

Commentaire sur la réponse de Clémence concernant le texte de Lacassagne, Salès- Wuillemin, Castel & Jebrane (2001)

Edith Salès-Wuillemin
Philippe Castel
Marie-Françoise Lacassagne

Groupe de Psychologie Sociale des Discours (GPSD)
Laboratoire de Psychologie Clinique et Sociale (LPCS)
Université de Bourgogne,
36 rue Chabot-Charny,
21000 Dijon, France
E-mail: Edith.Sales-Wuillemin@u-bourgogne.fr

Ahmed Jébrane

Laboratoire de Topologie, UMR 5584
Département de Mathématiques
Université de Bourgogne, BP 138
21000 Dijon Cedex, France

Le commentaire réalisé par Clémence appelle deux réponses: la première concerne la catégorisation sociale et plus précisément: 1- des éléments d'explication concernant l'étape du processus où nous nous situons dans l'article cité en référence; 2- la mise en évidence des biais d'assimilation et de contraste grâce à une analyse statistique des listes de mots induits produits par les sujets dans les deux conditions.

La deuxième concerne plus spécifiquement les fondements sur lesquels repose la méthode utilisée pour réaliser le codage des données. Ce deuxième volet de notre explication a nécessité la réalisation d'une étude complémentaire que nous présentons de manière synthétique directement dans cet écrit.

Processus de catégorisation et biais perceptifs

Contrairement à ce que souligne Clémence, le but de cette étude n'était pas de mettre en évidence le processus de catégorisation et les distorsions perceptives qui en découlent.

L'objectif était double: il s'agissait d'une part de montrer que le processus de catégorisation, et donc la mobilisation d'une perception catégorielle, a une incidence directe sur les traits associés aux éléments dont l'appartenance catégorielle est déjà connue du sujet; il s'agissait d'autre part de montrer que ces traits peuvent être analysés comme des traces de l'existence des biais de distorsion perceptive auxquels nous faisons référence (assimilation, contraste et discrimination).

Nous sommes ainsi bien là dans le cadre d'étude de la catégorisation sociale; simplement, ce qui est appréhendé n'est pas un processus mais les conséquences de sa mise en œuvre sur la représentation sociale mobilisée.

Par ailleurs, contrairement à ce que souligne Clémence, il s'agissait bien d'une situation dans laquelle une perception catégorielle était mobilisée. En effet, nous précisions par écrit aux sujets des deux conditions que l'étude portait sur la perception que les groupes d'individus interrogés pouvaient avoir de différents groupes ethniques et nationaux, c'est donc bien l'appartenance catégorielle des sujets qui était mobilisée par l'intermédiaire de cette consigne, la tâche implicite étant une tâche de comparaison catégorielle. Même si, comme le souligne Clémence, chacun des deux groupes ne savait pas à quelle catégorie l'autre était confronté.

Enfin, seules les réponses des étudiants d'origine nationale française ont été retenues.

L'ensemble de ces éléments nous permet d'attester que l'étude réalisée porte sur une partition catégorielle, elle met en regard les réponses des sujets ayant pour tâche de définir leur endogroupe (celui des français) ou un exogroupe (celui des maghrébins).

L'analyse statistique que nous faisons pour mettre en évidence les biais perceptifs a fait également l'objet de remarques de la part de Clémence. Nous reprendrons les éléments relevés point par point:

Indicateurs du consensus (assimilation): nous avons considéré que le biais d'assimilation se révélait de façon d'autant plus marquée qu'un nombre significatif d'individus citaient le même trait en réaction au mot stimulus. Le seuil choisi était de 10%. Nous avons précisé que la probabilité que 10% des sujets produisent le même mot est très faible, et que la probabilité que ces sujets en citent 34 était quasiment nulle.

Il est vrai, comme le constate Clémence, que nous avons appliqué un strict calcul de probabilité sans prendre en compte le fait que les probabilités ne peuvent, dans ce cas, se multiplier arithmétiquement, comme si chaque "tirage" était indépendant du précédent. Il existe effectivement dans le cadre d'une tâche d'associations verbales, un lien sémantique très fort et difficilement quantifiable, entre les différents mots induits produits par les sujets. La formule mathématique proposée ne prend pas en compte cette donnée. Une analyse plus approfondie montre néanmoins que l'application d'une formule plus ajustée (même si elle avait existé) n'aurait pas changé les conclusions.

Si nous nous plaçons en effet dans une situation (hautement improbable) où (a) le premier mot détermine totalement tous les autres mots de la liste, et donc si nous ne prenons en compte que la probabilité que 10% des sujets produisent ce mot et (b) que nous prenons le parti de n'utiliser pour ensemble de référence que les mots induits effectivement produits par les sujets et non pas la totalité du lexique de la langue française, ce qui nous met dans une position défavorable, cette probabilité est extrêmement faible: nous obtenons une probabilité avoisinant

.005 pour le mot stimulus "Français" et une probabilité de .01 pour le mot stimulus "Maghrébins"¹

Si l'on prend maintenant en compte le nombre de mots différents produits par 10% des sujets dans la condition mot inducteur "français" leur nombre est de 34 (soit 7.11% des mots produits). Ce qui est en soi révélateur d'un très fort consensus. A l'opposé, dans la condition mot inducteur " maghrébins " ils ne sont qu'au nombre de 17. Ce total est deux fois moins élevé que dans l'autre situation, même s'il représente 4.48% des mots cités. Nous ne pouvons que conclure à un plus fort consensus dans le premier cas que dans le second cas, même si la différence n'est que tendanciellement significative (test de comparaison des proportions $Z = 1.615$; $p=0.10$; test de Khi2 = 2.608; $p=0.10$).

Indicateurs de l'opposition (contraste): nous avons considéré que le biais de contraste se révélait de façon d'autant plus marquée que les mots induits apparaissant dans une condition n'apparaissaient pas dans l'autre, ce qui se traduisait au niveau statistique par le recours à un indice de corrélation négatif.

Clémence fait remarquer que la démonstration aurait paru plus évidente si les mots induits produits dans les deux conditions avaient été des antonymes. Ce point de vue présuppose un point de référence commun aux deux groupes. Or ce que montre l'analyse de nos résultats c'est que les mots induits dans les deux situations font plus que s'opposer, ils sont produits en regard de points de référence totalement différents. C'est un résultat tout à fait en concordance avec ceux que nous avons obtenus dans deux études récentes qui faisaient varier les différentes dénomination d'un même référent dans une tâche d'associations verbales (Castel, Lacassagne, Salès-Wuillemin (2002) ; Salès-Wuillemin, Castel, Lacassagne (sous-presse)). L'analyse montre que si l'on fait simplement varier la dénomination de l'objet (au travers du mot inducteur), cela active chez le sujet des représentations contrastées de l'objet, ces représentations traduisant des points d'ancrages différents.

Sens et catégories syntaxiques

Clémence fait remarquer que nous avons appliqué "de manière rigide une norme de définition de termes grammaticaux (en définissant de manière grammaticalement rigide substantifs, adjectifs et verbes)" et en oubliant que "dans le sens commun, ces règles deviennent beaucoup plus flexibles, tout simplement parce que le sens l'emporte sur la règle syntaxique".

Cette "rigidité" nous avait conduit à n'opérer aucun regroupement sémantique: les données n'avaient pas été "lemmatisées" (opération qui consiste à regrouper les mots ayant la même racine, sans distinguer les catégories syntaxiques). L'analyse avait été réalisée sur la base du corpus initial. Ce choix de traitement génère une certaine lourdeur au niveau de l'analyse des résultats, mais avait à nos yeux le précieux avantage de ne pas introduire de distorsions dans les données.

Pour apporter des éléments de justification concernant les choix opérés, nous avons réalisé une étude complémentaire.

¹ Les sujets ont produit 478 associations à partir du mot Stimulus " Français" . 34 le sont par plus de 10 pour cent des sujets. Chaque sujet a proposé 15 mots au total . La probabilité pour qu'un sujet ait utilisé un mot (par hasard) sur les 478 différents produits, est au maximum de 15/478. Le nombre X de sujets parmi les 50 qui ont utilisé ce mot varie comme une loi binomiale B(50,15/478). La probabilité pour avoir 6 sujets ou plus qui utilisent ce mot est la somme des probabilités $p[X = k]$ avec $k=6,7,\dots,50$ soit $C(50,k) (p)^k (1-p)^{(50-k)} = 0.004646$ ce qui est de l'ordre de 0.005.

Les sujets ont produit 379 associations à partir du mot stimulus "Maghrébins ". 17 l'ont été par plus de 10 pour cent des sujets. Le même calcul que précédemment conduit à la binomiale B(50,15/379). On obtient une probabilité de 0.01374

Méthode

L'étude s'est déroulée en deux temps². Dans la première phase, un échantillon restreint de sujets a été sollicité pour donner, au moyen d'une épreuve d'associations verbales, tous les mots et expressions qu'ils considéraient comme "synonyme" du mot "maghrébin".

Au terme de cette phase nous avons retenu 12 mots différents: islamiste, musulman, intégriste, arabe, rebeu, beur, bougnoul, gris, immigré, nord-africain, étranger, clandestin (sélectionnés sur la base de leur fréquence).

650 sujets ont participé à la deuxième phase de l'étude. Ils étaient comparables à ceux qui avaient participé à la première phase. Ces sujets ont été répartis aléatoirement dans 13 groupes indépendants. Dans chacun de ces groupes, la tâche consistait à réaliser une association verbale contrainte à partir de "maghrébin" ou de l'un de ses 12 synonymes. Chaque sujet devait produire 5 substantifs, 5 adjectifs et 5 verbes à partir du mot stimulus.

Méthode d'analyse des résultats

Sur la base des mots induits produits par les sujets en fonction de leur catégorie syntaxique (substantifs, adjectifs, verbes), trois Analyses Factorielles de Correspondances (A.F.C) ont été réalisées. L'objectif étant de montrer que les places occupées par les mots inducteurs (et donc les oppositions qu'ils cristallisent) varient en fonction de la catégorie syntaxique des induits.

L'analyse que nous présentons ici se focalise sur l'axe 1 des oppositions (qui représente à lui seul 32,7% de la variance pour l'AFC réalisée à partir des substantifs induits, 35,6% de la variance pour l'AFC réalisée à partir des adjectifs induits, et 34,1% de la variance pour l'AFC réalisée à partir des verbes induits). Les coordonnées de chaque inducteur ont été répertoriées au sein de chacune des AFC réalisées.

Tableau 1
Coordonnées de chaque mot inducteur sur l'axe 1 (indépendamment du poids)

	substantifs	adjectifs	verbes
arabe	-0,36	-0,26	-0,05
beur	0,14	-0,18	0,24
bougnoul	0,04	-0,02	0,19
clandestin	1,96	1,64	0,24
étranger	0,53	0,06	-0,03
gris	0,32	-0,08	0,20
immigré	0,42	0,15	0,58
intégriste	-0,50	-0,79	-1,20
islamiste	-0,85	-0,75	0,23
maghrébin	0,18	-0,07	0,15
musulman	-0,80	-0,41	-0,92
nord-africain	0,51	-0,075	-0,16
rebeu	-0,15	-0,06	0,25

Une fois les coordonnées identifiées, elles ont été transformées en rang sur l'axe 1 des oppositions. Le mot inducteur ayant les coordonnées les plus négatives et donc situé le plus à gauche de l'axe, étant côté 1, le plus positif et donc situé le plus à droite de l'axe, étant côté 13.

² Pour une présentation détaillée du dispositif utilisé dans cette étude cf. Salès-Wuillemin, Castel, Lacassagne (sous-presse)

Un coefficient de corrélation sur les distributions de rang a été ensuite calculé (Spearman), ce qui nous a permis de comparer l'organisation des distributions de rang pour chacune des catégories syntaxiques de mots induits.

Tableau 2
Corrélation entre les distributions de rang

Distributions de rang comparées	coefficient de corrélation (Spearman)	seuil de probabilité p.
Substantifs / adjectifs	.857	p=.002
Adjectifs / verbes	.549	p=.051
Verbes / substantifs	.258	p=.390

Le Tableau 2 montre que la distribution des coordonnées des inducteurs activant des induits substantifs comparée à cette même distribution avec des induits adjectifs, est corrélée fortement et significativement. La même comparaison réalisée sur la distribution des verbes donne une corrélation qui n'est pas significative. La corrélation est tendancielle entre les adjectifs et les verbes. On peut donc conclure que la distribution des positions prises par les mots inducteurs au sein de l'AFC faite sur les verbes est différente des deux autres listes et en particulier de celle faite sur les substantifs. A titre illustratif, le mot inducteur "islamiste" a la coordonnée la plus négative (-0.85, rang 1) avec les substantifs, alors qu'il passe sur le versant positif (+0.23, rang 9) avec les verbes.

Plus largement, si l'on considère, pour ce qui concerne les mots induits de la catégorie des substantifs, les positions respectives prises par les mots inducteurs on note que le bloc "islamiste" (-0.85) et "musulman" (-0.80) s'oppose sur l'axe 1 à "clandestin" (+1.96).

Si l'on opère de même sur l'AFC des adjectifs, c'est le bloc "intégriste" (-0.79) et "islamiste" (-0.75), qui s'oppose à "clandestin" (1.64). Enfin, pour les verbes, c'est une autre configuration qui apparaît: le bloc "intégriste" (-1.20) et "musulman" (-0.92) s'oppose à "immigré" (+0.58). En d'autres termes, selon que l'on considère l'une ou l'autre des catégories de mots induits, l'opposition varie fortement. Pour les substantifs et les adjectifs elle se fait entre un ensemble que l'on pourrait dénommer "religion" qui se situe à un extrême du continuum, et "clandestin" à l'autre extrême. Si l'on considère par contre les verbes induits, on retrouve bien à un extrême l'ensemble "religion" mais il s'oppose cette fois-ci à "immigré". Or c'est cette opposition qui va nourrir la signification attribuée à chacun des termes qui la composent.

En conclusion, l'analyse des résultats montre que les traits associés aux éléments d'une catégorie ne reposent pas sur les mêmes oppositions selon que le sujet considère les actions des membres de la catégorie (au moyen de verbes), les caractéristiques des membres de cette catégorie (au moyen d'adjectifs), ou la catégorie en elle-même et dans son ensemble (au moyen de substantifs).

Il ressort de cette analyse que si le mot est en lui-même porteur d'une signification, la catégorie syntaxique à laquelle il appartient véhicule un autre type de signification qui met en saillance des dimensions bien particulières de l'objet et donc réorganise l'axe des oppositions dont rend compte l'AFC.

Le choix réalisé de n'opérer aucun regroupement des termes y compris à partir de leur racine commune se trouve ici pleinement justifié ; prendre cette précaution c'est en effet ne pas oublier que si " le sens peut parfois l'emporter sur la stricte règle syntaxique ", la forme adoptée par le référent est en soi révélatrice de la place qui lui est accordée et des composantes qui sont mises en saillance (volontairement ou pas) par le sujet.

Références bibliographiques

- Castel, P., Lacassagne, M.F., Sales-Wuillemin, E. (2002). Categorical points of view in social representation. *Language Sciences*, 24, 667-678.
- Lacassagne, M.-F., Salès-Wuillemin, E., Castel, Ph. (2001). La catégorisation d'un exogroupe à travers une tâche d'association de mots, *Papers on Social Representations, Textes sur les Représentations Sociales*, 10, 7.1-7.11.
- Sales-Wuillemin, E., Castel, P., Lacassagne, M.-F. (Sous presse) L'étude des représentations sociales: effet du mot inducteur sur les éléments activés lors d'une tâche d'association verbale, *European Journal for Semiotic Studies*.