44%
Verbes	18%	25%	32%
Adjectifs	8%	8%	8%
Adverbes	5%	13%	17%

À partir de l'analyse linguistique, il est possible de mieux appréhender ce qu'est le DEM, ou les DEM. Bien entendu, il faut poursuivre les recherches comparatives entre courriel, forums et chats pour vérifier la stabilité de ces observations.

3.2.5. Usage du Texte

Dans le forum-C2i, nous avons relevé plusieurs exemples de texto, dont voici un extrait :

« je c ke c chian de devoir se connecter mé en fait on é noté ossi sur le nombre d'échange et ce qu'on dira dc si vous avez un ordi à disposition ca seré cool d'envoyer des infos sur d sites par ex pr le dossier ou se genre de truc pr le site je pense ke ca devré allé vite je vou envéré un message ds les jours suivants sur chaque page k on peu faire et chacun pourra dire ce ki lui plé le plus de faire moi je m en fou voila. »

(forum entre étudiants, travail en groupe, 7/2/05)

Différents phénomènes y apparaissent comme ceux relevés par Guimier de Neef (in Guimier de Neef & Véronis, 2004). De manière non-exhaustive, nous en exposons quelques-uns ici :

- modifications de type phonétique (*c, mé, é, ossi, seré, envéré...*),
- suppressions de la lettre finale d'un mot (*chian, vou, fou, peu...*),
- des troncations intermédiaires, des abréviations, des réductions aux squelettes consonnantiques, (*dc, ordi, pr, ds*) etc.

L'extrait analysé présente également une absence totale de ponctuation (aucun point, aucune virgule, absence d'apostrophes, etc.). Ces usages s'accompagnent d'un problème de traitement automatique du langage. Nous avons donc soumis cet extrait à Cordial et à Nomino. Les résultats, ainsi que le corrigé, apparaissent dans le tableau 5 ci-dessous :

Tableau 5 : Problèmes de repérage automatique de Texto.

	Cordial		Nomino	
	Texto	Corrigé	Texto	Corrigé

Noms	33	15	27	14
Verbes	12	21	17	22
Adjectifs	2	3	2	2
Adverbes	4	3	2	2
Autres	47	57	50	59

Cordial et Nomino traitent tout élément qu'ils ne reconnaissent pas comme un « nom ». En confrontant les statistiques des résultats pour le texto et le corrigé, on comprend pourquoi il faut analyser avec précaution les statistiques obtenues pour le corpus C2i (cf. tableau 3). Certaines unités du texto ont été mal catégorisées par l'outil de TAL. Il s'agit essentiellement de verbes et de pronoms qui ont été catégorisés de manière erronée en noms. De plus, cet exemple est bien sûr très bref. Toutefois, il faut imaginer les implications pour les futures recherches en traitement automatique du langage naturel. Selon Guimier de Neef (in Guimier de Neef & Véronis, 2004), il faut des règles qui permettent de repérer, entre autres :

- la segmentation des messages : frontières des mots, décoller et interpréter les formes agglutinées ;
- le lexique : augmentation en volume (abréviations, variantes orthographiques, ambiguïté : *c = sait* ou *s'est*, etc.) ;
- l'analyse syntaxique : absence de marques formelles (apostrophes, ponctuation, majuscules, traits d'union etc.), désambiguïser les abréviations.

Enfin, il faut noter que tous les phénomènes n'ont pas été pris en compte, comme ceux concernant les procédés dynamiques des mécanismes d'écriture. De plus, il existe une réelle difficulté d'interprétation lorsque des procédés sont cumulés (par exemple agglutination de mots et écriture phonétique). Enfin, la question est de savoir comment segmenter des unités lexicales en l'absence de séparateurs. Lorsque la ponctuation est totalement absente, cela paraît impossible. Ce type de difficulté augmente avec la longueur du message.

Véronis (in Guimier de Neef & Véronis, 2004) a indiqué divers problèmes liés au traitement de ce type de langage par des outils informatisés. Les principes actuels du TAL ne montrent-ils pas leurs limites ? Comment fera-t-on pour traiter des accès lexicaux de manière convenable ? Les traitements automatiques doivent-ils demeurer toujours séquentiels, et, si oui, comment traiter des phénomènes concernant précisément le manque de séquentialité réelle (par exemple, dans les chats) ? Comment maîtriser l'explosion combinatoire ?

4. Conclusion

Il est important de poursuivre ces recherches, dans une double direction. Sur un plan plus pragmatique, comprendre comment se forment des communautés en ligne. Les étudiants échangent en tant que groupe et forment ensuite des communautés. Comment un individu réagit-il et évolue-t-il au sein de ces communautés ? Il est important de comprendre davantage ces processus, en tant qu'enseignants et tuteurs pour en tirer un meilleur usage didactique. Il faudrait, également, effectuer plus de recherches sur des aspects mêlant les techniques de TAL et l'étude des textos. Ces derniers commencent effectivement à envahir l'espace communicationnel virtuel dans l'enseignement supérieur.

Afin de poursuivre le travail dans ces deux directions, élaborer d'autres corpus s'avère nécessaire. En fonction de l'évolution des usagers, il devrait alors être possible de confirmer que les textos font leur apparition dans le contexte pédagogique asynchrone dans lequel l'enseignant est également impliqué. Il y a quelques années cet usage n'aurait pu être envisagé. Penser que le langage écrit des SMS ferait son apparition de manière de plus en plus importante dans les espaces asynchrones était inimaginable. On peut le comprendre pour le chat, mais plus difficilement dans un forum de discussion universitaire. Il est vrai que le corpus C2i présenté ici contenait uniquement des discussions entre pairs. Mais nous sommes sûre que nous verrons pointer ce type de discours prochainement dans les forums entre étudiants et enseignants. Nous n'avons sans doute pas pris encore la mesure de cette révolution langagière.

En tout état de cause, comme le souligne Herring (1996), les études doivent être menées de manière pluridisciplinaire, afin de mieux comprendre cette évolution langagière moderne. Selon elle : « Ce secteur [...] continuera à l'avenir à rassembler plusieurs disciplines

universitaires et à faire en sorte qu'elles apprennent les unes des autres. Parallèlement, l'émergence de sous-spécialisations sera favorisée par l'expertise disciplinaire appliquée aux questions portant sur la CMO. Un approfondissement multidisciplinaire est nécessaire pour nous aider à mieux comprendre ce phénomène technologique aux implications profondément humaines. » (p. 10, notre traduction).

Par ailleurs, il est remarquable que les usages langagiers évoluent toujours de manière aussi créative. Nous pensons que ce domaine constituera un terrain de recherches fructueuses pour bien des années futures. Comme le souligne Crystal (2001) : « Il est vraiment remarquable qu'autant de personnes se soient si rapidement adaptées aux exigences linguistiques de ces situations nouvelles et aient su exploiter tout le potentiel du nouveau médium de manière suffisamment créative pour favoriser l'émergence de nouvelles formes d'expression. Cela s'est produit en quelques décennies seulement. J'en conclus que la faculté linguistique de l'être humain est en bonne forme. L'apparition du parler Net nous révèle ce que l'*Homo loquens* a de meilleur. ».

Références :

- Anis, J. (1999), *Internet, communication et langue française*, Paris, Hermès.
- Anis, J. (2002), *Communication électronique scripturale et formes langagières : chats et SMS*, consulté sur : <http://oav.univ-poitiers.fr/rhrt/index.htm>
- Anis, J., de Fornel, M., Franekel, B. (organisateur) *La communication électronique : Approches linguistiques et anthropologiques*, Colloque international, EHESS, Paris, 5-6 février 2004.
- Bakhtine, M. (1979), trad. fr. 1984), *Esthétique de la création verbale*, Paris : Gallimard.
- Breton, P. (1997), *L'utopie de la communication*, Paris, Découverte.
- Crystal, D. (2001), *Language and the Internet*, Cambridge, Cambridge University Press.
- Gadet, F. (1996), Une distinction bien fragile : oral/écrit, *TRANEL*, 25, 13-27.
- Gibson, W. (2003), *Pattern Recognition*, G. P. Putnam's Sons.
- Guimier de Neef, É., Véronis, J. (2004, organisateurs) *Le traitement automatique des nouvelles formes de communication écrite*, journée d'étude, Paris, 5 juin 2004 : <http://www.up.univ-mrs.fr/veronis/je-nfce/index.html>
- Haliday, M.A.K. (1989), *Spoken and written language*, Oxford University Press.
- Herring, S. C. (ed), (1996), *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*, Amsterdam/Philadelphia: John Benjamins.
- Marcoccia, M. (2000), *La représentation du non-verbal dans la communication écrite médiatisée par ordinateur*, consulté sur : <http://www.montaigne.u-bordeaux.fr/GRECO/colloque-2000.html>
- Panckhurst, R. (1997), *La communication médiatisée par ordinateur ou la communication médiée par ordinateur ?*, *Terminologies nouvelles*, 17, 56-58.
- Panckhurst, R. (1998a), *Analyse linguistique du courrier électronique*, Actes du colloque Les relations entre individus médiatisées par les réseaux informatiques, GRESICO, Vannes, 10-11/9/98, Paris, L'Harmattan, 47-60.
- Panckhurst, R. (1998b), *Marques typiques et ratages en communication médiée par ordinateur*, Actes du colloque CIDE 98, INPT, Rabat, 15-17/04/98, Paris, Europa Productions, 31-43.
- Panckhurst, R. (1999a), *Analyse linguistique assistée par ordinateur du courriel*, in *Internet, communication et langue française*, Anis, J. (Ed.), Hermès, 55-70.
- Panckhurst, R. (1999b), *La Communication médiée par ordinateur : un discours autre ?*, in *L'autre en discours*, Bres, J., Delamotte-Legrand, R., Madray, F., Siblot, P (éd.), Dyalang-Praxiling, Service des publications de l'Université Paul-Valéry, Montpellier 3, 307-331.
- Panckhurst, R. (2001), *Distance, open and virtual lifelong learning : shaping the transition within a French University*, 20th World conference on open learning and distance education, Düsseldorf, Allemagne, 1-5 avril, Proceedings, ISBN-NR.3-934093-01-9, 2001, ICDE - Oslo, Norway & FernUniversität Hagen, Allemagne.
- Panckhurst, R., Bouguerra, T. (2003a), *Communicational and methodological/linguistic strategies using electronic mail in a French University*, Proceedings, 8th International Symposium on Social Communication, Santiago de Cuba, 20-24 janvier, 548-554.

- Panckhurst, R. (2003b), Computer-mediated communication and linguistic issues in French University online courses, Online Educa, Berlin, 3-5 décembre, Proceedings, 454-457, ISBN 3- 9808909-2-9.
- Schneuwly, B., Bronckart, J.-P. (éds., 1987), Vygotsky aujourd'hui, Neuchâtel : Delachaux et Niestlé, 237 pp.; 2ème édition, 1997.
- Vygotsky, L. S. (1985), Pensée et langage, Paris : Terrains/Éditions Sociales, [traduction française par F. Sève].