



**HAL**  
open science

# WEBCOMPLETION - PROTOCOLE DE NORMES ASSOCIATIVES SUR INTERNET

Jean-Philippe Meiye, José Rouillard, Dominique Vaufreydaz

► **To cite this version:**

Jean-Philippe Meiye, José Rouillard, Dominique Vaufreydaz. WEBCOMPLETION - PROTOCOLE DE NORMES ASSOCIATIVES SUR INTERNET. École Sémantique, CNRS, Feb 1998, Asnelles-sur-mer, France. inria-00326146

**HAL Id: inria-00326146**

**<https://inria.hal.science/inria-00326146v1>**

Submitted on 1 Oct 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# WEBCOMPLETION - PROTOCOLE DE NORMES ASSOCIATIVES SUR INTERNET

Jean-Philippe Meije, José Rouillard, Dominique Vaufreydaz

Laboratoire CLIPS-IMAG / UMR CNRS 5524

Université Joseph Fourier

B.P 53 - 38041 Grenoble Cedex 9 France

Jean.Philippe.Meije@imag.fr, Jose.Rouillard@imag.fr, Dominique.Vaufreydaz@imag.fr

## 1. INTRODUCTION

Nous présentons une nouvelle méthode d'approche de construction des normes associatives verbales fondée sur WebCompletion, un protocole d'expérimentation sur Internet<sup>1</sup>. Avec l'avènement du Web<sup>2</sup> via Internet, il est désormais possible d'optimiser les espaces de rencontres virtuelles en offrant à l'utilisateur la possibilité de pouvoir exécuter à distance et en temps réel des protocoles d'expériences. En ce sens, WebCompletion est une alternative qui envisage de pallier les insuffisances des méthodes classiques actuelles notamment celles qui sont fondées sur la passation papier crayon - Rosenzweig [15], rétro-projection - Oléron et LeGall [14], auto-projection - Meije [10], etc. Celles-ci connaissent en effet des difficultés en termes de temps de dépouillement des données, et ainsi de réalisation d'un panel d'enquête représentatif des 4 catégories socioprofessionnelles : élèves, lycéens, étudiants et salariés.

## 2. PROBLEMATIQUE

Depuis le recueil des normes d'associations verbales de Kent-Rosanoff en 1910, cité par Rosenzweig [15], les normes associatives constituent désormais une étape non seulement charnière dans le

processus de validation des phénomènes cognitifs, mais aussi un domaine de recherche en soi. En effet plusieurs travaux ont montré que la force associative, à savoir la nature des relations sémantiques que peuvent entretenir les mots dans une langue donnée, reflète certaines propriétés des activités cognitives, dans le domaine par exemple de la mémoire sémantique ou à long terme (Collins et Quillian, [01] ; Fraisse, [04] ; Meyer, [13] ; Loftus, [08] ; Le Ny, [06]).

De façon officielle<sup>3</sup>, deux normes d'associations verbales ont été produites dans le cadre de la langue française (Rosenzweig ; Oléron et LeGall [14]). Les limites et les difficultés communes de ces travaux (les normes ronéotées incluses) concernent les points suivants :

- Nombre de sujets : aucune étude n'a excédé le nombre de 300 sujets,
- Rapport homme vs. femme : aucune étude n'est fondée sur un nombre de sujets homme vs. femme significatif,
- Catégories socioprofessionnelles : aucune étude ne prend en compte la catégorie socioprofessionnelle des élèves et lycéens,
- Temps de dépouillement : aucune norme n'a mis un temps de dépouillement inférieur à un mois,
- Mode de dépouillement : aucun dépouillement ne s'est effectué sur un mode automatique ou semi-

<sup>1</sup> - Internet définit le réseau mondial via lequel les ordinateurs distants interconnectés échangent électroniquement des informations.

<sup>2</sup> - Web, ou WWW, est l'acronyme de World Wide Web. Il définit le protocole de communication et d'affichage des données hypermédiés (texte, son, image, vidéo) sur Internet.

<sup>3</sup> - Il existe cependant certaines normes d'associations verbales sous forme de document ronéotypé (document interne d'équipes) : Lieury et Coll. [07], Meije [09] ou Ferrand et Alario (à paraître) [03].

automatique. Voir le tableau comparatif ci-dessous.

	N.Items	N.Sujets	Rap. H vs. F	Tps. Dépou	Mode Dép	Elèves	Lycéens	Etudiants	Salariés
Rosenzweig	100	288	non	~ 2 ans	manuel			x	
Oléron & Legal	238	219.5	non	~2 ans	manuel				x
Lieury & Coll.	122	297	non	~1 ans	manuel			x	
Meiye & Brulex	110	125	non	~4 mois	manuel			x	
Ferrand & Alario	366	89	non	~4 mois	manuel			x	

Fig. 2.1 – Tableau comparatif des normes associatives verbales du français

WebCompletion met à profit les technologies client-serveur<sup>4</sup> pour répondre de façon significative aux manques que nous venons d'énumérer.

### 3. HYPOTHESES

La progression fulgurante du réseau Internet à l'échelle planétaire conjuguée à celle qu'on observe en local au niveau de la socio-culture française doit permettre à WebCompletion de réaliser un corpus significativement représentatif sur les points suivants :

- population d'enquête de l'ordre du millier,
- rapport homme vs. femme,
- 4 catégories socioprofessionnelles (élèves, lycéens, étudiants et salariés),
- méthode de dépouillement automatique.

### 4. METHODE

Les caractéristiques de la méthode sont les suivantes :

- Nom de Domaine de l'Expérimentation (NDE)  
L'expérimentation est limitée au territoire français.
- Sujets  
Les utilisateurs de langue maternelle française, âgés de 8 à 85 ans et disposant d'un accès Web sur Internet.

- Matériel  
100 mots issus des 5 normes d'associations verbales du français.
- Centre d'expérimentation et URL (adresse)  
Centre Multicom du CLIPS-IMAG - <http://www-multicom.imag.fr/Normes>
- Procédure
  - Identification anonyme selon les critères : sexe (F et M), 5 tranches d'âge (8 à 85 ans) et 4 catégories socioprofessionnelles (élève, lycéen, étudiant, salarié)
  - Consigne : demander au sujet de lire le mot qui apparaît sur l'écran de l'ordinateur puis d'écrire immédiatement le premier mot qui lui vient à l'esprit, en le saisissant sur le clavier de l'ordinateur dans la zone prévue à cet effet. Une seule réponse est autorisée et la touche "Entrée" permet non seulement de valider la réponse, mais aussi de passer à l'item suivant.
  - Note spéciale de la consigne : Pour des raisons de biais expérimental, il est demandé à chaque sujet de ne pas passer deux fois cette expérience.
  - Randomisation : les mots n'apparaissent pas en ordre alphabétique, mais aléatoirement.
  - Durée de l'expérience : environ 10 minutes.
  - Phase de Pré-test : constituée de 15 mots à compléter. Elle sert à familiariser le sujet à la tâche.
  - Phase de Test : définit le test proprement dit. Elle est constituée de 125 mots à compléter.
- Méthode Action Serveur
  - CGI méthode Post-Action - le CGI (Commun Gateway Interface) définit la méthode qui permet de traiter les requêtes

<sup>4</sup> - Architecture logicielle dans laquelle une application s'exécute sur une machine (serveur) et répond à d'autres machines distantes (clients).

du client vers le serveur et inversement. Ce script, qui a une fonction de récupération et d'écriture des données, peut connaître des problèmes de performance lorsque le réseau est saturé, c'est pourquoi nous avons envisagé une deuxième méthode CGI fondée sur les Applets Java.

- CGI méthode Applets Java - contrairement à la première, cette méthode est basée sur le concept de transfert des savoir-faire ou extension client-serveur, à savoir que l'on dote le logiciel client (Netscape ou Internet Explorer) de certaines fonctionnalités fondamentales dont dispose le serveur, permettant d'exécuter des actions en local. L'avantage principal de cette méthode est double, d'une part se « couper » du serveur en exécutant localement une tâche et d'autre part, être à l'abri de tous les problèmes sous-jacents du trafic réseau (Meiye, Rouillard et Vaufreydaz, à paraître) [12].

- Connexion multi-utilisateur - "n" utilisateurs peuvent se connecter à un instant "t" et exécuter l'expérience simultanément et chacun à son rythme.

## 5. RESULTATS

Les principaux résultats concernent les points suivants :

### □ Dépouillement des données

- Chaque fichier résultat codé au format texte (.txt) est directement recodé au format du tableur Excel (.xls). A partir de là, des macros-commandes appropriées permettent de classer les données en fonction du critère item associé, sexe, âge et catégorie socioprofessionnelle.

- Les fichiers incomplets, à savoir plus de 10 réponses manquantes ou réponses spéciales dites HS (Hors Sujet) ne sont pas pris en compte.

- Chaque fichier est ensuite transféré directement dans le corpus général sur lequel les réponses des sujets sont calculées automatiquement.

### □ Tendances actuelles

- Les résultats auxquels nous sommes parvenus semblent confirmer en partie notre hypothèse de base, particulièrement en ce qui concerne l'augmentation significative du nombre de sujets, l'abaissement du temps investi dans le dépouillement des données (procédure automatique) et l'équilibre du rapport homme vs. femme dans le panel de la population d'enquête.

- En revanche, le problème de la représentativité des catégories socioprofessionnelles demeure. En effet, on s'aperçoit que seules les catégories socioprofessionnelles des étudiants et des salariés semblent significativement représentées contrairement à celles des élèves et des lycéens qui restent quasiment absentes (voir le tableau 5.1 ci-dessous).

	N.Items	N.Sujets	Rap. H vs. F	Tps. Dépou	Mode Dép.	Elèves	Lycéens	Etudiants	Salariés
Rosenzweig	100	288	non	~ 2 ans	manuel			x	
Oléron & Legal	238	219.5	non	~2 ans	manuel				x
Lieury & Coll.	122	297	non	~1 ans	manuel			x	
Meiye & Brulex	110	125	non	~4 mois	manuel			x	
Ferrand & Alario	366	89	non	~4 mois	manuel			x	
WebCompletion	125	> 1000	oui	30 sec.	auto			x	x

Fig. 5.1 – Tableau comparatif des normes associatives verbales du français

- Autrement dit, ce résultat suppose que la progression fulgurante du réseau Internet qu'on observe à l'échelle planétaire ne se retrouve pas au niveau des foyers français, sans doute à cause des coûts des communications qui restent encore élevés, et du fait que notre appel à participation s'est principalement porté sur le milieu universitaire.

## 6. CONCLUSION

□ La démarche que nous proposons dans ce domaine particulier de la recherche cognitive semble prometteuse. En effet, à côté des méthodes d'approches classiques, WebCompletion offre la possibilité au chercheur de réaliser un gain de temps considérable pour élaborer un corpus qui soit hautement significatif sur les critères suivants : le nombre de sujets, rapport homme vs. femme, le temps de dépouillement et le nombre de catégories socioprofessionnelles.

□ Mais il semble que la non représentativité de la catégorie socioprofessionnelle des élèves et lycéens soit liée au coût des communications et abonnements Internet encore trop élevé pour un foyer français moyen.

□ Comme l'a si bien noté Jodelet [05] dans sa revue critique, une norme associative peut intégrer plusieurs aspects dont la variation culturelle. En ce sens, lorsque Oléron et LeGall dans les années 1960-1961, demandent à leurs sujets (appelés militaires) quelle association ils font du mot "vérité" 85.2% des sujets répondent "Brigitte Bardot". 38 ans après, peut-on concevoir un instant cette réponse comme appropriée ? Mais l'intérêt de cet exemple doit nous sensibiliser sur le fait que les normes associatives sont finalement des images mentales propres à une socio-culture. Sur cette base, on peut donc dire qu'une norme associative qui ne prend pas en compte les images mentales des différentes catégories

socioprofessionnelles d'un groupe donné, n'est pas représentative de ce dernier.

□ Ainsi, en attendant que les coûts des communications baissent pour permettre d'une part une augmentation significative du nombre d'abonnements des foyers français sur Internet et d'autre part, de se doter d'une clé d'accès permettant de toucher massivement les catégories socioprofessionnelles des élèves et lycéens, une solution consiste à effectuer les mêmes expérimentations sur le terrain auprès de ces deux catégories, mais à partir d'un outil qui offre les mêmes avantages technologiques que WebCompletion. Cette solution complémentaire existe et nous l'avons simulée à travers le protocole d'expérimentation en local que nous avons baptisé ExpositiveEdit (Meiye, [10], Meiye et Coll. [11]).

## 7. BIBLIOGRAPHIE

- [01] COLLINS A. M., QUILLIAN M.R., *Retrieval time from semantic memory*, Journal of Verbal Learning and Verbal Behavior, 8, 240-247, 1969.
- [02] CONTENT A., MOUSTY P., RADEAU M., *Brulex, une base de données lexicales informatisée pour le français écrit et parlé*, l'Année Psychologique, 90, 551-566.
- [03] FERRAND L., ALARIO F-X., (à paraître), *Normes d'associations verbales pour 366 noms d'objets concrets*, l'Année Psychologique, 1998.
- [04] FRAISSE P., NOIZET G., FLAMENT C., *Fréquence et familiarité du vocabulaire*, in Problèmes de psycholinguistique, Paris, PUF, 157-167, 1963.
- [05] JODELET F., *L'association verbale*, In P. Fraisse et J. Piaget (Eds.), *Traité de Psychologie Expérimentale (Tome VIII, ch. 27)*, Paris, PUF, 93-137, 1965
- [06] LE NY P., *La sémantique psychologique*, Paris, PUF, 1979.
- [07] LIEURY A., IFF M., DURIS P., *Normes d'associations verbales*, Document ronéoté du Laboratoire de Psychologie Expérimentale et Comparée, 28 rue Serpente, 75006Paris, 1-39, 1976.
- [08] LOFTUS E.F., *Nouns, adjectives and semantic memory*, Journal of Experimental Psycholog.

- [09] **MEIYE J-P.**, *Normes associatives verbales du français, Comparatif et prise en compte de Brulex*, Document ronéoté du laboratoire CLIPS-IMAG, Equipe GEOD, Grenoble, 1-22, 1996.
- [10] **MEIYE J-P.**, *ExpositiveEdit, Cognitive Processing Test*, Document ronéoté et logiciel interne du laboratoire CLIPS-IMAG, Equipe GEOD, Grenoble, 1996.
- [11] **MEIYE J-P., CAELEN-HAUMONT G., DUBOIS D., JOUX C., (à paraître).**, *Normes associatives verbales du français : nouvelles méthodes d'approche.*
- [12] **MEIYE J-P., ROUILLARD J., VAUFREYDAZ D., (à paraître).**, *WebTDL, Simulation du paradigme de la décision lexicale sur le Web en modalité texte, image et son.*
- [13] **MEYER D.E.**, *On the representation and retrieval of stored semantic information*, *Cognitive Psychology*, 1, 242-300, 1970.
- [14] **OLERON G., LEGALL F.**, *Associations Verbales, Normes sur 238 mots*, Document ronéoté du Laboratoire de Psychologie Expérimentale et Comparée, 28 rue Serpente, 75006Paris, 1-24, 1961-1962.
- [15] **ROSENZWEIG M.R.**, *Etudes sur l'association des mots*, *l'Année Psychologique*, 57, 1, 23-32, 1957.