



HAL
open science

Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse

Madeleine Akrich

► **To cite this version:**

Madeleine Akrich. Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse. 2012. halshs-00702086

HAL Id: halshs-00702086

<https://shs.hal.science/halshs-00702086>

Preprint submitted on 29 May 2012

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.

N° 025

2012

Les listes de discussion comme communautés en ligne :
outils de description et méthodes d'analyse

Madeleine Akrich
Centre de Sociologie de l'Innovation
Mines ParisTech
madeleine.akrich(a)mines-paristech.fr

CENTRE DE SOCIOLOGIE DE L'INNOVATION
MINES PARISTECH / CNRS UMR 7185
60 Boulevard Saint-Michel
75272 Paris cedex 06 FRANCE
<http://www.csi.mines-paristech.fr/>

PAPIERS DE RECHERCHE DU CSI

Cette collection a pour but de rendre aisément disponible un ensemble de documents de travail et autres matériaux de discussion issus des recherches menées au CSI (CENTRE DE SOCIOLOGIE DE L'INNOVATION).

Tous les droits afférant aux textes diffusés dans cette collection appartiennent aux auteurs.

Des versions ultérieures des papiers diffusés dans cette collection sont susceptibles de faire l'objet d'une publication. Veuillez consulter la base bibliographique des travaux du CSI pour obtenir la référence exacte d'une éventuelle version publiée.

CSI WORKING PAPERS SERIES

The aim of this collection is to make easily available a set of working papers and other materials for discussion produced at the CSI (CENTRE DE SOCIOLOGIE DE L'INNOVATION).

The copyright of the work made available within this series remains with the authors.

Further versions of these working papers may have been submitted for publication. Please check the bibliographic database of the CSI to obtain exact references of possible published versions.

CENTRE DE SOCIOLOGIE DE L'INNOVATION
MINES PARISTECH / CNRS UMR 7185
60 Boulevard Saint-Michel
75272 Paris cedex 06 FRANCE
<http://www.csi.mines-paristech.fr/>

Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse

Madeleine Akrich

Partie I

La question des méthodes dans l'analyse des groupes électroniques de discussion

« A quelles conditions les pratiques de communication au sein de cette entité imaginée que constitue la « communauté virtuelle » pourraient-elles conduire à un engagement social, culturel ou politique des individus qui s'y identifient ? A quelles conditions serait-il possible de dépasser le simple niveau d'un « engagement électronique » des usagers du collectif en ligne pour atteindre celui des prises de décision durable dans la communauté ancrée géographiquement, celui de la formulation des projets sociaux et politiques concrets et, par conséquent, d'une action effective dans la Cité ? »

Cette interrogation, qui conclut le chapitre proposé par Serge Proulx (2006), a été la nôtre depuis qu'avec Cécile Méadel, nous avons commencé à étudier les listes électroniques de discussion concernant des thématiques de santé. A la lecture des échanges de nombreux groupes, nous avons en effet le sentiment que s'élaborait là quelque chose qui allait au delà d'une conversation – par essence éphémère et sans autre visée qu'elle-même – entre des personnes qui partagent des intérêts communs : on y découvrirait des personnes à la fois soucieuses de leurs interlocuteurs et profondément engagées dans un travail d'investigation personnelle et collective.

Comment rendre compte de la dynamique informationnelle, cognitive, politique à l'œuvre dans ces groupes ? Comment rendre compte, en amont de cela, de la composition – au sens d'un processus actif et constamment réactivé – du groupe lui-même ? La tâche est de fait ardue ; passé le moment d'émerveillement face à la richesse des échanges, le chercheur se trouve vite enseveli par un matériau qui déborde de toutes parts, comme le raconte Nancy Baym (2000a) dans son livre sur un groupe de fans de soap opera :

"Once data were collected, I immediately faced a second problem: I had way too much data. Although it was clear to me that analyzing so many messages closely was impossible, the best ways in which to narrow them down were not immediately apparent. The challenge was to narrow the data down in a way that retained the coherence of both the group and of the discussion." p. 26

Les méthodes apparaissent donc cruciales pour être en mesure d'apporter un début de réponse aux questions soulevées. D'où l'intérêt de reprendre la question à partir des travaux existants. Au delà de cette première motivation, le choix des méthodes comme point d'entrée dans l'analyse des groupes électroniques de discussion relève d'un parti pris de notre part : si j'en crois l'échantillon dont je dispose sur mon ordinateur et qui a été prélevé en interrogeant des bases de données avec des termes comme « discussion

group » « discussion list » « forum », 90% de la littérature consacrée aux groupes électroniques de discussion utilise le mot « communauté » ou « community ». Il y a certes eu des discussions sur l'intérêt de cette notion, sur son caractère approprié, sur ce qu'elle charrie éventuellement avec elle comme théorie implicite¹ etc. Cependant, tout ce travail analytique n'a pas – semble-t-il – eu d'effet notable sur l'utilisation de la notion, si ce n'est de contribuer à la banaliser. Or, au delà de ce consensus apparent, il existe au fond des manières extrêmement différentes d'envisager ces groupes : l'entrée par les méthodes vise à interroger la manière dont, en pratique, les chercheurs construisent leur objet. Notre pari est qu'au travers de ce prisme apparaîtront de nouvelles lignes de démarcation, autrement peu visibilisées.

Bien sûr, tous les chercheurs ne visent pas les mêmes objectifs en s'intéressant aux groupes électroniques de discussion ; comme on le verra dans la suite, on peut distinguer grosso modo trois postures :

Un certain nombre d'auteurs, plutôt des sociologues, partent d'une question exogène par rapport au média – par exemple s'interrogent sur le passage du patient d'une position passive à une position active (Fox *et al.* 2005), ou sur la mobilisation des connaissances par les patients dans leur évaluation d'un traitement (Wikgren 2001) – et vont utiliser le média comme intermédiaire commode permettant d'avoir accès aux acteurs, sans oublier de se demander au passage, mais plutôt comme une question annexe, ce que le média est susceptible d'apporter spécifiquement comme transformation par rapport à la question posée.

D'autres auteurs, davantage intéressés par les sciences de la communication ou les linguistes, considèrent les groupes électroniques de discussion comme une modalité de communication comparable à d'autres, comme la conversation en face à face, l'échange épistolaire, le cours dispensé en présentiel etc. Ils vont plutôt s'employer à caractériser ce média par rapport à d'autres.

Enfin, les derniers, des sociologues, des chercheurs en communication et des informaticiens se concentrent sur l'objet « groupe de discussion », sur son fonctionnement, son organisation, sur ce qu'il fait émerger de spécifique et déploient leurs efforts vers la production de descriptions qui rendent compte des caractéristiques de cet objet. Ces derniers sont souvent impliqués non seulement en tant qu'observateurs des groupes mais aussi en tant que concepteurs de systèmes permettant d'en améliorer l'utilisation.

Nous allons donc traverser des disciplines et des champs de recherche sans manifester un respect démesuré des frontières supposés les séparer. À l'inverse, nous considérons que notre parti pris présente l'avantage si ce n'est de faire dialoguer ces divers courants de recherche, du moins d'essayer de saisir les différentes modalités par lesquelles se construit un objet de recherche.

¹ Par exemple : Blanchard (2004) ; Jones *et al.* (2004) ; Kling & Courtright (2003) ; Papadakis (2003) ; Proulx (2006) ; Wellman & Gulia (1999) ; Wilson & Peterson (2002)

Comment aborder donc les corpus constitués par ces groupes de discussion ? En première instance, on pourrait les considérer comme du texte : les sciences humaines et sociales ayant développé depuis longtemps des techniques d'analyse adaptées, le chercheur ne serait pas démuni face à ce matériau. Néanmoins, en plus d'être pléthorique, celui-ci présente la caractéristique d'être un matériau complexe et « sale » : l'interactivité propre au media se traduit notamment dans l'utilisation de techniques de reprise et de citation variées (Mondada 1999), qui sont en partie préinscrites dans les systèmes de gestion des forums et des listes et dans les logiciels de courrier électronique (Blanchard 2004), mais font aussi l'objet d'innovations de la part des utilisateurs. Ces pratiques ont pour conséquence une très grande redondance du matériau brut et des difficultés non négligeables pour en extraire les parties pertinentes.

Cependant, cette réduction du corpus au texte ne tient pas longtemps pour qui s'intéresse à ce que construit la circulation des messages : un message est certes un objet textuel, mais c'est aussi quelque chose qui circule dans un espace défini à l'initiative d'une personne donnée dans un temps donné et à l'intérieur d'une séquence d'autres messages, toutes dimensions manifestées par un certain nombre d'attributs du message présents dans les en-têtes. Il faut donc trouver le moyen de tenir ensemble – même de manière lâche – les deux faces du message, le message comme contenu et le message comme relation dynamique.

L'analyse de l'émergence de réseaux socio-sémantiques pourrait constituer une forme de réponse à ce défi : elle commence à se développer dans le domaine des blogs (Cointet 2009 ; Roth & Cointet 2010), s'inspirant en partie des travaux de la sociologie des sciences, de la scientométrie et de l'analyse des réseaux sociaux. Mais comme nous allons le voir, ces développements n'ont pas encore trouvé leur traduction dans le domaine qui nous intéresse.

L'analyse des groupes électroniques de discussion se trouve donc face à un double défi méthodologique : celui constitué par la masse de données d'une part et d'autre part celui qui tient à la double nature du message, à la fois contenu et relation.

Dans la suite, je vais m'attacher à décrire un panorama des méthodes qui ont été développées par les chercheurs en sociologie ou en communication pour analyser les groupes électroniques de discussion en essayant de voir quelles réponses ils ont tenté d'apporter à ces défis : sans anticiper sur la description, disons qu'en première approximation, les méthodes en question se divisent en deux grandes catégories, celles qui s'intéressent au contenu de ce qui s'échange et celles qui s'intéressent au réseau de relations qui résulte des échanges ; rares sont les travaux qui à défaut d'articuler complètement ces deux types d'approches les mobilisent au moins de manière conjointe. En d'autres termes, le social et le sémantique sont encore à réconcilier dans les études sur les groupes de discussion.

Quels objets d'étude ?

La littérature qui traite de ces questions s'intéresse en fait à trois formes d'échange électronique collectif : les **bulletin boards** (newsgroup usenet), les **listes de discussion** et les **forums**. La question des spécificités de chacun de ces supports et du coup, du caractère plus ou moins transposable des méthodes d'analyse, n'est en fait pas ou peu abordée par les travaux dans ce domaine : Josiane Jouët (1989) et Nancy Baym (1998) sont parmi les rares auteurs à attirer l'attention sur l'importance des caractéristiques techniques des systèmes (temporalité, interfaces avec l'utilisateur). Les différences entre ces trois types de groupes tiennent pour une part aux protocoles techniques qui sont utilisés en « back office », et pour une part aux interfaces qui sont proposées aux utilisateurs. Nous nous concentrerons sur ces dernières :

- les bulletin boards et les forums sont consultables en ligne, alors que les abonnés aux listes de discussion reçoivent les messages sur leur boîte mail, ce qui induit une implication différente de l'utilisateur :

"Bulletin board systems (BBSs—also known as conferencing systems) are another form of asynchronous communication that refine email discussion lists in a number of ways. Most BBSs allow participants to create topical groups in which a series of messages, similar to email messages, can be strung together one after another. (...) BBSs differ from email discussion lists in another way. Email is a "push" media—messages are sent to people without them necessarily doing anything. In contrast, conferencing systems are "pull" media—people must select groups and messages they want to read and actively request them." (Kollock & Smith 1999)

- Cependant, il est possible aussi de consulter les listes de discussion sur un site ; par ailleurs, les groupes usenet – repris par Google – permettent la réception d'emails récapitulatifs qui reprennent l'ensemble des messages postés dans la journée ; enfin, un grand nombre de forums proposent de recevoir des alertes mails quand de nouveaux messages sont postés sur certains sujets. Les différences qui, à l'origine, étaient très marquées et supposaient des formes d'engagement contrastées des participants s'effacent en partie : comme l'utilisateur de listes, l'utilisateur de forums peut déléguer au système le soin de l'intéresser et n'a plus forcément à entrer dans une démarche proactive de connexion alors que, symétriquement, l'utilisateur de listes lassé de voir sa boîte encombrée de messages peut consulter sa liste comme un forum.
- L'affichage des messages sur un forum ou un newsgroup est organisé par fil ou sujet de discussion actif, alors qu'il est généralement chronologique dans le cas des listes ; cependant sur l'interface web des listes, il est possible d'adopter une présentation similaire à celle des forums.
- Toutes les listes de discussion supposent une inscription afin de recevoir les mails et de pouvoir poster ; une grande partie d'entre elles réservent par ailleurs la consultation de leurs archives à leurs abonnés : elles constituent donc souvent des

espaces privés. A l'inverse, les newsgroups ne nécessitent généralement pas d'inscription et sont publics alors que les forums se trouvent dans une situation intermédiaire : il faut souvent s'inscrire pour pouvoir intervenir même si les archives sont publiques ; certains sont privés et accessibles seulement aux inscrits. Là encore, plutôt qu'une opposition franche entre deux modèles, on observe un continuum de situations, allant des espaces les plus clos aux espaces les plus ouverts.

- Du coup, les modalités d'appréhension des participants les uns par les autres peuvent être a priori assez différentes : la fréquentation assidue d'un même groupe fermé permet de construire des repères et de différencier les intervenants, alors qu'un forum pourrait apparaître comme un espace où se croisent des inconnus qui échangent brièvement et s'éloignent aussi vite. Si tel était le cas, le contenu même de chaque message serait amené à jouer un rôle déterminant sur l'échange dans les forums, alors que l'accumulation des contacts passés, c'est-à-dire la connaissance que les participants ont les uns sur les autres, pourrait avoir une certaine influence sur les échanges des listes. En conséquence, la capacité à nouer des relations durables serait elle aussi contrastée selon les supports. En pratique cependant, au delà de tous les éléments d'identification que n'importe quel participant à un newsgroup fournit au travers de sa signature, de son ou ses adresses mails etc. (Donath 1999), des outils se sont progressivement mis en place sur les forums comme sur les newsgroups qui permettent « d'accrocher » au message des éléments d'identification de son auteur : catégorisation en fonction du nombre de messages postés, accès à un profil, reconstitution d'un historique de participation (usenet). Ce qui a permis de constater que beaucoup de ces forums ou groupes sont fréquentés assidûment par un certain nombre de participants : non seulement les locuteurs ne sont pas des anonymes interchangeable mais ils constituent des repères dans des espaces qui se spécifient dans le même temps. En définitive, que les espaces soient ou non clôturés par des barrières techniques, ils finissent souvent par s'organiser autour d'un groupe de participants identifiés.
- Les newsgroups (usenet) présentent cependant des particularités liées à l'histoire et à l'organisation des groupes (Pfaffenberger 1996) : sans entrer dans les détails, disons que l'ensemble des groupes est structuré à l'intérieur d'un espace public unique dont les lignes internes sont plutôt de l'ordre de la démarcation que de l'ordre de la frontière ; peu d'entre eux sont modérés et, à la différence des listes de discussions, aucun n'appartient à un propriétaire qui aurait la possibilité de prendre des décisions radicales comme la fermeture du groupe par exemple. Cet ensemble de caractéristiques a favorisé le développement de pratiques spécifiques, notamment le fait que beaucoup d'utilisateurs interviennent sur plusieurs groupes à la fois, sans que ces groupes aient forcément de rapport thématique entre eux : il y a donc pour les utilisateurs un monde global Usenet qui fait sens et dans lequel ils se déplacent à leur guise.

Cette discussion est motivée par le fait que, pour l'essentiel, les travaux qui se sont attachés à explorer la dimension relationnelle des groupes, qui ont développé des approches quantitatives et proposé des outils de visualisation de l'activité des groupes, se sont appuyés sur les newsgroups Usenet². Cette focalisation sur Usenet peut être mise en relation avec plusieurs éléments : d'abord l'antériorité des newsgroups, leur développement spectaculaire³, la disponibilité des données du fait du caractère public des échanges, l'existence d'un outil, Netscan, permettant une première exploration statistique des groupes. Cet outil a été élaboré par un sociologue, Marc Smith, travaillant pour Microsoft⁴. Il était destiné à aider l'internaute à se repérer dans la multitude des groupes, des participants et des fils de discussion. Cet outil a été mis à disposition du public entre 1999 et 2007, date à laquelle Microsoft a décidé de le retirer, semble-t-il en raison du déclin relatif des newsgroups.

² On peut citer, relevant de ce genre : Arguello *et al.* (2006) ; Beaudouin & Velkovska (1999) ; Brush *et al.* (2005) ; Burke, M. *et al.* (2007) ; Burnett, Gary & Bonnici (2003) ; Burnett, Gary & Buerkle (2004) ; Fiore *et al.* (2002) ; Fisher (2005) ; Fisher *et al.* (2006) ; Gleave *et al.* (2009) ; Golder & Donath (2004) ; Himelboim, Itai *et al.* (2009a) ; Himelboim, Itai. *et al.* (2009b) ; Jones *et al.* (2004) ; Raban & Rabin (2009) ; Sack (2002) ; Shumar & Smith (2004) ; Smith, C. B. *et al.* (1998) ; Smith, M. A. & Fiore (2001) ; Turner, Tammara Combs & Fisher (2006) ; Turner, T. C. *et al.* (2005) ; Viegas, F. B. & Smith (2004) ; Welser, H. *et al.* (2007) ; Welser, H. T. *et al.* (2009) ; Whittaker *et al.* (1998).

³ Alors qu'en 1980, environ 10 messages étaient échangés par jour, en 1988, on en comptait 1800 (ou articles selon la terminologie consacrée), ce qui paraissait énorme (<http://communication.ucsd.edu/bjones/Usenet.Hist/Nethist/0014.html>), et en 1998, il y avait 17000 groupes et 3 millions d'utilisateurs (Whittaker *et al.* 1998) ;

⁴ Sur la construction de l'outil, voir Smith, M. (1999)

Statistiques proposées par Netscan :

Sur les groupes : pour des périodes dont la longueur est ajustable par l'utilisateur, le système fournit le nombre de messages, de participants (posters), de participants déjà présents sur la période précédente, le nombre de réponses (messages envoyés en réponse à un autre message), le nombre de messages sans réponse, le nombre de répondeurs (auteurs qui ont envoyé au moins un message de réponse), la longueur moyenne (en lignes) des messages, les groupes proches (c'est-à-dire qui ont des messages communs : phénomène du cross-posting quand un message est envoyé simultanément sur plusieurs groupes) ; le nombre de messages en commun. L'ensemble de ces données pouvait faire l'objet d'une comparaison sur deux périodes choisies par l'utilisateur. Une partie d'entre elles était aussi accessible sous la forme de graphiques.

The screenshot shows the Netscan website interface. At the top, there are navigation links: Daily Technical Newsgroups, My Usenet, My Settings, NG Search, NG Report, Author Profile, Thread View, Treemap, About, and Help. Below this is a search bar with the text "Discuss Netscan microsoft.public.research.netscan.discussion" and a "Send us your Feedback" link. The search filters section includes "Report Time Span" with options for Day, Week, Month, Quarter, and Year. The search criteria are "Newsgroups where name contains cultur" and "For date mm/dd/yyyy: 5/31/2007". There is a "Search" button and a "Save Search" checkbox. Below the search filters, the results are displayed for the month of May, comparing to the previous time period. The table has columns for Posts, Posters, Returnees, Replies, Repliers, URM, AdjURM, AvgLineCt, XPosts, XPTgs, and Authors. The table lists 414 newsgroups, with the top results being soc.culture.israel, soc.culture.jewish, soc.culture.usa, it.cultura.single, soc.culture.belgium, soc.culture.romanian, soc.culture.scottish, soc.culture.scottish, soc.culture.singapore, soc.culture.iranian, soc.culture.polish, soc.culture.china, soc.culture.indonesia, soc.culture.argentina, soc.culture.canada, soc.culture.venezuela, soc.culture.turkish, it.cultura.religioni, soc.culture.quebec, soc.culture.europe, soc.culture.indian, soc.subculture.bondage-bdsm, soc.culture.palestine, soc.culture.malaysia, soc.culture.cuba, soc.culture.vietnamese, soc.culture.british, soc.culture.spain, soc.culture.iraq, and soc.culture.russian.

414 Newsgroups	Posts	Posters	Returnees	Replies	Repliers	URM	AdjURM	AvgLineCt	XPosts	XPTgs	Authors
soc.culture.israel	14591	783	269	11956	682	1287	1256	64	13776	246	Metrics
soc.culture.jewish	14276	929	330	12515	822	707	695	59	14026	288	Metrics
soc.culture.usa	13185	1704	619	9572	1410	2412	2398	82	12830	526	Metrics
it.cultura.single	10974	230	92	9995	197	212	206	20	879	34	Metrics
soc.culture.belgium	10635	390	171	9892	356	146	143	33	4144	88	Metrics
soc.culture.romanian	8891	174	46	3364	133	5162	5161	1189	5812	40	Metrics
soc.culture.scottish	8058	346	157	7600	324	149	147	48	6756	75	Metrics
soc.culture.singapore	7930	540	185	6579	396	572	569	64	4696	136	Metrics
soc.culture.iranian	6555	516	191	5212	432	344	343	67	4171	163	Metrics
soc.culture.polish	6149	289	112	4900	239	545	520	47	684	76	Metrics
soc.culture.china	5925	474	144	4621	351	687	686	81	4770	172	Metrics
soc.culture.indonesia	5400	370	111	4386	272	432	432	52	1849	135	Metrics
soc.culture.argentina	4767	252	103	4213	204	207	205	62	2779	62	Metrics
soc.culture.canada	4641	674	232	3332	572	767	731	97	4516	273	Metrics
soc.culture.venezuela	4554	283	99	3146	220	970	969	81	2204	133	Metrics
soc.culture.turkish	4221	375	117	3492	300	259	256	39	2401	108	Metrics
it.cultura.religioni	4115	288	117	3750	247	138	123	37	2383	35	Metrics
soc.culture.quebec	4035	333	171	3500	298	160	158	38	2740	64	Metrics
soc.culture.europe	3832	557	176	2718	434	716	686	69	3726	252	Metrics
soc.culture.indian	3645	698	152	2216	400	781	779	75	3001	197	Metrics
soc.subculture.bondage-bdsm	3381	229	80	2998	149	117	112	29	145	37	Metrics
soc.culture.palestine	3331	357	134	2552	301	405	400	78	3198	175	Metrics
soc.culture.malaysia	3244	478	135	2118	290	725	724	78	2752	116	Metrics
soc.culture.cuba	3209	224	84	1577	167	1370	1367	82	1772	139	Metrics
soc.culture.vietnamese	3129	254	94	1597	188	1181	1179	93	650	62	Metrics
soc.culture.british	3098	533	178	2190	423	573	572	79	3000	232	Metrics
soc.culture.spain	3010	304	122	2273	257	378	377	55	2951	103	Metrics
soc.culture.iraq	2764	459	154	1880	379	581	579	119	2294	225	Metrics
soc.culture.russian	2619	355	107	2001	287	348	329	72	2295	146	Metrics

Sur les fils de discussion à l'intérieur des groupes : nombre de messages durant la période choisie et nombre total de messages, âge en jours du fil, présentation en arbre du fil qui permet de visualiser qui répond à quoi.

Sur les auteurs ou participants : premier jour de participation, dernier jour, nombre de messages pour le jour et l'avant-dernier jour recherchés, la semaine et l'avant-dernière semaine recherchées, le mois, le trimestre, l'année..., le nombre de groupes auxquels l'auteur a participé ; puis, par groupes, le nombre de jours d'activité, le nombre de messages, le nombre de fils lancés, le nombre de fils auxquels l'auteur a participé

Enfin pour chaque groupe, les 40 participants les plus actifs sont listés avec leurs statistiques accessibles.

Mon travail se concentre à l'inverse sur des listes de discussion : dans quelle mesure puis-je m'appuyer sur des outils, des méthodes, des résultats qui ont été élaborés pour des groupes Usenet ? On l'aura compris, il n'y a pas de réponse tranchée à cette question : il y a indubitablement des similitudes entre les différents supports et la convergence s'est plutôt accrue avec le temps. Cependant, il existe aussi de vraies différences notamment dans la manière dont l'utilisateur interagit avec le dispositif : du point de vue qui est le mien – c'est-à-dire i) qui cherche à comprendre comment à partir d'interactions répétées peut se construire quelque chose qui ressemblerait à une action ou une position collective et ii) qui ne présuppose pas l'existence d'une structure liée au format de communication – la question qui doit alors prévaloir est de savoir quel sens les outils et les résultats produits par ces approches ont par rapport à l'expérience de l'utilisateur. Nous garderons donc en mémoire cette interrogation qui, de fait, est plutôt absente des travaux du domaine.

Des groupes sur quelles thématiques ?

Bien que les groupes électroniques de discussion se soient développés sur les thèmes les plus variés, les travaux qui les concernent ne sont pas également répartis. Les recherches qui ont porté sur les newsgroups Usenet à un niveau relativement macroscopique (analyse structurale d'un ensemble de groupes) distinguent généralement quatre grandes thématiques : les groupes orientés vers la technique dans lesquels ceux qui concernent l'informatique tiennent une place importante ; les groupes formés autour des hobbies ou intérêts les plus divers (du patchwork à l'investissement financier en passant par le jeu ou la généalogie) ; les groupes débattant de sujets « politiques » ; les groupes consacrés aux questions de santé.

Ceci étant, les travaux qui se sont penchés de façon plus spécifique sur certains groupes ne sont pas représentatifs de cette variété : beaucoup d'attention a été portée en général à l'internet médical – une revue entière, le *Journal of Medical Internet Research*, lui est par exemple consacrée – et cet intérêt se manifeste aussi en ce qui concerne les forums, les groupes usenet ou les listes de discussion. La médecine et la santé sont des domaines phares de l'internet tant du point de vue de l'offre que de la demande – c'est l'un des premiers motifs de consultation d'internet. Très tôt, un certain nombre d'acteurs se sont inquiétés ou réjouis de ce qu'ils pouvaient percevoir comme une remise en cause des relations traditionnelles entre médecins et patients, les premiers n'ayant plus le monopole de l'accès aux connaissances.⁵ La surreprésentation de ces travaux dans la revue que je propose n'est donc qu'en partie liée à mon objet d'étude.

⁵ Sur ces questions, voir Akrich & Méadel (2004).

Autre domaine assez présent dans la littérature, celui de l'éducation⁶ : le développement des TIC a donné lieu à un certain nombre d'expériences pédagogiques qui se sont appuyées sur des forums ou des listes de discussions. En contraste, ces travaux seront assez peu présents dans notre analyse, du fait du cadrage très spécifique qui est le leur.

Le domaine de la technique et tout particulièrement celui de l'informatique et des communautés Open Source est un autre espace qui a suscité un nombre significatif de travaux, avec des questions qui touchent notamment à la création et à la survie de formes d'organisations tout à fait originales.

Enfin, les groupes centrés sur des hobbies ou intérêts partagés et ceux qui débattent de questions politiques sont assez peu étudiés. La question générale du rôle d'internet sur les processus démocratiques n'a certes pas été oubliée⁷, mais lorsqu'il s'agit d'entrer dans une discussion plus précise des dispositifs, l'attention se porte plutôt du côté institutionnel, de celui des organisateurs de consultations et de débats publics, voire des techniciens chargés de concevoir les interfaces appropriées ou encore du côté des utilisateurs au sens quasi-technique du terme, puisqu'il s'agit de savoir comment ils perçoivent l'espace de parole qui leur est proposé⁸. Les forums et listes politiques « spontanés » semblent n'avoir pas été beaucoup analysés, alors que, plus récemment, le phénomène des blogs politiques suscite une attention soutenue.

Du texte au réseau

Nous avons souligné plus haut que le message avait deux composantes : son contenu, et son en-tête qui permet de spécifier son origine et sa place dans la dynamique des échanges. J'ai choisi d'ordonner cette présentation en partant des analyses de contenu – les plus proches des méthodes utilisées de manière très générale en sociologie – pour aboutir aux analyses de réseau.

Avant d'entrer dans le vif du sujet, signalons tout de même qu'à côté de travaux qui prennent appui sur le matériau même constitué par les groupes électroniques de discussion, certains auteurs ont recouru à d'autres approches qui court-circuitent ce matériau : les enquêtes par questionnaire et les entretiens.

Assez rarement, il s'agit de recueillir des éléments sur les pratiques : Butler, B. *et al.* (2007) ont réalisé un questionnaire dont la première partie était consacrée au temps passé par les participants aux activités liées aux groupes de discussion et au repérage

⁶ On peut voir notamment Caviale & Bruillard (2009) ; Edwards & Huff (2003) ; Goos & Bennison (2008) ; Groth (2005-2006) ; Hara *et al.* (2000) ; Jeong (2003) ; Liu & Tsai (2008) ; Marra *et al.* (2004) ; Zrebiec (2005)

⁷ Flichy (2008, 2010)

⁸ Benvegno & Brugidou (2008) ; Flichy (2008, 2010) ; Jensen *et al.* (2007) ; Kavanaugh *et al.* (2005) ; Monnoyer-Smith (2006) ; Stromer-Galley (2007) ; Wright & Street (2007)

d'activités qui auraient comme objectif de contribuer à la bonne marche du groupe de discussion (envoi de messages d'encouragement, de félicitations, de rappel à l'ordre etc.). La plupart du temps, les questionnaires visent à interroger les motivations de la participation (Ridings & Gefen 2004), la manière dont les participants perçoivent l'effet des groupes de discussion notamment par rapport à d'autres media (Turner, J. W. *et al.* 2001, Matzat 2004), la forme des relations nouées grâce à ce media (McKenna *et al.* 2002). Les entretiens reprennent de manière plus approfondie certaines de ces thématiques (Broom 2005; Ziebland *et al.* 2004) ou vont chercher à élucider les raisons de la non-participation (Sandaunet 2008).

En règle générale, ces travaux ne sont pas centrés sur le groupe lui-même, son fonctionnement, ce qu'il produit : la dimension collective de l'échange est absente alors qu'est privilégiée une analyse qui constitue une dichotomie entre le groupe et l'individu, le premier étant considéré comme une ressource potentielle pour le second ; il s'agit surtout de s'interroger sur la « valeur » de la ressource et les « coûts » d'accès à cette ressource. Ils sont donc assez étrangers à la question qui nous occupe ici, puisque, d'une certaine manière, ils en considèrent la réponse comme acquise.

Comprendre la nature des échanges : les analyses de contenu

Un grand nombre de travaux s'appuient sur ou prennent pour objet l'analyse de discussions dans des forums, des newsgroups ou des listes de discussion. Ils se différencient cependant les uns des autres à la fois par la problématique qui les sous-tend et les méthodes qu'ils adoptent.

Du point de vue des méthodes, on peut distinguer en effet des analyses de contenu « classiques » qui supposent la lecture intégrale du matériau, sa restitution sous une forme synthétique illustrée par des exemples ; des analyses qui reposent sur un codage à partir d'une grille établie a priori et qui sont susceptibles de déboucher à la fois sur des statistiques et une typologie ; des analyses qui reposent sur l'élaboration de catégories intermédiaires issues de la confrontation au matériau. Dans un certain nombre de cas, ce travail peut suivre une période d'acculturation assez longue, voire même une forme d'ethnographie virtuelle (Hine 2000), dans le sens où l'observateur interagit avec les observés, soit qu'il participe à leurs activités, soit qu'il se soit déclaré comme ethnographe et ait eu des échanges sur ce thème, voire même en ayant emprunté les deux postures. Dans ces différentes configurations, la « réduction » du matériau prend des formes spécifiques.

Les formats classiques d'analyse

J'appelle analyses de contenu « classiques » des analyses qui, sur le fond, s'interrogent assez peu sur leurs propres fondements : le chercheur a à sa disposition un matériau – qu'il s'agisse d'entretiens, de documents d'archives ou de discussions sur un groupe électronique – et il va essayer de construire une interprétation à partir de ce support. Qu'il ait fait a priori des hypothèses fortes quant aux formes d'interprétation possibles

ou qu'il laisse davantage ces hypothèses se construire dans la confrontation avec le matériau, à la limite peu nous importe ici. En revanche, on constate que ces types d'analyse sont en fait utilisées dans deux configurations assez différentes : il peut s'agir de creuser une problématique qui n'est pas spécifique des groupes électroniques de discussion mais pour laquelle ils offrent soit un accès privilégié à du matériau, soit une mise à l'épreuve intéressante ; il peut s'agir aussi de rendre compte de ce qui se passe dans un groupe particulier, d'en faire une forme d'ethnographie.

Quelques exemples d'études dans lesquelles l'analyse de contenu vise à éclaircir une question exogène

Wikgren (2001) se demande quel rôle les groupes de discussion jouent dans la transformation d'un patient « passif » en un patient « expert » actif, en charge de lui-même et de sa maladie : pour cela, il collecte 14 fils de discussion dans différents newsgroups, repère les citations et liens à différentes sources de savoir, et s'intéresse au contexte de citation.

Kitchin (2002) cherche à apprécier jusqu'à quel point les membres de groupes d'Alcooliques Anonymes adhèrent à la « doctrine » de l'association ; d'un groupe Usenet, il extrait 34 fils de discussion pertinents sous cet angle, et se livre à une recension et une analyse des décalages entre les positions qu'ils affichent et la doctrine.

Mahoney (1998) cherche à appréhender l'expérience des proches en charge d'une personne atteinte de la maladie d'Alzheimer et analyse environ 400 messages portant sur ce thème.

Akrich & Méadel (2002) s'intéressent à la place occupée par le médicament dans la relation du patient à sa pathologie, à son corps et à son médecin, en prenant un certain nombre d'exemples dans trois listes de discussion sur des pathologies variées. Il s'agit de montrer d'une part, que le médicament occupe une place assez différente dans chacun des trois cas, et d'autre part, que les listes sont des espaces dans lesquels s'élaborent collectivement une forme d'expertise partagée sur le médicament.

Dans ces différents cas, le groupe électronique de discussion est considéré comme un support sur lequel s'expriment, et, le cas échéant, se confrontent et se transforment un certain nombre d'expériences, d'opinions et de connaissances. Si ce ne sont pas la liste ou le groupe en tant que tels qui intéressent le chercheur, cela ne signifie pas pour autant qu'il néglige l'importance du média et considère le groupe comme une simple fenêtre sur le monde : la question de ce que l'échange produit spécifiquement sur les objets d'étude est régulièrement soulevée. La réduction du matériau est obtenue par une sélection de fils de discussion pertinents sur la base notamment du vocabulaire utilisé.

Des analyses de groupes en tant que tels

Radin (2006) s'attache à décrire un forum autour du cancer du sein qui, à l'époque où elle mène son analyse, totalise un demi-million de messages après 10 ans d'activité : pour obtenir un corpus traitable, elle choisit de se concentrer sur un mois d'activité.

Baym (2000a) adopte une autre stratégie en isolant une discussion sur un thème particulier, discussion qu'elle juge représentative des aspects importants dans le fonctionnement du groupe mais qui ne compte que 524 messages, à comparer aux plus de 33 000 collectés.

L'une comme l'autre mobilisent en fait une palette de méthodes même si la présentation de matériaux extraits de leur sélection tient une part notable dans leurs descriptions : entretiens, enquêtes en ligne, petite analyse statistique viennent en appui à la démonstration. N. Baym multiplie les approches puisqu'elle s'intéresse selon les chapitres de son livre à l'activité même des participants, notamment à leur travail interprétatif ; aux relations, éventuellement critiques, que ces participants entretiennent avec l'objet de leurs discussions, le soap opera ; à la gestion des relations interpersonnelles entre les participants ; au style de communication développé par les participants et notamment à la place de l'humour etc. C'est grâce à ce kaléidoscope de méthodes et de questions qu'elle parvient à produire une description assez convaincante et à faire toucher du doigt ce que cela signifie que de participer à ce groupe.⁹

A la différence du cas précédent, les chercheurs cherchent à prendre en compte le groupe en tant que tel, mais de façon à première vue paradoxale, ils concentrent le travail d'analyse de contenu sur une portion très faible du matériau pertinent par rapport à leur objectif. Il s'agit plutôt ici d'un travail d'exemplification : la familiarité acquise par la fréquentation du groupe, le recours à d'autres médiations permettent au chercheur de déterminer les grandes lignes selon lesquelles sa démonstration va s'organiser et que les extraits choisis vont permettre d'appuyer.

Les codages a priori

Une analyse en profondeur des échanges suppose, comme nous venons de le voir, de construire un échantillon très restreint de messages : d'autres auteurs arbitrent différemment entre « largeur » de l'échantillon et « profondeur » de l'investigation, et choisissent de recourir à une forme de synthèse qui relève davantage d'un processus de sommation que d'une exemplification. L'utilisation de codages – selon une grille préétablie, souvent empruntée à des travaux antérieurs – est l'une des techniques privilégiées de ce point de vue.

Généralement, il s'agit de décrire la nature des activités dans lesquelles les participants sont engagés : Loader *et al.* (2002) et Coursaris & Liu (2009) se sont attachés à l'identification et à la quantification des différentes formes de soutien (accompagnement, estime, informationnel etc.) apportés dans des groupes consacrés au diabète pour les premiers et au SIDA pour les seconds. Burnett, Gary & Beurle (2004) applique à deux groupes dans le domaine de la santé une catégorisation décontextualisée par rapport à l'objet des groupes qui distingue les activités hostiles, les comportements interactifs de collaboration non informationnels (neutre, humour, émotionnel) / informationnels. Maloney-Krichmar & Preece (2005) ont codé les 492 messages d'une semaine de

⁹ Ce qui n'a qu'assez peu à voir avec le modèle d'émergence des communautés en ligne qu'elle développe ailleurs Baym (1998).

discussion sur un forum consacré aux personnes qui ont des problèmes de genou à l'aide d'une grille tirée d'un travail de 1947 sur les différents rôles joués par les membres d'un groupe (par exemple fournisseur d'information, fournisseur d'opinion, commentateur du groupe, etc.) ; compte tenu du fait que, comme les auteurs le reconnaissent, les mêmes personnes prennent différentes postures, parfois à l'intérieur du même message, on peut considérer qu'en fait il s'agit encore d'une description d'activités ou de comportements. Pour Welser, H. T. *et al.* (2009), il s'agit d'évaluer l'apport des forums type Questions/Réponses, c'est-à-dire sur lequel des personnes supposées rencontrer une difficulté technique viennent soumettre leur problème auquel d'autres personnes supposées expertes apportent des réponses : leur codage vise à faire la distinction entre des réponses appuyées sur des éléments tangibles, des opinions, voire des réponses sans rapport avec la question.

À côté de ces travaux qui visent plutôt le contenu des messages, d'autres approches, principalement linguistiques, s'attachent plutôt à la forme : nous avons mentionné le travail de Mondada (1999) sur l'utilisation de moyens techniques propres au media pour créer ou rendre visibles des formes d'interactivité dans le message lui-même ; on peut voir aussi Waseleski (2006) qui s'intéresse à l'utilisation des points d'exclamation : cet auteur procède à un codage de messages – issus de deux listes de discussion pour les professionnels des bibliothèques et de la documentation – basé sur une description d'actions génériques (action ou appel à l'action, excuse, challenge, coalition, faits, opinion, préférences personnelles, sarcasmes, remerciements) afin de contextualiser l'usage des points d'exclamations et d'en spécifier la signification ; il s'interroge sur la dimension genrée de cet usage, mais peu sur la spécificité éventuelle des usages dans les listes qu'il étudie. En fait, ce « terrain » linguistique nouveau lui permet de relativiser l'interprétation habituelle du point d'exclamation comme marqueur d'excitabilité et de mettre en valeur un usage qui indiquerait plutôt des intentions amicales.

Plusieurs remarques peuvent être faites :

- le choix d'un certain type de codage renvoie à « l'espace académique de référence », et à la place qu'y occupe l'analyse des discussions électroniques : s'agit-il en particulier de décrire et de caractériser différents groupes électroniques ou de dresser des comparaisons entre des groupes électroniques et des organisations ou des pratiques d'une autre nature ?
- l'analyse est centrée sur les activités sociales ou langagières et ne s'intéresse pas au groupe en tant que tel ni à ce en quoi les « fonctions » assurées reposent ou pas sur l'existence de processus collectifs¹⁰.
- le message est ici considéré uniquement sous l'angle de l'action, de ce qu'il produit sur ceux qui le reçoivent, que ce soit en raison de son contenu ou de sa forme ; on se

¹⁰ Par exemple, dans l'étude de Coursaris & Liu (2009), les 5000 messages codés ont été sélectionnés aléatoirement dans 26 newsgroups

trouve donc assez éloigné des analyses de contenu « classiques » décrites plus haut et dans lesquelles le contenu même du message était central.

La production de catégories ad hoc

Face à l'alternative que dessinent les options précédentes – analyse en profondeur d'un petit corpus, codage superficiel d'un gros corpus – certains auteurs ont cherché une troisième voie. Celle-ci passe souvent par l'utilisation d'une « aide » à l'analyse sous la forme d'un logiciel de traitement.

Nudist et Nvivo sont les plus utilisés, tout particulièrement de la part de chercheurs qui se réclament de la « grounded theory » : dans cette perspective, la catégorisation doit émerger d'un processus complexe et itératif de confrontation avec les données et non être imposée de l'extérieur par le chercheur. Ces logiciels de traitement des données qualitatives permettent aux chercheurs de constituer et de valider les catégories ainsi que de naviguer sans se noyer dans un large corpus.

Alceste repose un principe différent et sépare plus nettement les phases d'analyse du texte de celles de l'interprétation : le logiciel calcule les cooccurrences des termes dans des séquences de texte. Sur cette base, il construit des ensembles de vocabulaire « cohérents » – au sens où les termes qui composent un ensemble ont une forte probabilité de se retrouver associés dans le texte – et classe les différents énoncés dans ces ensembles. L'approche s'apparente à une analyse de réseau telle proposée par exemple par Leximappe ; cependant le logiciel va plus loin et découpe en quelque sorte des « clusters » à l'intérieur du réseau et permet de classer les différents messages dans ces clusters. En principe donc, le logiciel permet de dégager sur des masses de données considérables les grands thèmes autour desquels s'organise le corpus.

Les travaux qui s'appuient sur ce type d'outils ou tout du moins qui s'attachent à la construction de catégories à partir d'une confrontation avec le corpus débouchent sur deux grands types d'analyses : un premier ensemble de travaux visent à produire une catégorisation des « activités » du groupe qui se prête à une comparaison entre groupes ou entre groupes électroniques de communication et d'autres types d'organisation ; le second ensemble est davantage orienté vers une analyse de ce qui est spécifiquement produit par l'activité d'un groupe particulier.

Décrire les activités des groupes

Wiseman & Grunwald (2008), qui s'intéressent à un forum de discussion sur la boulimie (bulimia nervosa), ont effectué un travail de codage à partir des intitulés des 2000 fils de discussion (14000 messages) sur une période de 16 mois (durée totale du forum : 7 ans). Le codage a été réalisé en deux temps : d'abord, les fils de discussion ont été classés par un certain nombre de caractéristiques structurelles et les catégories de codage ont été établies par échantillonnage et comparaison d'un certain nombre de fils ; puis l'ensemble des fils a été codé, ce qui a permis de distinguer trois types principaux de fils : des fils orientés vers des problèmes, des fils orientés vers la communication (échanges qui ne sont

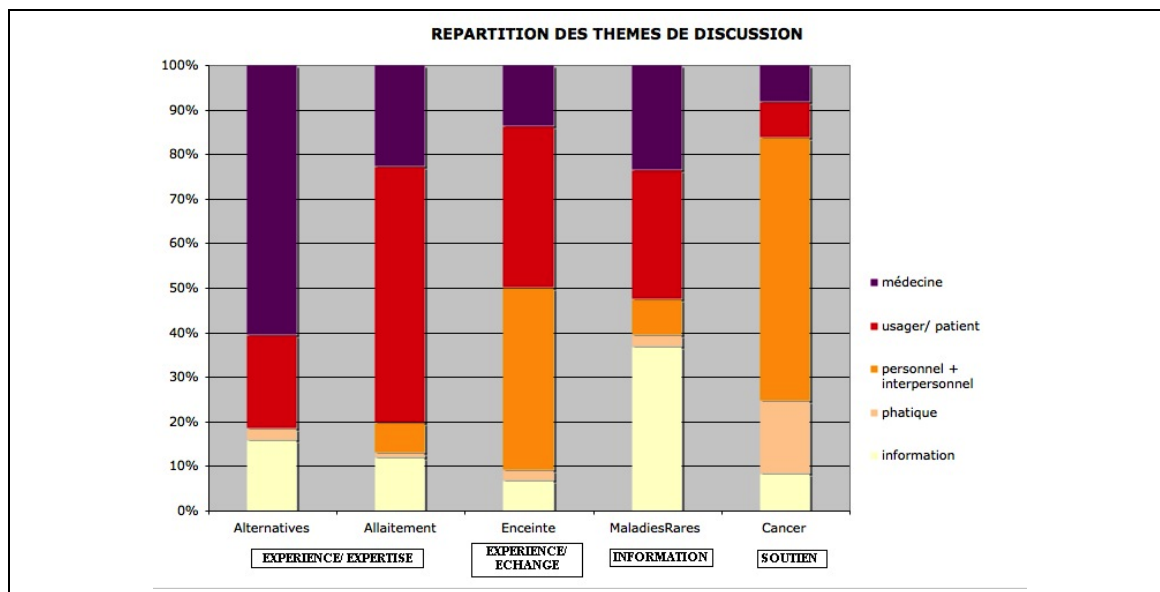
pas directement connectés à la bulimia nervosa), et des fils dits de métacommunication (sur le fonctionnement du forum).¹¹

Au démarrage de notre travail sur les listes, mené en collaboration avec Cécile Méadel (Akrich & Méadel 2001, repris dans Akrich & Méadel 2009), nous avons utilisé une approche un peu similaire : notre objectif était de comparer un certain nombre de listes de discussion dans le domaine de la santé ; après un processus d'acculturation à ces listes basé sur la lecture d'un nombre important de messages, nous avons défini six catégories descriptives¹² (phatique, personnel et interpersonnel, collectif, médecine, patient/usager, information) qui permettent de classer les fils (20000 messages environ) et de caractériser les listes : contrairement à la vision un peu monolithique que donne la littérature sur les groupes concernés par les problèmes de santé, cette catégorisation permettait de contraster fortement des listes comme on peut le voir sur le graphique suivant.¹³

¹¹ Au delà d'une description des activités de la liste, l'objectif de cette analyse était de croiser cette catégorisation avec une description « structurale » des fils basés sur d'autres paramètres, des paramètres temporels (mois, jour et heure de la création du fil) et des paramètres « utilisateurs » (nombre de participants au fil, nombre de consultations du fil par des internautes), sans que des hypothèses explicites sur l'intérêt de ce croisement soient proposées, croisement qui d'ailleurs ne donne pas de résultats remarquables.

¹² Le thème « phatique » renvoie aux manifestations de la civilité ordinaire et se manifeste par des mots comme « bonjour », « vœux », « coucou », « bienvenue ») ; le thème « personnel et interpersonnel » concerne soit les évocations de la vie privée, soit les relations personnelles qui s'établissent entre certains membres de la liste) ; le thème « collectif » contient ce qui a trait à la vie de la liste (abonnement, sondage, co-listiers...)) ; le thème « médecine » contient tous les mots appartenant clairement au registre médical : « chimio », « prozac », « lombalgie », « scanner », etc.) ; le thème « patient/ usager » englobe les différents aspects de la question de santé abordée vus du point de vue de l'utilisateur : « sommeil », « douleur », « droit », « travail », « congé » etc.) ; le thème « information » fait référence aux sources d'information (presse, télévision, web, journaux scientifiques etc.) ou à des événements externes à la liste (conférences...). Comme le thème « collectif » n'était pas discriminant dans l'échantillon de listes présentées sur le graphique, il n'a pas été pris en considération.

¹³ Les noms donnés aux listes sur le graphique correspondent à la question de santé traitée par le groupe : le groupe désigné par le nom « Alternatives » s'intéresse aux médecines dites « douces ».



Appréhender la spécificité d'un groupe

Pour Barker (2008), la question est de savoir si les groupes de patients sur Internet contribuent aux processus de médicalisation. Le codage réalisé sur un corpus de 1800 messages (un an de participation sur un forum consacré à la fibromyalgie, pathologie controversée) lui permet de mettre en évidence quatre grands thèmes en rapport avec sa problématique sur lesquels se construit par l'échange et s'exprime dans l'échange une forme de constat partagé : i) la réification de la maladie, c'est-à-dire la certitude des membres que leurs symptômes renvoient à une même maladie, qui est à la fois le présupposé sur lequel le forum a été établi et qui est en même « enactée » par les échanges eux-mêmes ; ii) un même scepticisme vis à vis des médecins qui ne reconnaissent pas la réalité du trouble, sur fond d'une dépendance à la médecine ; iii) une expertise incorporée des malades qui cherche à s'aligner sur l'expertise médicale embrassant son langage rationnel et son dualisme corps/ esprit ; iv) un « empowerment » des membres par l'apprentissage collectif qui ne trouve pas à s'exercer dans le monde extérieur.

Dans un domaine différent, celui de l'autisme, Huws *et al.* (2001) produisent une analyse un peu comparable qui leur permet de mettre en évidence le travail de construction, mené conjointement par les parents présents sur la liste, de ce qu'est l'autisme. ¹⁴

Conclusion

Comme nous venons de le voir, les méthodes utilisées par les chercheurs qui s'intéressent d'une façon ou d'une autre aux contenus échangés dans les groupes

¹⁴ dans un registre comparable, Harry et al. (2008) se sont attachés à décrire, en s'appuyant sur le logiciel Alceste, les thèmes autour desquels échangent des diabétiques sur un forum ; Benvegnu & Brugidou (2008) ont quant à eux analysé les contributions du public sur un forum destiné à animer un débat public autour de la maîtrise de l'énergie.

électroniques de discussion gèrent de manière différenciée la question de la réduction du matériau. Certaines privilégient la « profondeur » et en conséquence sont amenés à effectuer des sélections drastiques qui fragilisent du coup la démonstration : il n’y a pas de commune mesure, ni même d’articulation possible entre ce qui est mobilisé à l’appui de l’argument et ce dont ces extraits sont censés rendre compte. D’autres, à l’inverse, privilégient la « couverture », mais du coup restent à la surface des échanges. Ces approches pourraient être complémentaires et permettre des approches graduées mais sont de fait peu utilisées de concert. Seules des méthodes « équipées » – encore en partie à définir – pourraient permettre de résoudre différemment l’équation posée par l’abondance du matériau.

Au delà, on observe qu’au travers de ce type d’investigations, il s’agit de répondre à deux types principaux d’objectifs :

- décrire dans un langage générique les « activités » du groupe, de manière notamment à établir une comparaison avec d’autres espaces ;
- décrire les contenus qui circulent dans le groupe, ces contenus étant selon les cas pensés indépendamment du média – qui apparaît alors comme un simple intermédiaire – ou au contraire intimement liés au média, qui est ce par quoi du sens est produit collectivement.

Dans tous les cas, ces types d’analyses partagent deux présupposés ou points aveugles :

- les participants à ces espaces de discussion sont presque toujours indifférenciés. Ceci se traduit de plusieurs manières. Le fait que la distribution des participations soit extrêmement inégale dans tous ces groupes n’est pas pris en compte : les locuteurs sont finalement considérés comme des représentants interchangeables d’une même classe de personnes définies par l’intérêt ou la situation qui les a poussés à se joindre au groupe. Du coup, la dynamique propre de la discussion n’est pas non plus réellement présente ; les confrontations ou désaccords éventuels sont peu évoqués, l’espace de la liste ou du groupe est un espace assez homogène de ce point de vue.
- la liste ou le groupe sont des espaces sans Histoire : ils sont considérés comme des systèmes stables, ce qui dispense par exemple de justifier les procédures d’échantillonnage basées sur une variable temporelle.

Comprendre la dynamique des échanges : modélisation et expérimentation

À la différence de ceux dont nous venons d’envisager les travaux, plusieurs chercheurs se sont posés la question des conditions qui permettraient d’expliquer la dynamique des groupes. Ces approches se situent à deux niveaux différents :

- un premier ensemble de recherches (Butler, B. S. 2001; Jones *et al.* 2004; Whittaker *et al.* 1998) se placent à l’échelle du groupe tout entier : ce sont les caractéristiques de ce groupe qui détermineraient sa viabilité.

- d'autres travaux (Arguello *et al.* 2006; Burke, M. *et al.* 2007; Burke, Moira & Kraft (2008; Burke, Moira *et al.* 2010; Joyce & Kraft 2006) considèrent que les caractéristiques intrinsèques de chaque message expliquent en partie le fait qu'il reçoive ou pas une réponse, et partant de là influe sur la dynamique des échanges.

Les premiers ont essentiellement travaillé sur la formulation d'hypothèses explicatives et leur traduction en critères modélisables ; ils ont ensuite testé leurs hypothèses sur de très larges corpus.

Les seconds ont mené une démarche similaire, mais se situant au niveau du message qu'il s'agit alors de caractériser ; ils se sont par ailleurs engagés de manière complémentaire dans une approche expérimentale, consistant à manipuler des messages afin de produire des variations contrôlées qui permettent de tester spécifiquement l'influence de telle ou telle variable.

Les modèles de fonctionnement des groupes

Trois modèles du fonctionnement des groupes ont été proposés, celui du « common ground » (Whittaker *et al.* 1998), celui de la théorie des ressources (Butler, B. S. 2001), et celui de surcharge informationnelle (Jones *et al.* 2004).

Dans le modèle du « common ground » développé dans le domaine de la communication en face-à-face, l'idée est que pour que les participants à une conversation puissent se comprendre les uns les autres, il faut qu'ils partagent une certaine connaissance les uns des autres, des règles de la discussion et des sujets sur lesquels ils vont discuter.

Whittaker *et al.* (1998) se posent la question de la manière dont on pourrait traduire cette exigence dans le domaine des newsgroups Usenet, alors que les participants sont infiniment plus nombreux que dans une conversation usuelle et qu'ils sont par ailleurs changeants : ils vont proposer un certain nombre de critères de substitution comme le fait que le groupe soit modéré, qu'il ait produit des FAQs, qu'il ne soit pas trop grand, que les messages ne soient pas trop courts et qu'il n'y en ait pas trop provenant d'autres groupes (cross-posting) etc. ; par ailleurs, la fluidité de la conversation est évaluée par le nombre de messages par fil de discussion. Ayant opérationnalisé les hypothèses, ils procèdent par test sur un corpus de 2,15 millions de messages échangés dans 500 groupes pendant 6 mois. Le test infirme en partie les hypothèses : les messages courts et le cross posting semblent augmenter l'interactivité plutôt que la diminuer. Les auteurs sont assez lucides sur les limites du type d'analyse qu'ils ont développé :

« The current analysis is silent about the effects of conversational content: which specific topics or conversational styles encourage large responses and which fail to elicit a response? What factors encourage or discourage flaming? What are people's reactions to flagrant examples of cross-posting? (...) Why do people contribute to certain discussions but not to others? How long do people lurk before they first post? And how is dominance viewed? Why do certain people post multiple messages and how are they perceived by others for doing so? » p. 263

Pour Butler (2001), adepte de la théorie des ressources, toute structure sociale ne peut perdurer que si elle est capable de convertir les ressources (temps, connaissance, énergie, argent, ressources matérielles) mises à disposition par les membres en bénéfiques (information, soutien) pour ces mêmes membres. De ce point de vue, la taille de la structure (qui déterminerait les ressources disponibles), l'activité de communication proprement dite (qui serait un indicateur de l'activité de « transformation » des ressources en bénéfiques) sont des éléments cruciaux qui expliqueraient la durabilité de l'organisation.

Les résultats de son étude qui s'est effectuée sur 206 newsgroups sur 4 mois sont loin de valider complètement cette hypothèse puisque par exemple les groupes les plus grands sont à la fois ceux qui gagnent le plus de participants dans une période de temps mais aussi ceux qui perdent le plus de participants. Dans cette analyse, se trouve de fait construite une dualité entre une structure sociale et des individus : Butler raisonne dans une perspective qui est celle de la comparaison entre les performances de différents types d'organisations, parmi lesquelles les communautés en ligne. Or une des caractéristiques de ces groupes est précisément qu'il est impossible de maintenir cette distinction : il n'y a pas d'autre organisation que celle que crée et recrée en permanence l'échange de messages entre un certain nombre de participants.

La troisième approche (Jones *et al.* 2004) part presque du point où s'arrête Butler, à savoir des effets négatifs que peut engendrer une taille trop importante des groupes : les individus ont une certaine capacité d'absorption et de traitement de l'information, qui lorsqu'elle est atteinte génère des comportements nouveaux, allant du raccourcissement des messages jusqu'à l'évitement ou au retrait, mettant éventuellement en péril le fonctionnement du groupe.

La difficulté est ensuite d'opérationnaliser la notion de charge cognitive : ils construisent un modèle assez compliqué qui prend en compte la longueur totale des messages, la longueur spécifique des messages (en dehors de la citation d'autres messages et des fichiers attachés), le nombre de messages. De même, la question de la taille du groupe n'est pas simple à traiter : doit-on considérer que la taille est en rapport avec le nombre de messages, le nombre de participants, le nombre de participants ayant répondu à d'autres etc. ? En tout état de cause, ils construisent trois hypothèses : lorsque le groupe fonctionne près du niveau maximal – compte tenu des capacités individuelles – (i) si le nombre de messages augmente, leur complexité doit décroître ; (ii) les messages simples auront plus de chance de recevoir une réponse que les messages compliqués ; (iii) des participants auront tendance à arrêter ou à réduire leur participation. Les hypothèses sont ensuite testées sur 600 groupes Usenet représentant 2,6 millions de messages sur 6 mois. Leur analyse repose sur un certain nombre de développements méthodologiques explicités dans une annexe assez longue. On peut percevoir la complexité de la moindre opération. Par exemple, l'identification des messages qui sont des réponses à d'autres messages nécessite des développements importants car aucun critère simple ne suffit : ils sont conduits à prendre en compte (i) le fait que le sujet contienne une chaîne de caractères du type RE : avec ses variantes ; (ii) le fait que le message contienne des parties indentées avec des « > » et/ou des expressions du type « Machin a écrit : », (iii) le fait que l'en-tête du

message contienne des indications dans des champs de type Reference : ou In Reply to : qui tracent en principe les relations entre messages. En revanche, la mesure de la stabilité du groupe pose problème : elle est construite en prenant en compte la part des participants présents un mois donné et qui était présents le mois précédent. Or, les auteurs n'ont pas intégré le fait que la durée choisie pour mesurer la stabilité du groupe a un effet important sur le résultat, en raison de ce que la participation est très inégale entre les locuteurs, et que ces inégalités sont elles-mêmes inégales selon les groupes (plus ou moins grande concentration des messages sur un nombre limité de locuteurs). Le choix d'une durée courte pour le calcul de la stabilité sélectionne de fait les très gros locuteurs, groupe dont la taille varie selon les listes.¹⁵

Ces trois modèles ont pour point commun de faire l'hypothèse d'un effet systémique dans le fonctionnement des groupes de discussion qui tient à des caractéristiques génériques des individus qui participent : ils ont besoin de se connaître pour pouvoir communiquer ; leur comportement est dicté par une perspective utilitariste – leur participation à un groupe est conditionnée par le bénéfice qu'ils en retirent mise en regard de l'investissement qu'ils consentent ; ils disposent de capacités cognitives limitées. Ces différents modèles se différencient nettement quant aux mécanismes sous-jacents qui sont à l'origine de l'effet systémique : ils font d'ailleurs des prédictions différentes quant au rôle de la taille du groupe sur sa stabilité. Ou construisent différemment les chaînes de causalité : par exemple, pour Butler, les messages courts augmentent l'interactivité, alors que pour Jones, les messages courts sont une réponse à l'augmentation de l'interactivité qui sature les capacités cognitives des individus. Ce qui attire l'attention sur la question du rapport entre les analyses statistiques et les interprétations de type causal qui en sont tirées.

Ces modèles partagent une même façon de considérer les participants comme une masse indifférenciée ; ici, de surcroît, les contenus échangés ne sont pas pris en compte à l'exception de quelques traits formels. Si la dimension historique se trouve réintroduite, elle semble se réduire à un enchaînement quasi-mécanique de causes qui sont à la fois le produit et le déterminant des comportements individuels. Par rapport à

¹⁵ On se trouve là devant un problème récurrent : les groupes de discussion sont des systèmes complexes dans lequel un certain nombre de variables sont liées) ; l'interprétation est délicate car en définitive, ce que l'on prend pour l'expression d'une caractéristique du groupe peut en fait être le résultat du jeu d'une autre variable. Ici par exemple, on croit mesurer la stabilité alors qu'en fait l'indicateur construit est en partie une mesure du degré d'inégalité dans la répartition des participations. Dorat *et al.* (2007) ont proposé de ce point de vue une analyse intéressante : ils ont voulu montrer que les propriétés observables dans la structure des échanges d'un groupe de discussion peuvent être modélisées à l'aide d'un nombre limité d'hypothèses. Pour cela, ils sont partis d'une groupe de discussion sur le logiciel libre dont ils ont extrait quelques caractéristiques (le nombre de messages, la distributions des liens (degree distribution), la distribution des structures de fils, la distribution des auteurs etc.) ; ils partent de certaines de ces caractéristiques et construisent un réseau aléatoire respectant la caractéristique en question et montrent que ce réseau présente des similitudes très importantes avec le réseau réel.

la question qui nous occupe – rendre compte de l’activité d’un groupe de discussion et comprendre ce en quoi il « fait » groupe – la plus value apportée par ce type d’analyse est limitée. Elle porte principalement sur les aspects de méthode : les différents travaux ont nourri une discussion sur la manière dont on pouvait traduire des éléments de description de l’activité (comment caractériser l’interactivité, la stabilité ?) ou des messages (qu’est-ce qu’une « réponse ») sous la forme de critères opérationnalisables, ce qui ouvre des perspectives pour un traitement en masse des données. Sur le plan de la compréhension de la dynamique des échanges en revanche, à supposer que l’on puisse mettre en évidence au niveau statistique des régularités, cela n’aide guère à répondre aux questions que soulèvent, dans la pratique, la constatation de « destins » de groupes clairement différenciés.

Les analyses centrées sur les messages

Burke, M. *et al.* (2007) partent en fait d’une perspective très proche de celle de Butler qu’ils citent d’ailleurs à l’appui de leur argumentation :

« In order to survive, online communities, like face-to-face groups, must meet the needs of individual members and the group as a whole, and conversation is the mechanism through which this occurs. It is through reading archived conversations that newcomers determine whether the potential benefit of membership is worth the cost of participation, and groups need a steady supply of newcomers to replenish membership levels over time. Current members solicit information and steer the group toward topics they care about by starting conversations. People start conversations hoping to derive benefit from the group; depending on the response they get members and prospective members will either continue to participate or they will leave. » p. 21

De leur point de vue, donc, les réactions du groupe aux « initiatives » des membres nouveaux ou anciens revêtent un caractère stratégique¹⁶. Ce qui les conduit à s’interroger sur les caractéristiques des messages qui pourraient conditionner le fait de recevoir une réponse.

Ils développent dans une série d’articles un modèle complexe qui intègre des données sur le contexte de réception (nombre de messages postés le jour de l’envoi du message) et sur le message lui-même (s’il s’agit d’un témoignage, s’il comporte des questions, proximité thématique du message avec les sujets débattus dans la liste, complexité linguistique, choix de mots). Comme il ne s’agit évidemment pas de lire l’ensemble des messages, chacun de ses paramètres doit être traduit sous la forme de critères dont le traitement peut être automatisé : par exemple, les premiers messages des nouveaux arrivants sont classés dans la catégorie « témoignage » s’ils contiennent beaucoup de pronoms à la

¹⁶ Dans un premier article (Joyce & Kraut 2006), ils montrent que parmi ceux des 2777 nouveaux arrivants dans les six groupes étudiés qui ont reçu une réponse à leur premier message, 56% posteront à nouveau contre 44% pour ceux qui n’ont pas reçu de réponse ; cependant dans un article plus récent, Burke, Moira *et al.* (2010), s’appuyant sur un corpus plus important (12000 nouveaux arrivants), les pourcentages sont très différents : 45% / 16,7%. Ce qui ne manque pas de soulever quelques interrogations.

première personne, si l'âge est mentionné, si la personne fait référence au fait d'avoir suivi pendant un certain temps les échanges sans intervenir (*lurking*), etc. Les questions sont repérées par les points d'interrogation, mais aussi par certaines expressions du type « je me demande », « je cherche ». La proximité thématique est calculée en agrégeant les fréquences dans le groupe entier des mots utilisés dans le message (calcul pondéré par le nombre de mots). La complexité linguistique est mesurée par la longueur du message, la longueur moyenne des phrases, et le pourcentage de mots « longs ». En revanche, la caractérisation de ce qu'est une « réponse » est extrêmement fruste : bien que travaillant à partir du même type de données que Jones *et al.* (2004), à savoir les groupes Usenet, ils n'ont apparemment pas pris en compte les développements méthodologiques de ce dernier et en sont restés à la prise en compte des données d'en-tête du message, malgré le fait qu'elles ne repèrent qu'une partie des liens entre messages.

L'analyse menée dans Arguello *et al.* (2006) leur permet de dégager un certain nombre de résultats : les messages qui comportent des témoignages et des questions ont plus de chance de recevoir une réponse, de même que ceux qui expriment des émotions, des processus mentaux ; les longues phrases découragent mais pas les longs messages ; contrairement à ce que pourraient laisser supposer les analyses de Jones (2004), le volume des échanges au moment où le message est envoyé n'a pas d'incidence sur la probabilité d'obtenir une réponse. Cependant leur analyse met aussi en évidence un certain nombre « d'incohérences » par rapport aux hypothèses de base qui ont été faites : dans leur premier message, les nouveaux arrivants utilisent davantage que les autres des formes rhétoriques qui prédisposent en principe à recevoir une réponse ; et pourtant, ils reçoivent en moyenne moins de réponses que les « anciens ». Cela suggère un modèle d'échanges plus complexe qui fait intervenir non seulement les formes rhétoriques, mais l'identité des locuteurs et/ou le contenu précis des messages. Par ailleurs, des tris des données effectués en considérant les grands domaines d'intérêt des groupes (santé, politique, sports) révèlent des différences importantes entre les groupes « santé » dans lesquelles le taux de réponse est élevé et les groupes « politiques » dans lequel il est beaucoup plus faible.

Ce qui précède montre aussi toutes les difficultés de ce type d'analyse statistique extrêmement complexe qui agrège une multitude de données¹⁷ ; on voit en particulier qu'une relation qui apparaît au niveau global peut résulter de l'agrégation de deux populations qui ont des caractéristiques différentes du point de vue étudié, d'où la fragilité des interprétations qui peuvent être tirées de l'observation de ces régularités, et ceci d'autant plus que les phénomènes observés sont de faible amplitude : +7% de réponses pour les messages avec des présentations des personnes ; +8% pour ceux dans lesquelles les pronoms à la première personne sont utilisés ; +6% pour des messages avec des questions. Limites dont d'ailleurs les auteurs sont conscients, ce qui les amène à des énoncés alambiqués :

¹⁷ Dans la première version de l'étude Arguello *et al.* (2006), ils considèrent 8 groupes et environ 6000 messages de début de fils) ; dans la seconde version Burke, M. *et al.* (2007), 99 groupes et près de 41000 messages.

« Finally, while we think it is plausible that the relationships we have identified in this paper between messages and the likelihood of getting a reply and between getting a reply and posting again are causal ones, which can be the basis of interventions to improve the success of online groups, our data are correlational and therefore do not necessarily prove causal relationships. (...) However, we cannot rule out the possibility that the relationships between messages and the likelihood of getting a reply and between getting a reply and posting again are artifacts, conditions on the existence of some unmeasured additional variables that cause both the independent variables and outcomes. » Arguello et al. (2006), p. 966

C'est aussi cette difficulté dans l'établissement des causalités compte tenu du jeu combiné de plusieurs variables à l'intérieur de leur propre modèle qui les conduit à entrer dans une démarche expérimentale (Burke, M. *et al.* 2007; Burke, Moira *et al.* 2010)¹⁸, reposant sur la manipulation de messages.

Il s'agit de tester « l'efficacité » relative (i) de références au newsgroup / références à la relation du locuteur au sujet d'intérêt commun du newsgroup, (ii) des questions explicites (point d'interrogation)/ questions implicites et (iii) des questions ouvertes / questions fermées. Pour chaque étude des messages datant d'au moins un an et appartenant clairement à l'une de catégories ont été extraits des groupes, puis manipulés pour les faire rentrer dans une autre catégorie. Le nombre de messages testés varie de 100 à 300 selon les études, et chaque groupe ne reçoit que quelques messages. Cela leur permet de produire des affirmations du type : *« when individuals attest to previous group participation and make specific requests for information, community responsiveness increases, but claims of shared identity with the group have no impact. »* Burke, Moira *et al.* (2010), p. 4.

Je ne chercherai pas ici à discuter des résultats de ce travail, mais plutôt à mettre l'accent sur les hypothèses implicites sur lesquels il repose : les newsgroups sont en un sens considérés comme des boîtes noires qui sont supposées produire – au moins de manière statistique – le même type d'outputs quand elles sont soumises aux mêmes inputs. À nouveau, on constate que la dimension temporelle n'est pas prise en compte :

¹⁸ Pour parachever ce panorama, signalons que ce groupe de chercheurs utilise aussi une autre méthode d'investigation qui consiste à confronter les messages à une population de sujets : s'agissant d'analyser l'impact de la politesse sur la propension qu'a un message à générer des réponses, ils recrutent environ 200 personnes sur le web à qui ils demandent d'évaluer sur une échelle allant de 0 à 7 le degré de politesse/ d'impolitesse de 48 messages chacun (au total 576 messages dans 12 newsgroups). Ils concluent sur le fait que 1) le degré de politesse moyen dépend des groupes) ; et 2) que l'effet de la politesse sur la propension à répondre est lui aussi variable : l'impolitesse semble augmenter les chances de réponses sur les groupes politiques, alors que c'est la politesse qui produit cet effet dans les groupes techniques. D'une certaine manière, leur analyse semble révéler l'existence de cultures langagières spécifiques, ce qui même du point de vue qui est le leur n'est pas sans soulever des questions sur les stratégies de recherche qu'ils ont employées en général.

les newsgroups sont censés être des systèmes stables ; pas plus que n'est prise en compte la diversité potentielle des participants : de ce point de vue, le faible échantillon constitué semble une limite importante de l'étude dès lors que l'on fait l'hypothèse de cette diversité.

Conclusion

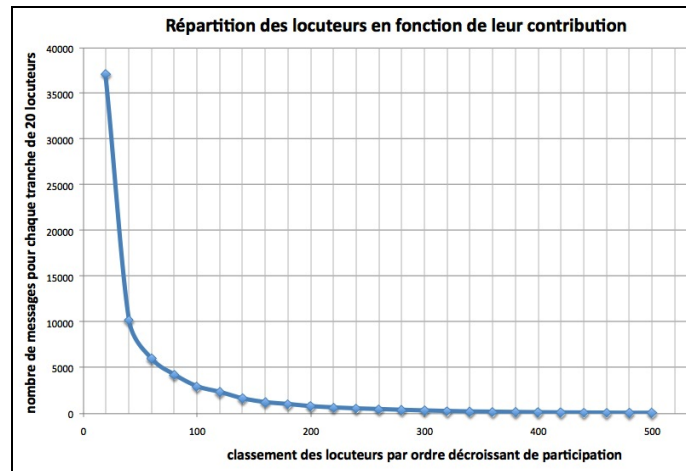
Au fond, tous ces modèles reposent sur des approches « structuralistes », au sens où ils font l'hypothèse qu'il doit être possible d'expliquer la dynamique globale des groupes à partir de la prise en compte d'un certain nombre de paramètres, plus ou moins indépendants des locuteurs et de la nature des discussions.

Ils construisent donc de fait une dualité entre un objet « macro », l'entité newsgroup, et les individus qui y interviennent, le newsgroup ayant en quelque sorte une vie propre. Bien entendu, cette apparente autonomie du newsgroup est en fait l'expression de déterminismes autres, qui agissent les individus dans leur ensemble : conventions socio-linguistiques, capacités cognitives par exemple. On peut se demander sur quoi ces analyses sont, en fait, susceptibles de nous apporter de la connaissance : d'un certain point de vue, on pourrait soutenir qu'il s'agit plus de traduire et de mettre à l'épreuve de nouveaux espaces de communication un certain nombre d'hypothèses sur le fonctionnement « ordinaire » des relations sociales que d'éclairer spécifiquement le fonctionnement des groupes électroniques de discussion. Par exemple, lorsque Burke, Moira *et al.* (2010) concluent que les questions « ouvertes » - c'est-à-dire qui laissent plus de champ aux interlocuteurs et sont donc plus « polies » - obtiennent moins de réponse que des questions plus fermées qui spécifient de manière directe ce qu'attend le demandeur, parce que les questions fermées demandent moins de travail d'élaboration de la part des répondants éventuels, sont-ils en train de nous parler des groupes de discussion ? Ou d'un mécanisme que l'on pourrait repérer dans d'autres contextes ? Le fait qu'ils citent, à l'appui de leur démonstration, une linguiste travaillant sur la communication dans le domaine de l'aviation oriente finalement vers cette dernière interprétation. Autrement dit, leur approche serait perpendiculaire à la nôtre, puisqu'à l'inverse, nous sommes intéressés à comprendre ce qui, de manière spécifique, se construit dans certains espaces de discussion.

Comprendre la morphologie des relations : les analyses en réseau

Différents auteurs (Adamic & Huberman 2002 ; Adamic & Huberman 2001; Barabbas & Réka 1999; Ekeblad 2001; Huberman & Adamic 1999; Shirky 2003) ont mis en évidence, dès la fin des années 90, l'importance de la loi de puissance dans les phénomènes en lien avec internet : la distribution des liens entrants et des liens sortants sur les sites web et des blogs, la répartition des visiteurs entre les sites, la distribution des sites en fonction du nombre de pages qu'ils comprennent, la répartition des locuteurs d'un groupe de

discussion en fonction de leur participation¹⁹ (nombre de messages postés), la répartition des fils de discussion en fonction du nombre de messages, la répartition des newsgroups consacrés à l'informatique en fonction du nombre de messages que chacun totalise (Raban & Rabin 2009) et même la répartition des délais de réponse aux messages postés (Kalman *et al.* 2006).



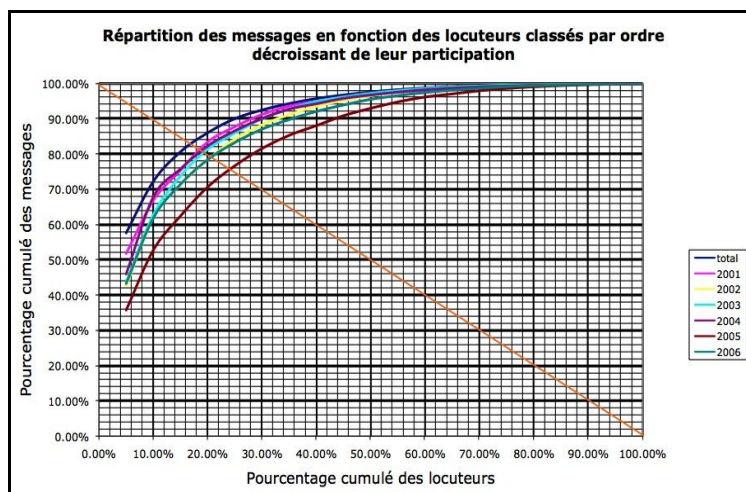
Graphique 1 : Répartition des locuteurs de la liste Naissance en fonction de leur contribution

A titre d'illustration, je présente une courbe issue de mes données. Comme on le voit, une répartition en loi de puissance se caractérise par son extrême inégalité. Une minorité d'entités (ici les locuteurs) représente une part très importante de la « richesse » (ici les messages) et à l'inverse un très grand nombre de locuteurs représentent une part très faible des messages.

La même distribution en loi de puissance peut aussi être représentée par un graphique du pourcentage cumulé des messages en fonction du pourcentage cumulé des locuteurs, comme dans le graphique ci-après dans lequel figurent les courbes pour chacune des années d'observation de la liste précédente. Cela permet de repérer pour chacune des courbes, à l'intersection entre la courbe et la droite d'équation $y=1-x$, le pourcentage X de locuteurs tel que l'ensemble de ces locuteurs représentent 100-X% des messages

En effet, les lois de puissance sont aussi connues par la loi de Pareto, ou la loi des 20 80 qui dit qu'environ 80% des effets en économie sont le produit de 20% des causes. La détermination du point exact de correspondance donne une idée du plus ou moins grand niveau d'inégalité de la distribution (ici varie entre 75% et 83%)

¹⁹ Whittaker *et al.* (1998) dessine dans son article la courbe de la répartition des locuteurs en fonction des nombres de messages qu'ils ont envoyés, sans relever qu'il s'agit d'une courbe en loi de puissance.



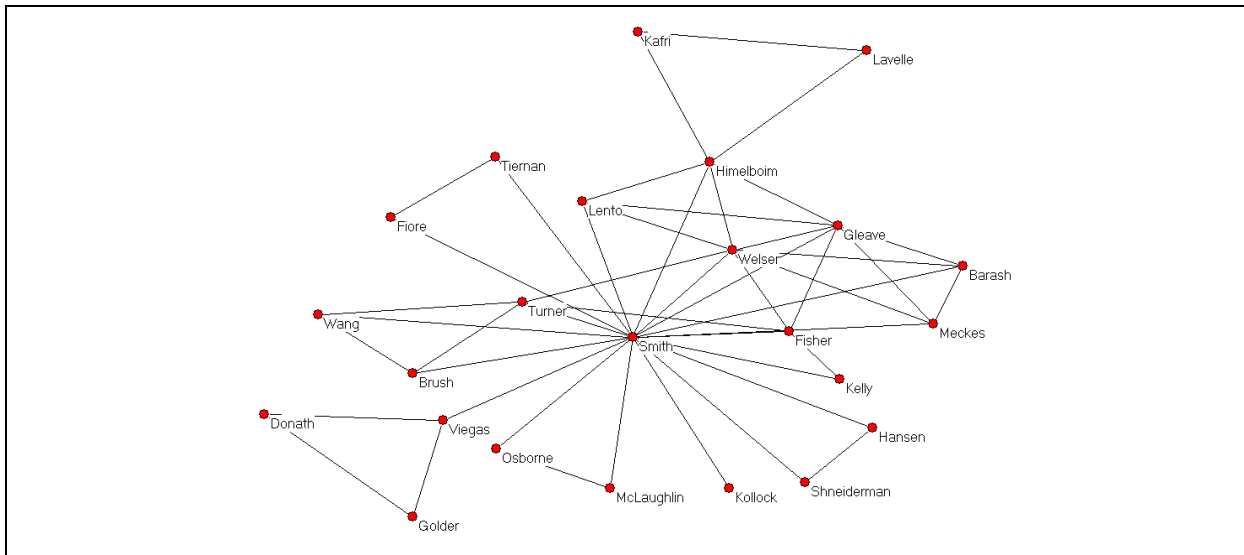
Graphique 2 : Répartition des messages en fonction des locuteurs (liste Naissance)

L'explication la plus fréquemment retenue pour rendre compte de ces distributions, notamment dans le domaine du web, est celle de l'« attachement préférentiel » (Barabasi & Réka 1999), qui repose sur une représentation des relations entre les différentes entités sous la forme d'un réseau : la probabilité qu'un nouvel entrant dans le réseau se lie avec un autre membre du réseau est proportionnelle à la somme des liens que ce membre entretient avec les autres membres du réseau. Mais il semble que cette question soit encore controversée : dans un article qui a fait récemment la couverture des *Proceedings of the National Academy of Sciences*, D'Souza *et al.* (2007) proposent un nouveau modèle qui introduit un principe d'optimisation et montre comment l'attachement préférentiel peut être considéré comme une conséquence de cette optimisation.

Quoi qu'il en soit, l'omniprésence des distributions en loi de puissance sur le web n'apparaît comme un principe explicatif en soi mais plutôt comme un constat qui n'épuise pas plus la spécificité des phénomènes à observer et à analyser que les distributions en loi normale dans d'autres domaines : ce constat doit alimenter la réflexion méthodologique et borner les efforts interprétatifs. En pratique, cependant, peu de travaux vont réellement au delà du constat lui-même.

En revanche, la question connexe du réseau a suscité tout un ensemble de recherches qui se sont déployées dans trois directions principales : la caractérisation des différents groupes ; l'identification de « rôles » ; l'analyse de la morphologie des fils de discussion.

Ces développements ont été en fait impulsés par un petit milieu de chercheurs nord-américains appartenant à des institutions diverses, universités, centre de recherche de Microsoft etc. mais formant un véritable réseau : sur la carte suivante, j'ai tracé ce réseau de relations à partir des co-signatures des articles des quelques 25 articles qui se trouvent dans ma base bibliographique. On constate que ce réseau est centré autour de Marc Smith, le sociologue qui a développé l'outil Netscan qui fournissait aux internautes comme au chercheur toute une série de statistiques sur les newsgroups Usenet.



Notons que dans tous ces domaines, le développement de formes de visualisation est une préoccupation constante en rapport avec la question que j'évoquais tout au début de ce parcours : comment produire de l'intelligibilité sur des masses de données ?

Derrière toutes ces analyses, on trouve une hypothèse commune : les messages qui circulent dans les groupes de discussion électronique dessinent des réseaux de relation entre les individus qui échangent. Bien que chaque message soit adressé à l'ensemble du groupe et sauf lorsque ce message vise à instaurer un nouveau sujet de conversation, dans la plupart des cas, il est adressé en réponse à un message précédent. En traçant les liens entre messages, il est alors possible de construire le réseau de relations entre les différents participants.

Signalons tout de suite quelques difficultés qui constituent des limites potentielles de ce type d'analyse :

- le message ne devient plus qu'un traceur d'une relation et son contenu n'est pas pris en compte. On se trouve ici dans une situation symétrique de celle que nous avons rencontrée sur le cas des analyses de contenu : dans ces dernières, seul le contenu est pris en compte alors que dans les analyses de réseau se concentrent sur l'en-tête du message. Cette réduction a néanmoins l'avantage de rabattre l'analyse des groupes électroniques de discussion sur des configurations largement explorées par les analyses de réseaux sociaux et de rendre ce travail homogène à celui qui peut être effectué sur les liens entre sites, les liens entre blogs et les réseaux sociaux au sens internet du terme.
- La signification du lien entre les messages mérite d'être interrogée : sur les newsgroups qui font l'objet de ces analyses en réseau, le lien résulte d'une contrainte imposée par le système. Lorsqu'un nouveau fil est lancé, il n'y a pas d'autre solution pour s'insérer dans la discussion que de « répondre » à un message. Ceci implique que le lien ne peut pas être considéré de manière univoque comme l'expression d'un

choix délibéré en faveur de tel ou tel locuteur²⁰. Le travail méthodologique de Jones *et al.* (2004), décrit précédemment, était motivé par le constat que les en-têtes ne rendent que partiellement compte des relations entre messages. Ekeblad (2001), qui essaie de tracer la trajectoire d'une discussion sur une liste, avec ces nombreuses ramifications, fait une analyse critique d'une méthode qui serait basée uniquement sur les liens de ce type ; bien qu'elle se place dans le cas d'une liste de discussion, les arguments peuvent être transposés sur le cas des newsgroups :

« Links between messages cannot be coded just by trusting that contributors have correctly used the response function of the mailer software. It often happens that the response function is used merely as a response function of the mailer software. It often happens that the response function is used merely as a convenience for getting the message addressed to the list, which can be totally misleading for the purpose of analyzing intermessage references. (...) The coding of reference links between messages also must allow for many-to-many relations: a posting may refer to several previous postings, as well as it may be referred to by several later postings. »

- Par définition et quoi qu'il puisse en être dit, cette approche qui s'inscrit clairement dans une perspective « réseaux sociaux » performe l'hypothèse que si A répond à B, c'est parce qu'il est B et non pas parce qu'il a soulevé un point qui intéresse spécialement A. Bien sûr, on peut rétorquer que l'un et l'autre sont en rapport, et que si A et B sont fortement liés, c'est sans doute qu'ils partagent des intérêts communs, fussent-ils conflictuels. Ceci étant, la méthode ne permet pas de tester ces différentes formes d'attachement. Mais on voit bien que l'on pourrait aussi considérer que toutes les personnes qui participent à une même discussion partagent des intérêts communs, ce qui ne produit pas la même définition des réseaux puisque les relations entre personnes telles qu'elles sont tracées par les messages ne sont pas transitives. Dans leur article consacré à l'élaboration d'une méthode permettant de repérer les experts dans un forum technique consacré au langage de programmation Java, Zhang *et al.* (2007) se démarquent de cette approche en considérant que, compte tenu de l'aspect question-réponse du Forum, les liens à considérer sont ceux qui unissent l'auteur du message initial et l'ensemble des personnes qui sont intervenues sur le sujet. Distinction reprise par Hansen *et al.* (2010) qui écrivent :

« There are two types of reply networks, depending on how you determine what constitutes a reply. The direct reply network connects a replier to the person they are

²⁰ Fisher *et al.* (2006) signalent que de fait, il ne s'agit pas de « répondre » à la personne qui a envoyé le message : *« Note that we use the word "reply" slightly loosely) ; within newsgroups, one does not directly reply to someone else, but rather posts a message in response to another's message. The distinction is subtle, but important: while we discuss interpersonal relations in terms of who replied to whom, we acknowledge that it is possible to post a follow-up message to a thread without giving any thought to the person who is being responded to. »* p. 3. Dans ces conditions, on peut se demander quel est alors le sens de tracer des réseaux de locuteurs) ; mais apparemment les auteurs ne perçoivent pas la dimension quelque peu paradoxale de leur démarche.

immediately replying to in the course of a thread. In contrast, a top level reply network connects all repliers within a thread to the original thread author. »

Sans préjuger des résultats que peut produire une analyse basée sur les liens entre messages, on ne peut manquer de constater que la première définition de ces liens a été de fait naturalisée dans la plupart des analyses, sans qu'il y ait de discussion autour de leur signification.

La caractérisation des différents groupes

Pour chaque locuteur, les analystes distinguent différents types de relations :

- Les réponses entrantes (In-Replies) : nombre de messages adressés en réponse à des messages du locuteur ;
- Les réponses sortantes (Out-Replies) : nombre de messages adressés par le locuteur en réponse à des messages d'autres participants ;
- les liens entrants : la liste des locuteurs qui lui ont répondu ;
- les liens sortants : la liste des locuteurs auxquels il a répondu.

Deux types de traitements sont effectués : la construction de réseaux accompagnée d'une analyse morphologique de ces réseaux ; un travail de comparaison de la distribution des liens entrants et des liens sortants. Ces traitements peuvent être accompagnés d'autres analyses statistiques.

L'hypothèse sous-jacente à ces analyses est que la forme que prend le réseau des relations entre les locuteurs synthétise certaines des caractéristiques du fonctionnement du groupe.

À partir d'un travail mené sur 9 groupes pendant 1 mois (de 188 à 2187 locuteurs), Fisher *et al.* (2006) produisent une première description d'un groupe, en mobilisant un certain nombre de données : le nombre de locuteurs ; le nombre de « réponses » pendant une durée donnée (ici un mois) – il s'agit de donner une idée de l'interactivité plutôt que de la simple activité – ; le pourcentage de fils avec deux messages et plus de 5 messages qui permettent d'avoir une appréciation de l'équilibre entre activités d'information et de discussion ; le pourcentage de locuteurs n'ayant envoyé qu'un message et de ceux qui ont envoyé plus de 10 messages qui apportent des informations sur le caractère plus ou moins « resserré » du groupe. À ces indicateurs, ils ajoutent les distributions des liens entrants et des liens sortants : comme bien d'autres paramètres, ces distributions suivent des lois de puissance. L'enjeu est ici d'en mesurer le degré d'inégalité : par exemple, une très forte inégalité dans la distribution des liens sortants accompagnée par une moins forte inégalité dans la distribution des liens entrants laisse supposer qu'un petit groupe de locuteurs répond à un nombre de locuteurs qui viennent poser des questions ou chercher des informations.

Enfin, ils tracent, pour chacun des locuteurs les plus importants, différents graphiques. Le premier représente le réseau défini par les locuteurs auquel « Ego » a répondu, enrichi des

relations que ces locuteurs entretiennent entre eux (Reply-to Network at distance 1) : cela permet de voir si « Ego » répond préférentiellement à des personnes liées entre elles de manière intense ou au contraire, s'il répond à des individus isolés. Le second graphique (Reply-to Network at distance 2) étend le réseau en intégrant l'ensemble des liens des locuteurs du précédent graphique : y figurent ainsi les personnes auxquelles les locuteurs précédents ont répondu, à l'exception d' « Ego », et les liens qu'ils entretiennent entre eux. Enfin, un histogramme donne la distribution des locuteurs présents sur le premier graphique, c'est-à-dire auquel le locuteur de départ a répondu, en fonction du nombre de personnes auxquelles ils ont répondu : il traduit d'une certaine manière une partie de ce qui est déjà représenté sur le second graphique et révèle si les personnes auxquelles « Ego » répond sont plutôt isolées, plutôt très connectées, ou de profils variés.

Sur les diagrammes reproduits dans la page suivante, on observe les profils de locuteurs supposés représentatifs de 5 newsgroups appartenant à des « genres » différents : informatique, le cerf-volant, l'art de la dispute sur internet, la politique, et le soutien. Le graphique du haut par exemple montre que le locuteur central répond à une multitude de locuteurs peu connectés entre eux qui viennent selon toute vraisemblance chercher des réponses à des questions techniques, et qu'il est par ailleurs connecté à quelques personnes ayant le même type de profil que lui. Situation bien différente de celle illustrée par le quatrième graphique, qui concerne un groupe politique et dans lequel on n'observe pas de polarité de ce type.²¹

²¹ Le même groupe de chercheurs propose dans un autre article Turner, T. C. *et al.* (2005) un schéma qui, d'une certaine manière, permet de faire plus ou moins les mêmes points: il s'agit de représenter chaque locuteur par deux variables, le nombre moyen de messages envoyés par ce locuteur sur les fils auxquels il a participé et le nombre de jours de présence. La forme et la localisation du nuage de points permet de faire la distinction entre groupes type « questions-réponses » et groupes type « discussion ». Cette représentation a été intégrée dans un outil de visualisation des groupes Usenet baptisé *Newsgroup Crowds* (Viegas, F. B. & Smith 2004)



On se trouve ici dans un cas où la différence entre les supports (newsgroups/ forums et listes de discussion) a une importance et empêche d'extrapoler une méthode d'un espace vers l'autre. En effet, dans les newsgroups et les forums, du fait de l'organisation matérielle qui présente les discussions d'une manière synthétique sous la forme de différents fils que l'on peut ensuite aller explorer, les fils sont susceptibles de durer très longtemps – parfois des années – et de constituer quasiment un sous-forum à l'intérieur du forum. De ce fait, il existe une importante variabilité dans le nombre de posts qu'un locuteur est susceptible d'envoyer sur un fil donné. En revanche, les listes de discussion sont accessibles de manière préférentielle par l'intermédiaire

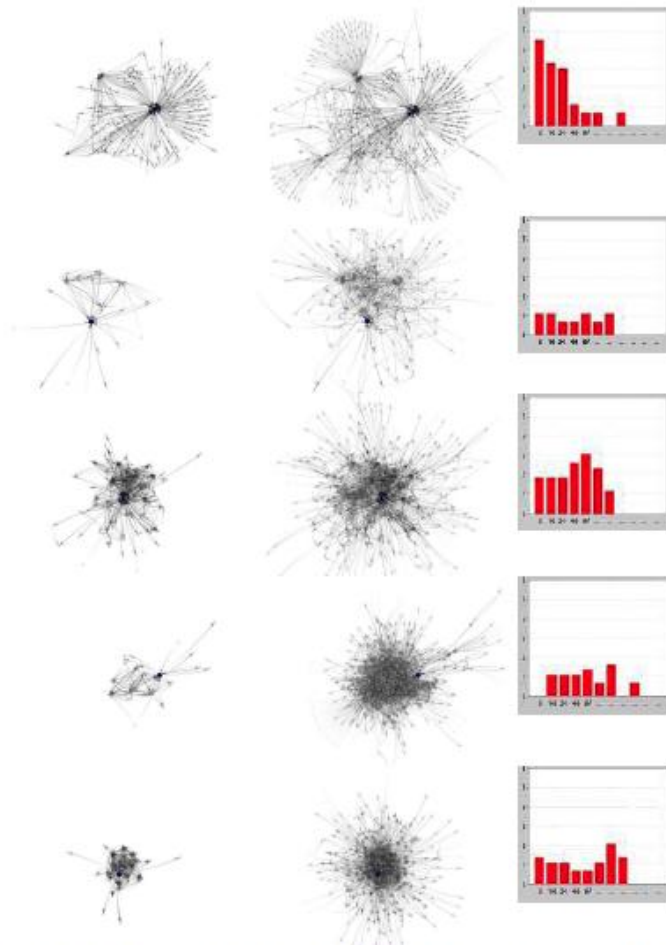


Figure 2. Reply-to network at distance 1 (left) and distance 2 (center), and out-degree distribution histogram, for one representative member for each of five newsgroups. From top to bottom, (a) *microsoft.public.windows.server.general*, (b) *rec.kites*, (c) *alt.flame*, (d) *alt.politics*, and (e) *alt.support.divorce*.

Fisher, D., Smith, M., & Welsler, H. (2006). *You Are Who You Talk To: Detecting Roles in Usenet Newsgroups*. HICSS '06: Proceedings of the 39th Annual Hawaii International Conference on System Sciences. IEEE Computer Society. 3, 1-10. (p. 5)

L'article propose ensuite un passage au niveau du groupe en construisant un profil agrégé : les auteurs considèrent les locuteurs qui ont répondu à plus de 10 personnes différentes et représentent la distribution moyenne des profils des personnes auxquelles elles ont répondu. Il s'agit de voir si, en moyenne, ces locuteurs répondent à des personnes isolées comme cela semble être le cas pour le groupe sur l'informatique, ou assez connectées, à l'instar de ce qui apparaît pour le groupe politique.

d'une boîte mail selon un classement a priori chronologique) ; du coup, la présentation par message conduit assez vite à ensevelir les messages précédents et par voie de conséquence les fils précédents : la variabilité du nombre de messages par fil et par locuteur est donc très faible et ne permet pas de faire des différences.

On peut distinguer deux types d'objectifs dans cette démarche :

- fournir des outils de description à l'analyste qui permettent d'identifier les locuteurs importants, de caractériser leur positionnement à l'intérieur du réseau de relations constituées par les échanges de mail et plus globalement, l'économie des relations à l'intérieur du groupe. De ce point de vue, le fait de se concentrer sur les locuteurs importants minimise la portée de la critique faite plus haut sur la méthode : (i) on peut considérer que l'effet statistique d'un grand nombre de liens minimise les biais éventuels, et (ii) c'est davantage la morphologie du réseau qui fait l'objet de l'analyse que son contenu en tant que tel.
- caractériser les groupes les uns par rapport aux autres, avec l'hypothèse que la nature de l'objet sur lequel s'est constitué le groupe prédétermine en partie la morphologie du réseau des relations.

Il y a donc une hésitation sur le statut de ce qui est observé et du coup, sur le cadre théorique adopté.

D'un côté, on a une analyse qui part du message et montre comment à partir de là, on peut reconstituer une figure d'ensemble qui soit intelligible : on se trouve ici clairement dans une perspective anti-structuraliste puisqu'aucune hypothèse n'est faite qui rapporterait les phénomènes observés à des causes générales externes, et la vision statistique qui en découle est compatible avec celle que Tarde, G. (1893 -1890) appelait de ses vœux (Latour 2010). L'analyse déploie deux caractéristiques importantes liées à l'approche en réseau : l'entre-définition des différentes entités qui le composent (le profil d'un locuteur est en quelque sorte défini par le profil des locuteurs avec lesquels il est en relation et réciproquement), et l'abolition de la distinction entre niveau micro et niveau macro (Callon & Ferrary 2006).

De l'autre, c'est-à-dire dès lors que l'on bascule dans une comparaison entre groupes, une forme de tentation structuraliste ressurgit : l'article suggère en effet (on ne peut pas parler de démonstration puisque seuls deux groupes sont mobilisés par classe) que les morphologies de réseau observées sont liées à la thématique des groupes appréhendée au travers de grandes classes telles loisir, informatique, soutien, politique. Il conclut en engageant à développer une analyse de « rôles » qui envoie un peu plus loin dans cette direction.

Des travaux ultérieurs permettent de voir où conduit cette tendance « structuraliste ». Par exemple, Himelboim (2008) poursuit ce travail de caractérisation par la comparaison de 15 groupes dans le domaine de la santé et de 15 groupes dans le domaine politique. Il fait une hypothèse massive en considérant que les groupes « politique » sont plutôt orientés vers l'échange d'opinions, alors que les groupes « santé » seraient orientés vers l'échange de faits/ d'informations, ce qui, en ce qui concerne ce dernier cas, ne résiste sans doute pas à un examen un peu rapproché des messages. Cette différence se traduirait par des structures contrastées du réseau

d'échanges et des conversations : dans le cas des groupes « politique », l'échange d'opinions donnerait lieu à des conversations longues, dans lesquelles les différents locuteurs interviendraient plus ou moins à parité, ceci induisant une relative homogénéité dans les profils des locuteurs ; dans le cas des groupes « santé », l'échange de faits conduirait à des conversations courtes et à des relations dissymétriques liées au fait que l'expertise est inégalement distribuée. L'analyse conduit à des résultats qui sont difficiles à expliquer, au sens où les causalités qui ont motivé les différents traitements sont difficiles à démêler : l'inégalité de distribution dans la répartition des liens apparaît corrélée avec le niveau d'activité du groupe (plus un groupe est actif, plus cette répartition est inégale) ; les groupes « santé » seraient plus inégalitaires dans la distribution des réponses que les groupes « politique », ce qui est interprété comme une confirmation de l'hypothèse selon laquelle les premiers seraient plus orientés vers l'information avec des inégalités d'expertise qui se traduisent dans les inégalités de distribution ; alors que les seconds seraient plus orientés vers la discussion. Cependant, la réalité des données conduit l'auteur à une conclusion « balancée » :

« The structure of a discussion network is affected by the type of information exchanged within it. Factual information invites a more centralized and hierarchical discussion, whereas an opinionated discussion invites relatively more egalitarian patterns. At the same time, one should also recognize that size of newsgroups held predictive power regardless of discussion type, and that most discussions were relatively low in diversity of contributors. This pattern suggests that there is considerable overlap between health communication and political communication, overlap that is sometimes overlooked when scholars treat the two domains as wholly distinct. » p. 173

Autrement dit, l'analyse ne permet pas vraiment de valider l'hypothèse structurale à l'origine de la démarche, et de faire une distinction tranchée entre les différents types de groupes.

La définition de rôles

Le pas suivant dans l'analyse des réseaux consiste à essayer de définir ce que les auteurs ont appelé des rôles, c'est-à-dire en première instance des profils particuliers de participation. Remarquons que cette démarche était en quelque sorte inscrite dans le travail de catégorisation des groupes puisque celui-ci partait justement de l'analyse des profils de participants. De façon congruente, différents paramètres vont être utilisés dans l'analyse : la proportion, dans les messages postés par un locuteur, de messages qui lancent de nouveaux fils et de messages qui répondent à d'autres messages ; le réseau à distance 1 des locuteurs auxquels le locuteur étudié a répondu (c'est-à-dire les liens à ces locuteurs additionnés des liens que ces locuteurs entretiennent entre eux) ; l'histogramme de répartition de ces locuteurs en fonction du nombre de liens sortants (nombre de personnes auxquelles chacun d'entre eux a répondu).

Welser, H. *et al.* (2007) cherchent sur cette base à déterminer ce qu'ils appellent la « signature structurale » du répondant, c'est-à-dire de locuteurs dont le comportement

dominant serait de répondre aux autres. L'originalité de leur démarche tient au fait qu'ils intègrent aussi une analyse de contenu de manière à pouvoir valider les indicateurs choisis : ils ont codé 10 à 15% des messages tirés au hasard pour les gros contributeurs et tous les messages pour les contributeurs n'ayant pas posté plus de 30 messages. In fine, le constat est mitigé : dans un seul des trois groupes qu'ils ont analysé, les indicateurs permettent de prédire de manière fiable le « rôle » de répondeur, et les auteurs débouchent sur le constat qu'un grand nombre de participants aux groupes jouent en fait différents « rôles », ce qui affaiblit la notion de rôle et renvoie plutôt vers une notion d'activité.

D'autres travaux (Gleave *et al.* 2009; Golder & Donath 2004; Himelboim, Itai *et al.* 2009a; Kelly *et al.* 2009; Turner, Tammara Combs & Fisher 2006; Turner, T. C. *et al.* 2005) se sont intéressés à cette question : différents types d'espaces ont été étudiés, des groupes techniques, des groupes politiques... pour essayer de mettre en évidence différents rôles. Ces rôles (catalyseur de discussion, discuteur, marginal, questionneur, allumeur de controverses, célébrité, novice, magnat, mentor etc.) sont définis de manière différenciée : toutes les appellations supposent un décalage par rapport à une description purement morphologique, mais l'intensité de ce décalage, autrement dit la part d'interprétation, est variable. Dans certains cas, cette interprétation repose sur des analyses qualitatives menées sur le contenu des échanges ou à partir d'entretiens.

Une même question se pose au sujet de ces travaux qui peut se décliner de différentes manières :

- la question de l'interprétation, c'est-à-dire de la manière dont on peut passer d'une description de la morphologie d'un réseau à l'attribution d'une « étiquette » qui renvoie à une signification sociale : on voit bien dans les différents travaux que cette liaison est fragile, toujours difficile à établir de manière satisfaisante, en l'absence d'une analyse des contenus effectifs des échanges. La volonté, manifestée par certains, de perfectionner l'analyse morphologique pour qu'elle produise des résultats de plus en plus conformes à ce que produirait une analyse du contenu dont on pourrait ainsi se dispenser paraît de ce point de vue un peu vaine .
- le choix du mot « rôle » pour désigner l'identification de profils particuliers induit une tension forte entre deux interprétations possibles qui font écho à la double perspective déjà soulignée dans la partie précédente : parfois, le rôle apparaît comme une configuration émergente qui naît des échanges ; d'autres fois, il apparaît comme une forme préexistante²² dont il s'agit de trouver la traduction dans l'univers

²² Pour Golder & Donath (2004), un rôle peut être défini de deux manières différentes, de façon formelle et de l'extérieur comme par exemple le rôle de parents, ou de façon informelle comme la conséquence d'interactions répétées et de pratiques acceptées (une personne dans un ami devient de fait celui qui planifie les activités du groupe). Dans tous les cas, (i) un rôle comprend deux composantes : « *The first component is related more to the individual and comprises the skills, privileges and responsibilities that that person enjoys* » ; *for this reason Goffman's analysis is important. The second is related more to the other people in a social context, and is embodied in the expectations those people have about the individual and his behavior* » ; Sheldon and

des groupes en ligne, parce que la prise en considération des rôles va permettre de mieux appréhender le fonctionnement et la dynamique des groupes :

« Building a catalog of social roles is an important first step towards understanding complex social systems. As a set of social role definitions is developed, these systems can be defined by the interaction of a relatively small set of roles that exist in different proportions, making up a variety of role ecologies. Faced with populations of millions of users, role identification allows these systems to be reduced to the interactions among a relatively small number of roles. With the capability to reduce social media system populations to simplified role ecologies, many practical and theoretical questions can be addressed. Are social media services better served by recruiting more leaders or a relatively larger population of moderate supporters? Are participants playing roles associated with enforcing informal social norms? » Gleave et al. (2009), p. 4

- pour résoudre en pratique cette tension, les auteurs proposent des montages méthodologiques qui sont censés permettre l'articulation entre ces deux définitions : Gleave et al. (2009) s'accordent sur un double processus, une analyse qualitative d'un côté qui doit permettre de repérer des rôles au travers de l'analyse du contenu des échanges, et, de l'autre, une analyse quantitative ou structurale qui s'attache à identifier des régularités comportementales. Dans des articles parus à un an d'intervalle (Turner & Fisher 2006; Turner. et al. 2005), Turner et ses collègues utilisent deux méthodes perpendiculaires pour identifier des rôles : dans le premier article, ils observent les configurations des relations ; dans le second, ils recourent de surcroît à une longue observation participante et à des interviews. De façon peu surprenante, les deux types de construction peinent à se recouvrir : ce ne sont pas seulement que les méthodes sont antagoniques, c'est simplement que chacune des constructions des « rôles » appartient à un monde spécifique, et qu'il n'y a pas de commune mesure possible entre eux.

Ces travaux ont permis de faire avancer la réflexion sur la manière dont on peut caractériser les différentes configurations de réseau dans lesquelles les locuteurs se trouvent : la prise en compte de certains paramètres permet notamment de repérer des singularités et d'introduire de la différence dans la maille continue des réseaux. Ils effectuent en revanche un saut interprétatif contestable quand ils convertissent ce qui, de notre point de vue, n'est autre qu'une analyse de « profil de participation » en analyse de « rôle », c'est-à-dire lorsqu'ils passent de critères morphologiques à des caractérisations sociales. Mais surtout, ce saut interprétatif correspond à un changement

Bettencourt's definition captures this succinctly » et (ii) s'appréhende de manière relationnelle. Gleave et al. (2009) définit le rôle de manière un peu différente : « We articulate a conceptual definition of 'social role' based on the symbolic interactionist approach to identity theory. That definition describes social roles as cultural objects that are "recognized, accepted, and used to accomplish pragmatic interaction goals in a community". Such roles operate as both sources of constraint and resources for action ».

de paradigme, puisqu'il fait basculer d'une analyse possiblement pragmatique à une analyse structuraliste.

L'analyse des structures conversationnelles : des fils au contenu

Dans les travaux précédents, les liens entre messages étaient en quelque sorte subsumés comme des liens entre locuteurs. Or, on peut aussi considérer qu'en faisant une telle opération, on perd une information cruciale autour de l'organisation des fils de discussion : quelques rares travaux ont essayé de s'en tenir à ce niveau du fil pour caractériser les activités se déroulant au sein des groupes de discussion.

Conein & Latapy (2008) sont partis d'une description des fils sous formes d'« arbres » dessinés en suivant la séquence des réponses : supposons qu'un locuteur envoie un message et que tous les locuteurs suivants répondent directement à son message initial, dans ce cas, le graphe se présentera sous forme d'éventail ; supposons à l'inverse que les locuteurs suivants répondent à chaque fois au dernier message qui a été envoyé, le graphe sera filiforme. En pratique, bien sûr, les fils d'un groupe de discussion présentent toutes les configurations possibles entre ces deux extrêmes.

Le propos des auteurs est de s'interroger, à partir d'une étude d'un newsgroup d'utilisateurs du logiciel libre, sur la contribution des outils de communication électronique sur la construction de formes de « cognition distribuée » qui seraient à l'origine de la production de nouvelles connaissances. Ils partent de l'hypothèse que (i) à l'intérieur d'un groupe comme celui qu'ils étudient, se créent des relations de conseil et des relations de coproduction de connaissances, et que (ii) ces deux types d'activités doivent se traduire par des morphologies différentes des fils.

Plus un fil est « large » (se déploie selon une forme qui rappelle l'éventail), plus l'échange consisterait en réponses connues à des questions posées ; plus le fil est « long », plus l'échange serait orienté vers la production de connaissances, cette tendance se renforçant avec la redondance des auteurs dans les fils. Deux personnes qui échangent longuement sur le même sujet seraient dans cette optique en train de construire de nouvelles connaissances.

L'analyse des structures des fils sur l'ensemble de la vie de deux groupes de discussion semble montrer qu'en fait ces deux formes d'interaction ne définissent pas des espaces séparés de communication, mais se nourrissent l'une l'autre : ce sont souvent des échanges de conseil qui donnent lieu au bout d'un certain temps à des discussions qui débouchent sur la production de connaissances. Le media en autorisant cette fluidité dans les échanges permettrait donc que se développe une véritable dynamique collective d'apprentissage et d'innovation. Si les conclusions sont intéressantes et stimulantes, on peut cependant regretter que l'échafaudage argumentaire soit affaibli par un saut interprétatif majeur, à savoir l'idée d'un couplage entre géométrie des fils et formes d'activité, couplage qui de toute façon, s'il était avéré, ne pourrait être valide que rapporté à des espaces bien spécifiés comme celui étudié. On retombe toujours sur la même difficulté, celle de l'articulation entre les deux faces du message, sa face relationnelle et son contenu.

Difficulté que quelques chercheurs se sont risqués à affronter : Ekeblad (2001) a tenté une autre forme d'exploitation des liens entre messages. Comme nous l'avons vu plus haut, son article critique assez vigoureusement le fait de considérer les liens tels que définis par la séquence des réponses comme représentatifs de la dynamique de la discussion. S'attachant à décrire la vie d'une liste de discussion sur les théories de l'apprentissage, elle invente une représentation graphique des liens entre les messages dans un fil de discussion ; ces liens ont été établis par la lecture attentive des messages : sont pris en compte le sujet débattu, les noms cités dans le message, les citations d'autres messages, voire les paraphrases d'un message précédent. Elle aboutit à des graphes tels qu'on peut en voir ci-dessous, dans lequel chaque petit rectangle correspond à un message, les traits représentant les liens entre messages. Cette représentation met à mal les interprétations univoques des liens entre messages : elle montre qu'il faut bien séparer l'architecture des liens telle qu'elle est imposée par le système technique et celle qui émerge des relations effectives entre les contenus des messages.

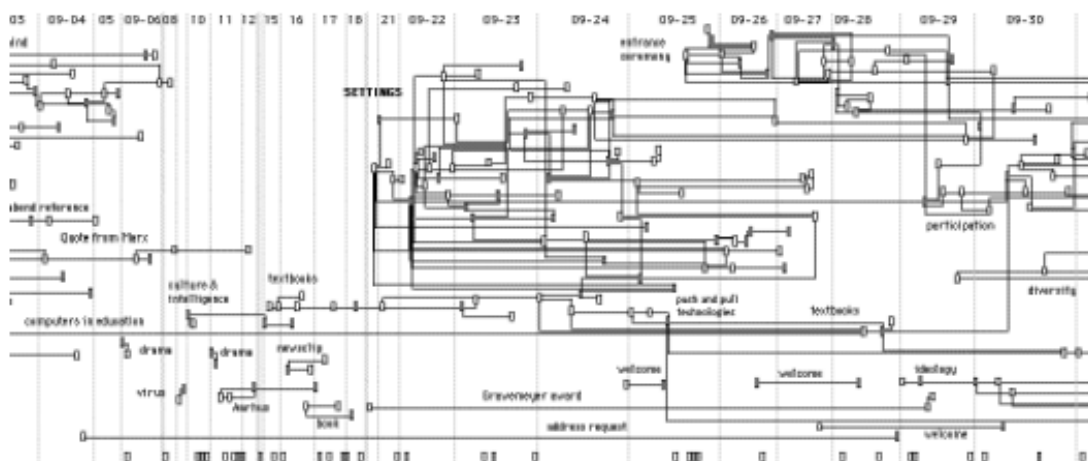
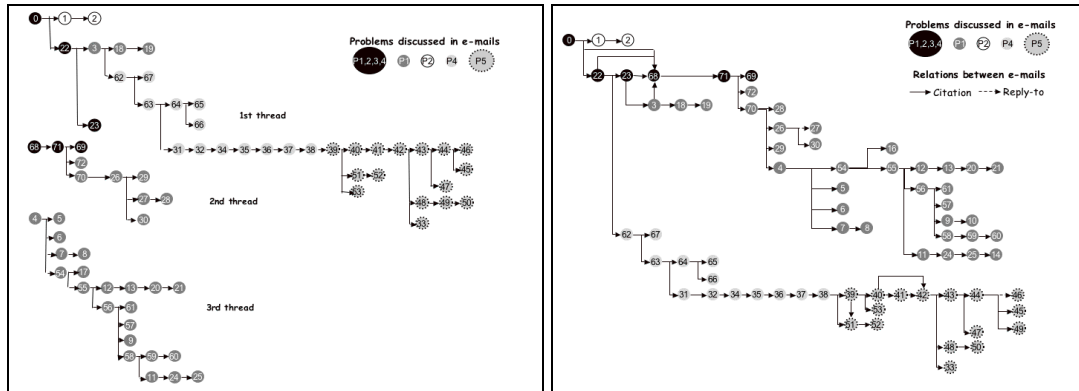


Fig. 7. The link map of the 1997 sample, including the Settings Multilogue.

Ekeblad, E. (2001). The Emergence and Decay of Multilogue. Self-Regulation of a Scholarly Mailing List.
<http://lchc.ucsd.edu/mca/Paper/eva/The%20Dynamics%20of%20Multilogue.htm>.

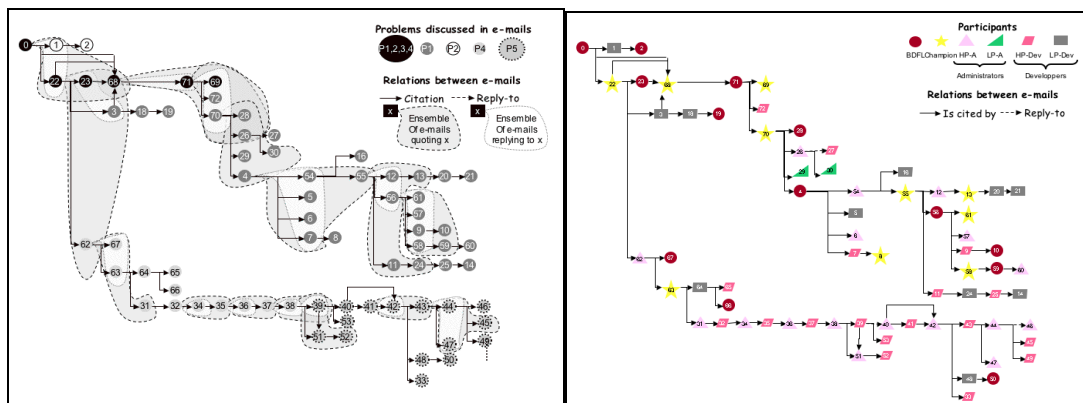
Barcellini et ses collègues (Barcellini *et al.* 2005; Barcellini, Flore *et al.* 2008) sont partis d'une position similaire à celle de Ekeblad : ils considèrent que les liens « techniques » entre les messages ne permettent pas de rendre compte de la dynamique d'échanges ; la « réponse » à une question soulevée dans un message émerge progressivement au fil des échanges caractérisés par des pratiques référentielles complexe. Ils vont développer une analyse centrée sur les inter-citations et non les liens entre messages.

Dans les graphiques présentés ci-dessous, ils comparent la structure de discussion que l'on obtient en considérant les liens entre messages à l'intérieur d'un fil d'une part (à gauche), et d'autre part les citations entre messages (à droite).



On constate immédiatement des différences flagrantes : tout d'abord, l'ensemble d'une séquence de discussion reliée du point de vue du contenu est en fait scindée en trois fils distincts ; ensuite, la structure même des échanges est très sensiblement différente d'une représentation à l'autre.

Ils vont beaucoup plus loin dans leur tentative de réconcilier les différents aspects dont nous avons vu qu'ils étaient traités de manière disjointe dans les recherches : leur analyse mixte ce travail de repérage et de traitement des citations, avec une approche thématique et une description des « fonctions » des différents intervenants – ils s'intéressent effectivement aux activités de conception dans des communautés Open Source, processus pour lesquels existe une distribution formelle des rôles.



Ces deux schémas extraits de Barcellini *et al.* (2005) représentent la discussion précédente, en faisant apparaître les thèmes (à gauche) et les intervenants (à droite).

Cette analyse leur permet d'inférer un certain nombre de résultats sur les activités de conception en ligne, sur le rôle des différents intervenants, sur l'organisation de la discussion, la prise de décision et le travail de coordination. Ils concluent néanmoins en

insistant sur la nécessité d'automatiser ce type de traitement qui repose pour le moment sur un travail « à la main ».

L'un des co-auteurs des travaux précédents, Warren Sack, a mis au point en 2000 un logiciel *Conversation Map* (Sack 2002), qui possède de ce point de vue des propriétés intéressantes, mais qui, comme tous les outils de traitement et/ou de visualisation²³ de ce type de corpus n'apparaît dans aucune publication autre que celles de l'auteur. Il s'agissait d'un navigateur de newsgroup Usenet²⁴ qui possédait les fonctionnalités suivantes : (i) tracé des réseaux sociaux correspondant aux liens entre messages ; (ii) calcul des réseaux sémantiques ; (iii) application de ces deux fonctionnalités à des individus, à des paires d'individus pour caractériser leurs relations, à des fils de discussion. Il est possible que l'outil, destiné à aider la navigation par le moyen de la visualisation, n'ait pas eu la souplesse nécessaire pour se prêter à des usages de recherche. Quoi qu'il en soit, on peut y voir l'amorce encore à développer d'outils capables d'intégrer la dimension socio-sémantique des groupes de discussion.

Cette intégration a fait l'objet de récents développements dans l'analyse des blogs (Cointet 2009). L'auteur développe une analyse en trois plans, social, sémantique, socio-sémantique à partir d'une part, des liens entre des blogs (citation, commentaire, liens dans le blogroll) et d'autre part, d'une caractérisation sémantique des billets. Parmi les intérêts de ce travail, on peut signaler une approche dynamique des relations, qui se traduit notamment par un modèle d' « émergence/ immergence » résolvant la question des relations entre micro et macro (qui se traduisait notamment sur les groupes électroniques de discussion par une dichotomie mal pensée entre individu et groupe) :

« Ces motifs et ces propriétés de haut niveau émergent des dynamiques individuelles mais ne sont pas pour autant inopérants par rapport à ces dernières. En continuité avec l'idée d'un individualisme méthodologique complexe (...), nous qualifierons d'immergentes les "causes induites" par les dynamiques de haut niveau sur les dynamiques de bas niveau. Dans cette perspective, Il serait vain de chercher lequel des deux niveaux précède l'autre dans une chaîne causale : les deux niveaux se co-déterminent l'un l'autre. Une telle modélisation permet également de résoudre le débat entre l'hypothèse d'une détermination du macro par le micro (hypothèse atomiste) ou du micro par le macro (hypothèse holiste) en les rendant compatibles au sein d'une boucle émergence/immergence qui lie de façon circulaire dynamiques de bas et de haut niveau. (...)

Micro et macro-sociologies se rejoignent donc en un même modèle qui intègre simultanément le rôle des motifs de haut niveau dans la détermination des dynamiques

²³ On peut citer par exemple l'outil *Themail*, qui traite le contenu des boîtes mail d'une personne : pour chacun des correspondants, il propose une sorte de frise temporelle qui pour chaque période va caractériser par des mots clés les échanges entre le correspondant et le titulaire de la boîte mail à partir d'une analyse de contenu. (Viegas, Fernanda B. *et al.* 2006)

²⁴ Il n'est plus disponible actuellement.

individuelles (ces motifs ou catégories ne sont pas données comme une variable exogène immanente, mais émergent des actions individuelles de bas niveau). Ainsi, les acteurs sont bien dotés d'un "libre-arbitre", ils ne sont pas des producteurs conscients d'un ordre "supérieur" (...) mais y contribuent par un effet de composition de l'ensemble de leurs actions ne s'exprimant que dans les limites induites par une forme d'ordre social préexistant. »

Il n'y a pas à notre connaissance l'équivalent d'une telle approche pour les groupes de discussion : cette approche est de toute façon à reconstruire partiellement en raison des différences qui existent entre les groupes et les blogs. Les blogs politiques étudiés par Cointet constituent un matériau beaucoup plus « propre » que les groupes à tout point de vue :

- Chaque unité d'analyse est pensée en tant que telle et déploie une argumentation qui est supposée se suffire, alors qu'une proportion importante de messages ne se comprennent que dans leur enchaînement dans le cours d'une discussion.
- Les registres de discours des blogs sont relativement homogènes : les formes de relation qu'ils manifestent avec d'autres blogs sont de nature argumentative, alors que les messages des groupes renvoient à une pluralité de relations, intellectuelles certes, mais aussi de soutien, de solidarité, d'amitié, d'information etc.
- Il n'y a pas à strictement parler d'équivalent des liens entre blogs : nous avons vu à quel point la considération des liens à l'intérieur d'un groupe de discussion est une question problématique, à la fois sur le plan conceptuel, mais aussi sur le plan technique.

Conclusion

Les analyses en réseau ont l'intérêt indéniable de fournir une vision « panoramique » de l'organisation des échanges. En exploitant le contenu relationnel des messages, elles permettent de décrire des asymétries dans la distribution des relations et de repérer les locuteurs importants dont l'activité est centrale dans le réseau.

Nous avons longuement insisté sur une double faiblesse de ces analyses :

- le fait qu'elles s'appuient sur une caractéristique liée à la configuration technique des groupes de discussion qu'elles transforment en élément signifiant de la structuration en réseau : tout l'édifice semble ainsi construit sur des fondations douteuses.
- la tentation « structuraliste » qui les habite et qui les conduit à postuler l'existence de relations univoques entre thématiques des groupes et caractéristiques structurelles des réseaux ou qui les amène à voir derrière certains profils de locuteurs la signature de « rôles sociaux » qui en deviendraient le principe explicatif.

Ces limites sont liées au découplage opéré entre l'aspect relationnel et le contenu des échanges ; comme nous venons de voir, certains travaux s'attachent dans la période récente à surmonter cette difficulté.

Vers une analyse dynamique des groupes de discussion

J'ai soulevé tout au long de cette description un certain nombre de problèmes posés par les différents partis pris méthodologiques. Ce travail n'a pourtant pas la prétention d'apporter une réponse à toutes ces questions.

L'objectif était de décrire l'articulation entre des choix de méthodes et des manières de concevoir l'objet de la recherche, en un sens de traiter la recherche comme une activité techniquement équipée et d'apprécier la manière dont cet équipement technique prédétermine en partie la configuration des objets produits. On a pu voir que la mise en œuvre de chaque technique d'investigation implique des choix importants : par exemple, indifférenciation/ différenciation des participants ; prise en compte du contenu/ prise en compte des relations. En même temps, ces techniques ne préjugent pas complètement de la nature même des interprétations qui peuvent être produites in fine : les analyses de réseau sont susceptibles de déboucher sur des interprétations structuralistes comme anti-structuralistes²⁵.

On peut néanmoins noter deux caractéristiques communes aux diverses méthodes, telles qu'elles ont été mobilisées jusqu'à présent :

Tout d'abord, il n'y a pas eu de réflexion très poussée sur les conséquences de la forte inégalité dans la distribution des participations à l'intérieur des groupes de discussion : or ce point nous paraît important au regard notamment de la question de la « communauté ». Avant même de se demander à quelles conditions certains groupes peuvent conduire à des formes d'engagement, quels qu'il soient, peut-être faut-il déjà interroger la notion de communauté, ou du moins se poser la question de ce que peuvent être des groupes dans lesquels existe une telle inégalité de participation. Les analyses de réseau ont mis en évidence un supposé contraste entre (i) des groupes plutôt orientés vers l'échange d'informations et dans lesquels la distribution des relations entre les locuteurs rend visible une asymétrie entre des « répondeurs » liés avec une multitude de « questionneurs » peu liés entre eux ; et (ii) des groupes orientés vers l'échange d'opinions et dans lesquels cette distribution des relations est beaucoup moins polarisée et où l'analyse en réseau débouche sur des « pelotes » de liens indémêlables (cf. graphique tiré de Fisher *et al.* (2006). Mais cette représentation est trompeuse dans la mesure où quels que soient les groupes, certains locuteurs occupent la plus grande partie de l'espace de communication. Que peut-on dire de ce sous-groupe ? Quelles relations entretient-il avec le reste des participants ? Quelle part prennent ces petits participants dans le dynamique du groupe ? etc. Ces questions n'ont

²⁵ Constat partagé par Paravel & Rosental (2003).

finalement été abordées nulle part et pourtant elles nous semblent importantes si l'on veut caractériser la « communauté » que constituerait le groupe de discussion.

La dimension historique des échanges n'est pas non plus prise en compte : la plupart des traitements sont effectuées sur des durées très courtes qui ont pour avantage de limiter la masse de données quelle que soit la méthode employée. Implicitement, les groupes sont considérés comme des entités dont les propriétés ne changent pas au cours du temps²⁶. Or, si l'on revient à la question de l'engagement, on ne peut se contenter de cette vision statique : il est nécessaire de comprendre comment dans ces groupes, s'enclenchent des dynamiques collectives et/ou individuelles, comment ils deviennent des lieux dans lesquelles se constituent des formes *d'agency*.

Dans la suite, nous allons, à partir du cas de deux listes de discussion, nous efforcer de produire une description de ces listes qui rendent intelligible leur fonctionnement en prenant en compte ces deux dimensions autrement négligées.

²⁶ Exception faite de Baym (2000a) qui consacre un chapitre de son livre à un « retour » cinq ans après sur le groupe qu'elle a étudié.

Partie II

Appréhender la dynamique collective des échanges

Notre objectif dans cette partie est de montrer comment à partir d'une série d'indicateurs relativement simples, il est possible d'appréhender la dynamique collective des échanges dans une liste de discussion. Il s'agit d'un travail de description de base qui vise à restituer les qualités ostensibles de l'activité d'une liste, plus ou moins telles qu'elles peuvent être perçues par un participant au travers des interfaces qui lui donnent accès aux messages, c'est-à-dire principalement sa boîte mail.

Les deux listes auxquelles nous allons consacrer cet effort de description concernent des questions en rapport avec la santé : l'une a rassemblé des personnes qui s'intéressaient à la grossesse et à l'accouchement, avec une attention particulière portée aux conditions de prise en charge médicale, l'autre concerne les parents confrontés au Trouble Déficit d'Attention/ Hyperactivité (TDAH). Par rapport à la question du lien entre espaces de discussion électronique et formes d'engagement, elles présentent des configurations différenciées :

- La liste sur la naissance a été créée sans qu'il y ait derrière d'organisation militante bien que, dès le départ, ses fondateurs entendent lui donner une coloration politique au sens large comme en témoigne la présentation de 2001 sur Yahoo groups²⁷

YAHOO! France Accueil - Yahoo! - Infos compte - Aide

Réglez vos comptes en direct YAHOO! Messenger

Bienvenue Invité Créer un compte - Ouvrir session

liste-naissance - liste de discussion sur le theme de la naissance [Rejoindre ce groupe !]

Description Catégorie : [Grossesse et naissance](#)

Bienvenue à tous les membres eGroups Pour en savoir plus

► **Accueil**
Messages
Tchat
Fichiers
Photos
Signets

Cette liste francophone est destinée à un échange d'informations et de points de vues concernant les actions menées autour de l'accouchement en France et dans les pays limitrophes ou francophones: projets de rencontres et de conférences, publications, soutien aux maisons de naissance, aux parents et aux sages-femmes.

La liste n'est pas modérée (tous les messages sont diffusés) mais toute demande d'adhésion est soumise à l'approbation des membres. Ce n'est donc pas une liste publique...

Rejoindre ce groupe !
Vous n'êtes pas reconnu en tant que membre de ce groupe.
Si vous l'êtes, c'est que vous vous êtes inscrit avec une adresse e-mail non convertie avec le compte Yahoo! que

Elle a finalement débouché sur/ été associée à la création de plusieurs associations militantes dans le domaine, mais ces associations ont été clairement distinguées de la liste et se sont dotées de leurs propres espaces de discussion : on est donc plutôt dans un processus d'essaimage.

- La liste sur le TDAH a été créée par une membre active d'une association. Cependant l'association a périclité assez vite, et s'appuyant sur la dynamique créée par la liste,

²⁷ page du groupe en novembre 2001, récupérée sur le site Web Archives.

certaines personnes ont fondé une nouvelle association. La liste est ici largement intriquée avec l'association : celle-ci est mentionnée dès la page d'accueil du groupe reproduite plus bas²⁸ ; les nouveaux venus reçoivent dès qu'ils s'inscrivent un formulaire qui leur permet d'adhérer s'ils le souhaitent.

The screenshot shows the Yahoo! Groups interface for the group 'thada_france'. At the top, it says 'YAHOO! Groupes FRANCE' and 'Accueil - Yahoo! - Aide'. Below the group name, it says 'Bienvenue, invité' and 'thada_france - thada en france'. There are links for 'Créer un compte - Ouvrir session' and '[Rejoindre ce groupe]'. A sidebar on the left contains navigation links: 'Accueil', 'Messages', 'Envoyer', 'Fichiers', 'Photos', 'Signets', 'Sondages', and 'Agenda'. The main content area has a 'Bienvenue invité(e)' message with links for 'Ouvrir session' and 'Rejoindre ce groupe'. Below this is a 'Description' section for 'déficit d'attention, hyperactivité, dyslexie, surdoudance, et autres ...' with the text 'Vous avez des questions ? ... N'hésitez pas à venir les poser sur ce forum'. It also mentions 'Vous pouvez aussi rejoindre l'Association HyperSupers Thada_France' with the address '2 sentier de la fontaine 77160 Provins' and 'Association loi 1901 N°0102001746' with the website 'www.thada-france.org'. On the right, there is an 'Infos groupe' section showing 'Membres : 248', 'Créé le : Septembre 22, 2001', and 'Langue principale : Français'. Below that is a 'Gestion du groupe' section with options like 'Figurer dans l'annuaire', 'Inscription restreinte', 'Non modéré', 'Tout le monde peut envoyer des messages', 'Archives publiques', and 'Les pièces jointes sont interdites'. At the bottom left, there is a star icon labeled 'Fondateur'.

Avant d'entrer dans le vif du sujet, précisons le statut, dans la démonstration, des deux listes que nous allons étudier : nous ne considérons pas que l'un ou l'autre liste soient représentatives d'une catégorie générale ; leur comparaison nous permet simplement de soulever des questions et de mettre à l'épreuve certaines interprétations. De manière congruente avec cette prise de position et avec la manière dont nous avons analysé la littérature dans le chapitre précédent, les résultats produits par l'analyse sur le fonctionnement des listes s'appliquent strictement aux listes étudiées : il ne s'agit pas pour nous de proposer un modèle général de fonctionnement des listes, mais plutôt de montrer comment avec certaines méthodes, il est possible de décrire l'activité d'une liste particulière.

Les outils que nous avons utilisés sont très variés – on pourrait parler de bricolage méthodologique – mais relativement simples (pas de modèle mathématique complexe) ou banalisés.

Il faut d'abord mentionner un logiciel de base Adyg qui a été réalisé par Frédéric Vergnaud²⁹ et qui permet de « nettoyer » les listes. Cela signifie d'abord homogénéiser les auteurs des messages : les individus utilisent des adresses mails et des pseudos variables dans le temps, ce dont les analystes de réseau présentés dans le chapitre précédent disent ne pas tenir compte car ce ne sont pas les individus en tant que tels qui les intéressent mais la structuration des groupes. En second lieu, il s'agit de reconstituer les conversations (messages qui figurent sous le même sujet), ce qui ne va pas de soi, car les sujets sont souvent déformés (ajouts de préfixes variables selon les logiciels de mail – Re :

²⁸ page du groupe en octobre 2002, récupérée sur le site Web Archives.

²⁹ Une première tentative avait été faite avec Andreï Mogoutov sous la forme d'un logiciel appelé Listman. En raison de problèmes récurrents, nous avons repris entièrement le processus.

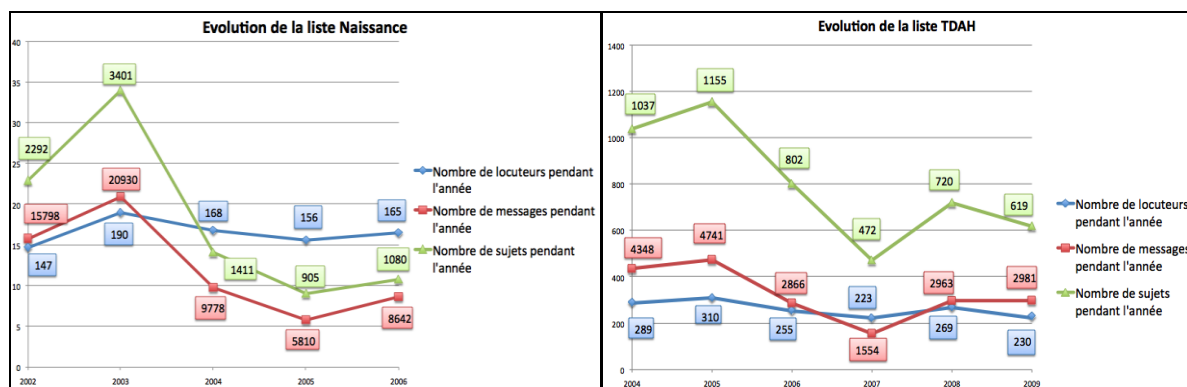
RE : Tr : Fwd : etc. –, introduction de caractères parasites, troncature du sujet). À partir de ce matériau de base reconstitué, le logiciel propose le calcul et l'exportation d'une série de statistiques portant sur les locuteurs et sur les conversations ; il permet aussi de faire de l'exportation de messages à partir de certains critères (locuteurs, mots dans le corps du message, mots dans le sujet ; nombre de messages dans la conversation ; nombre de locuteurs dans la conversation ; n° du message) ; enfin, il comporte un module d'aide à la lemmatisation.

À partir des fichiers produits par Adyg, plusieurs autres logiciels ont été mobilisés en fonction des questions posées et des besoins de traitement : Excel, Access, EndNote, Pajek (traceur de réseaux), Cordial Analyseur (analyse sémantique), Alceste (analyse textuelle).

Les listes, des espaces en constante évolution

Nous disposons des archives de ces deux groupes sur des durées longues :

- La liste Naissance a été créée au début 2000, mais elle a migré sur les Yahoo groups en avril 2001 et a été fermée en août 2007, ce qui représente environ 6 années d'archives, et de l'ordre de 70 000 messages.
- La liste TDAH apparaît sur les Yahoo groups à la même date que la liste Naissance ; cependant, les archives de 2001 à janvier 2004 ont été détruites : nous avons pu traiter les échanges de Janvier 2004 à Décembre 2009 soit une durée analogue de 6 années (20 000 messages)



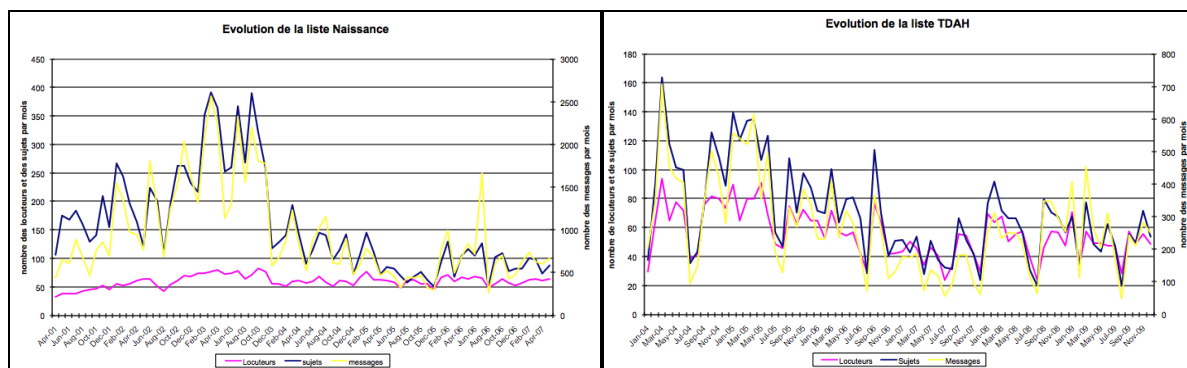
Graphiques 3 : Évolutions annuelles des listes Naissance et TDAH (locuteurs, messages, sujets)

Le volume d'activité, la multiplicité des intervenants et des sujets abordés sont les premiers éléments auxquels un abonné est confronté au travers de sa boîte mail. Ces différents éléments sont représentés sur les deux graphiques ci-dessus³⁰. On y voit que l'activité des listes peut fluctuer de manière importante d'une année à l'autre : le nombre

³⁰ L'échelle verticale n'a pas signification : les différentes variables ont été ajustées pour pouvoir être représentées à une échelle comparable.

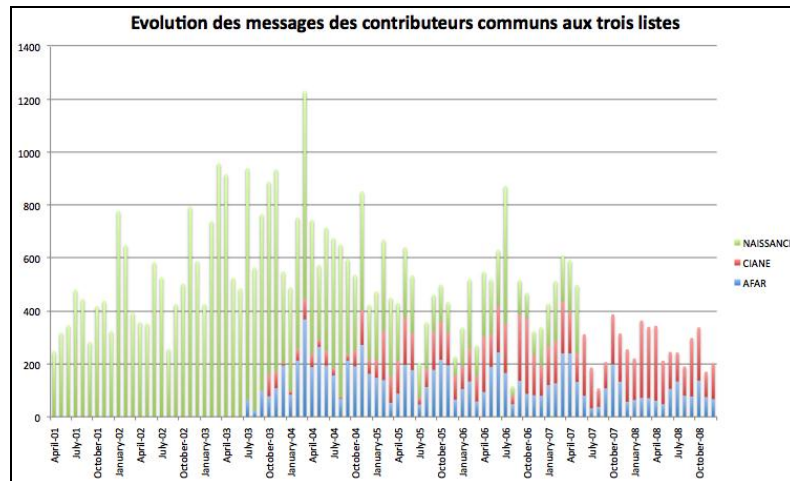
de messages par an varie d'un facteur 1 à 3 ou 4 entre la valeur minimale et la valeur maximale. Le nombre de conversations est grosso modo proportionnel au nombre de messages. En revanche, le nombre de participants (locuteurs) aux discussions est plus ou moins stable sur les deux listes, ce qui implique des changements de régime importants et relativement brutaux dans le fonctionnement des listes : entre 2003 et 2005, le nombre moyen de messages par locuteur sur la liste Naissance a plus ou moins été divisé par 3.

On note aussi que les niveaux d'activité moyen sont sensiblement différents sur les deux listes : sur la liste Naissance, le nombre moyen de messages par an et par locuteur varie de 100 à 280, alors que sur la liste TDAH, il se situe entre 35 et 65.



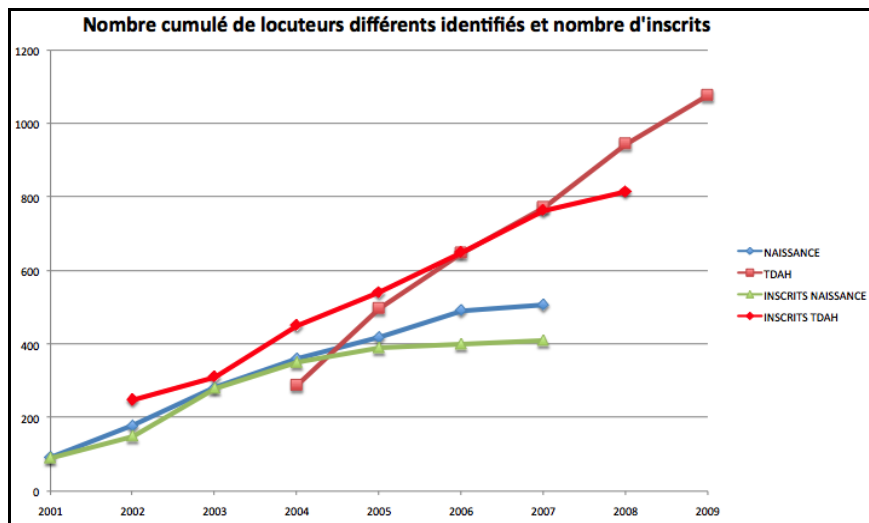
Graphiques 4 : Évolutions mensuelles des listes Naissance et TDAH

Ceci étant, considérées à une échelle différente, ces évolutions renvoient à des phénomènes différents : dans le cas de la liste Naissance, il apparaît clairement une rupture à la fin de l'année 2003 ; alors que dans le cas de la liste TDAH, la baisse moyenne d'activité semble davantage liée à une amplitude moins grande des périodes de forte activité. Le temps de la liste Naissance apparaît discontinu et donc inscrit dans une histoire : nous y reviendrons dans le chapitre IV ; disons simplement qu'à partir de 2003, certains des participants les plus actifs à la liste Naissance, dont le propriétaire, vont s'investir dans la création d'une association, l'AFAR, puis dans la participation à un collectif d'associations, le CIANE, toutes organisations dont le fonctionnement repose sur des listes de discussion ; ces participants ne vont pas quitter la liste Naissance, mais leur activité va évoluer progressivement, comme on le voit sur le graphique 5, en se déportant sur ces nouveaux espaces.



Graphique 5 : Évolution des messages des contributeurs communs aux trois listes, Naissance, CIANE, AFAR

Dernier élément général sur l'évolution des listes, la progression des inscriptions. Contrairement à ce que pourrait laisser penser la relative stabilité dans le fonctionnement des listes, le nombre d'inscrits n'a cessé d'augmenter tout au long du temps. On a représenté sur le graphique suivant l'évolution telle qu'on peut la reconstituer à partir des archives du web. Il est instructif de comparer les valeurs maximales du nombre d'inscrits avec le nombre de locuteurs différents que notre traitement a permis d'identifier sur la durée de fonctionnement de la liste. En ce qui concerne la liste Naissance, on constate que les deux courbes sont quasi-superposées, à l'exception de la dernière période, où le nombre de locuteurs identifiés dépasse le nombre d'inscrits. La courbe des inscrits pour TDAH est au départ plus haute que la courbe des locuteurs identifiés, mais compte tenu de la destruction des archives de la liste, il n'a pas été possible de reconstituer le « stock » initial entre 2001 et 2004 – il manque donc tous les locuteurs ayant participé durant cette période et ayant arrêté de participer en 2004.



Graphique 6 : Évolution des inscrits et des locuteurs identifiés

Deux remarques peuvent être tirées de l'observation de ces courbes :

- La proximité des chiffres (locuteurs identifiés grâce à leur participation ; inscrits) suggère qu'une majorité d'inscrits sont intervenus à un moment ou un autre sur la liste : on se trouve donc dans une situation bien différente de celles qui a été décrite par d'autres auteurs (Nonnecke & Preece 2000) pour lesquels le « lurking » est un phénomène massif. Ou du moins, cela amène à se poser des questions sur ceux que l'on considère comme des *lurkers* : pour une période annuelle donnée, la proportion d'inscrits qui ne postent pas n'est pas négligeable, puisqu'elle représente jusqu'à 75% des effectifs ; ceci étant, il est vraisemblable que la plupart de ces personnes ont participé précédemment, et qu'elles continuent de suivre les échanges sans ressentir la nécessité d'intervenir. Ceux qui, en fonction du dispositif d'observation construit, apparaîtront comme des lurkers ne correspondent cependant pas au portrait qui en est fait (Preece *et al.* 2004).
- Tout se passe comme si, une fois inscrit, l'abonné était attaché durablement à la liste qu'il y participe ou pas. La nature de cet attachement reste opaque, et il est de fait très rare qu'un locuteur resté passif pendant une année ou plus redevienne à nouveau actif.

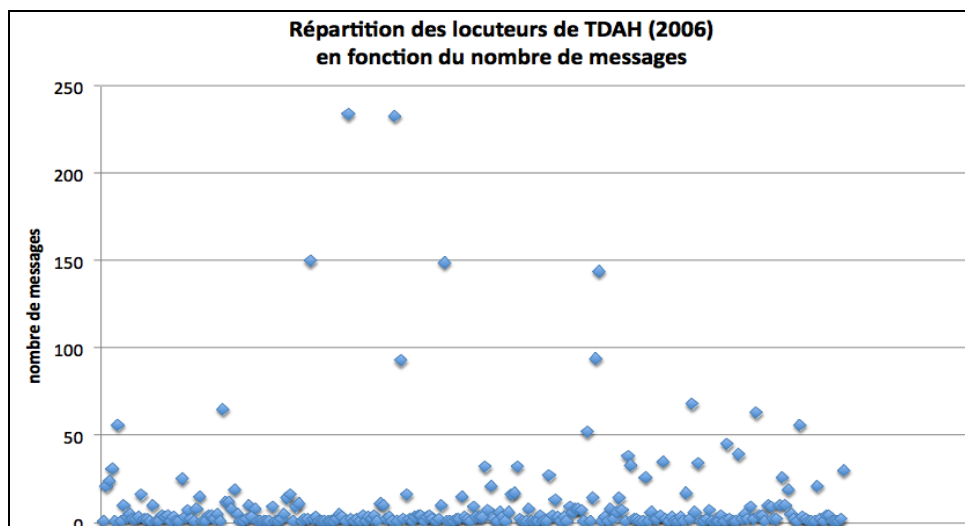
L'identité des participants est un autre élément crucial pour apprécier la dynamique des listes : y a-t-il une continuité de ce point de vue ou assiste-t-on à un renouvellement important des effectifs ? Rappelons que, dans le travail de Jones *et al.* (2004), seuls de l'ordre de 10 à 20% des participants étaient communs aux groupes étudiés d'un mois sur l'autre. Nous avons déjà exprimé un certain nombre de réserves par rapport à leur méthode. On peut les reformuler de manière un peu différente : les listes de discussion sont caractérisées par une forte inégalité dans la participation – des graphiques

représentant la répartition des locuteurs dans la liste Naissance ont été présentés plus haut des graphiques similaires pourraient être produits pour la liste TDAH – ce qui implique qu'un abonné à la liste voit dans sa boîte à lettres revenir de manière récurrente et importante certains noms, alors que d'autres ne sont présents que très épisodiquement. Évaluer la stabilité d'une liste de discussion ne se résume donc pas à évaluer la stabilité du groupe de participants, mais suppose d'évaluer aussi la stabilité dans la répartition des messages.

La question qui se pose donc est d'arriver à construire des instruments qui permettent de rendre visibles – et opératoires pour l'analyse – ces inégalités de répartition. C'est le détour que nous allons devoir faire avant de revenir vers la question de la stabilité.

La dynamique de participation

Pour rendre palpable autrement ce que signifie cette inégalité, nous avons positionné sur le graphique suivant les locuteurs (classés en ordre alphabétique) en fonction du nombre de messages qu'ils ont postés sur la liste TDAH durant l'année 2006.



Graphique 7 : Répartition des locuteurs de TDAH (2006) en fonction du nombre de messages

Confronté à ce type de graphique, n'importe quel observateur sera conduit à l'idée intuitive que les locuteurs présents dans le haut du graphique jouent un rôle important dans la dynamique des échanges et notamment dans leur « coloration ». Pour un abonné à la liste, ces locuteurs vont, si ce n'est personnifier la liste, du moins constituer une sorte de cercle particulièrement visible.

Pour pouvoir rendre compte de leur importance et analyser plus spécifiquement leur rôle, nous avons construit une catégorie, les « locuteurs dominants ». Nous sommes partis d'une analyse mensuelle : nous avons calculé pour chaque mois le nombre moyen de messages postés par des locuteurs ayant participé aux échanges, puis nous avons identifié les locuteurs dont la contribution était deux fois plus élevée que cette moyenne.

L'ensemble de tous les locuteurs qui, sur l'année, sont repérés en suivant cette procédure constitue le « groupe dominant ».

Quelques questions

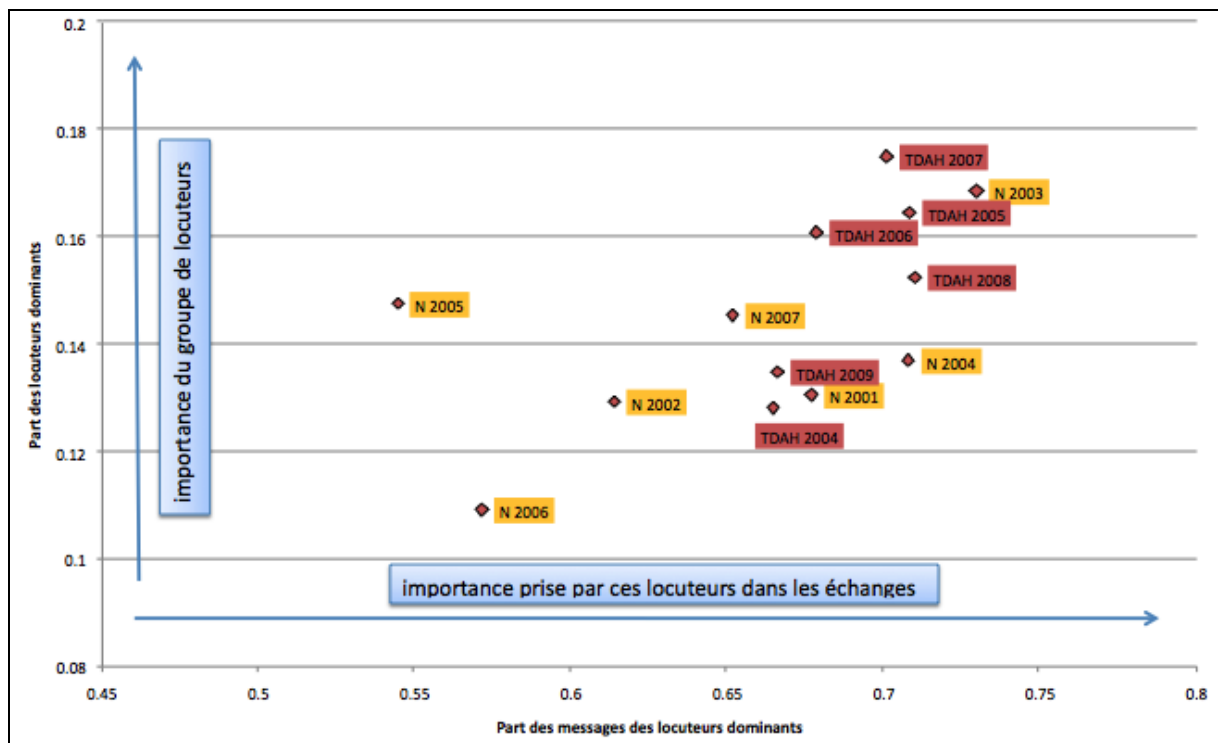
Pourquoi ne pas avoir pris une mesure annuelle ? La considération du point de vue de l'abonné a guidé notre choix : il nous a semblé que le point important était de repérer les locuteurs qui avaient joué un rôle visible dans la dynamique des échanges. Or, avec une mesure annuelle, on collecte des locuteurs qui peuvent participer de manière assez régulière mais peu remarquable ; à l'inverse, on manque des locuteurs qui n'ont été présents que quelques mois (parce qu'ils viennent d'arriver sur la liste ou s'en sont retirés) mais qui ont été très visibles pendant leur temps de présence.

Pourquoi ne pas avoir défini le groupe des dominants par un ratio de messages, c'est-à-dire comme l'ensemble des locuteurs les plus importants qui, à eux tous, représentent un certain pourcentage des messages, 80% par exemple si l'on se réfère à la loi de Pareto ? Ou à l'inverse pourquoi ne pas avoir pris les 20% de locuteurs les plus prolixes ? La réponse est très semblable à celle qui a été donnée à la question précédente : il n'y a aucune raison que l'ensemble ainsi déterminé ait un sens vis à vis de la perception qu'ont les participants de la liste. De surcroît, le ratio ainsi défini n'est pas spécifique de la liste étudiée : or, si la répartition des messages obéit sur toutes les listes à une loi de puissance, la forme de cette loi puissance, c'est-à-dire le niveau d'inégalité dans la répartition, n'a aucune raison d'être identique de l'une à l'autre des listes. Dans le graphique 2, on voit même que ce degré d'inégalité varie pour une même liste d'une année sur l'autre.

Limites de l'approche : de cette manière, on sélectionne quelques locuteurs qui ont pu avoir une brève « poussée » de fièvre participative et qui, ensuite, se retirent plus ou moins de la discussion. L'impact, faible, de cette « erreur » se fait sentir essentiellement sur le nombre de locuteurs dominants et la proportion que représente ce groupe (en nombre de personnes) par rapport à l'ensemble des participants. L'impact sur les analyses en nombre de messages est tout à fait négligeable.

Première remarque : la concentration des messages et l'importance du groupe des dominants sont variables

Sur le graphique suivant, on a fait figurer, pour chaque année d'observation de chaque liste, la part des messages représentée par les locuteurs dominants – que l'on appellera le **poids** du groupe dominant – tout comme la **taille** relative du groupe qu'ils constituent : ces valeurs sont variables, comme on peut l'observer puisque le groupe représente de 11 à 18% des locuteurs et de 54 à 73% des messages.



Graphique 8 :Évolution de la taille et de l'activité du groupe des dominants

Contrairement à ce que suggère Himelboim (2008), il n'y a pas de relation claire entre niveau d'activité d'un groupe et niveau d'inégalité dans ce groupe : certes le poids des dominants du groupe Naissance en 2003 est élevée alors que l'activité est maximale ; mais en 2004, cette activité chute brutalement, et pourtant cette part reste pratiquement identique alors que la taille du groupe dominant a diminué, ce qui indique une plus forte concentration de l'activité sur un petit nombre de personnes. Les positions de la liste TDAH en 2005 et 2007 sont relativement voisines et pourtant ces années occupent des points extrêmes du point de vue de l'activité. On pourrait multiplier les exemples du même type.

Deuxième remarque : les locuteurs qui appartiennent à un moment donné à un groupe dominant ont une durée de présence active sur la liste plus élevée.

De manière attendue, les locuteurs qui s'investissent le plus (en nombre de messages) sur les listes sont aussi des locuteurs qui restent actifs plus longtemps en moyenne. On a calculé pour tous les locuteurs, l'intervalle de temps en jours entre la première apparition sur la liste et la dernière apparition.

	locuteurs dominants	locuteurs non dominants
durée de présence active sur la liste TDAH	533	155
durée de présence active sur la liste Naissance	758	356

Tableau 1 : durée de présence active des locuteurs en jours

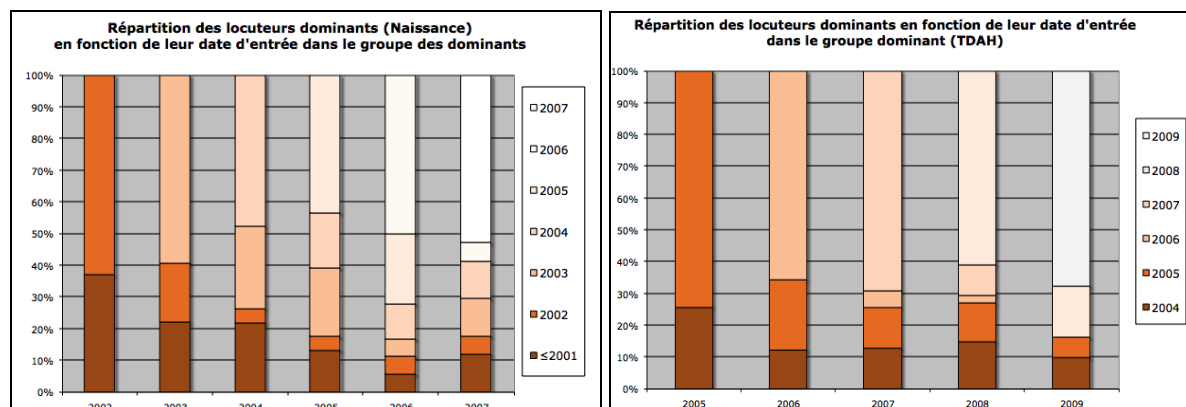
Le tableau 1 montre que les locuteurs dominants sont actifs en moyenne sur une durée de deux ans environ sur la liste Naissance (un an et demi pour la liste TDAH) alors que les non-dominants de Naissance sont actifs un an en moyenne (5 mois pour TDAH).

Il faut cependant garder à l'esprit (cf. tableau ci-dessous sur les données TDAH) qu'il y a des locuteurs non dominants qui restent actifs aussi longtemps que les locuteurs dominants, la différence de durée moyenne étant due au nombre important de ceux qui ne restent actifs que très peu de temps. La réalité des données empêche de fait toute caractérisation rapide des différents types de locuteurs (les locuteurs dominants restent longtemps, les autres s'en vont rapidement), car – et c'est un élément que l'on retrouvera tout au long de notre analyse – la variabilité à l'intérieur d'une classe quelconque (de locuteurs, de conversations...) est toujours très forte. Ceci signifie que les propriétés individuelles ne peuvent pas être inférées des propriétés du collectif.

durée d'activité en jours	non dominants	en %	dominants	en %
≥1000	48	5%	31	18%
≥500 et <1000	43	5%	31	18%
≥200 et <500	76	8%	38	22%
≥50 et <200	73	8%	58	33%
≥10 et <50	75	8%	16	9%
<10	587	65%	0	0%

Tableau 2 : répartition des locuteurs en fonction de leur durée d'activité

Troisième remarque : le groupe des locuteurs dominants se renouvelle continûment



Graphiques 9 : Répartition des locuteurs en fonction de leur date d'entrée dans le groupe dominant

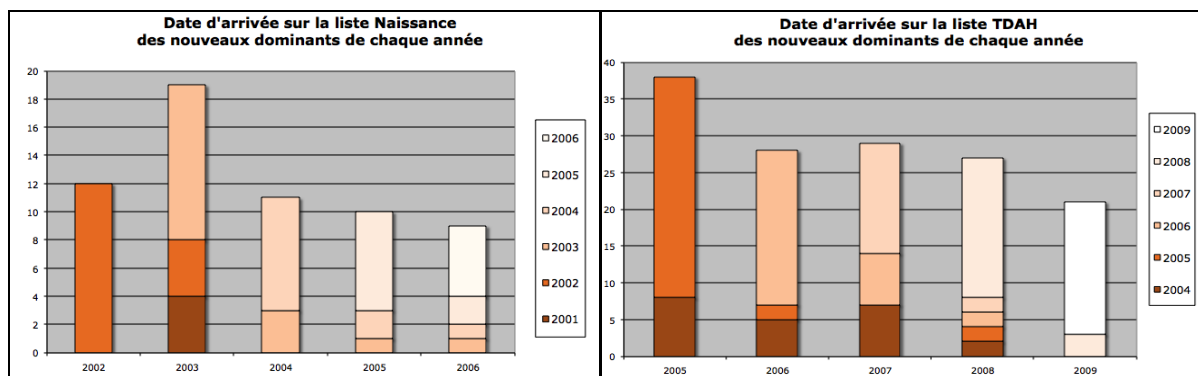
Sur ces deux graphiques, on a représenté la composition du groupe des dominants pour chaque année en fonction de « l'ancienneté » des membres : on constate qu'une part importante de ces membres n'étaient pas dans le groupe les années précédentes. Ce renouvellement est plus prononcé sur la liste TDAH (de l'ordre de 70% de nouveaux tous les ans) que sur la liste Naissance (de 50 à 60% de nouveaux par an).

Le renouvellement important des locuteurs dominants sur la liste TDAH est accompagné d'un écrasement très rapide de ce flux : à partir de 2006, très peu de nouveaux d'une année « survivent » à l'année suivante, laissant l'impression d'une césure entre un groupe plus ancien – composé de personnes arrivées avant 2006 – et qui se maintiendrait alors que les nouveaux arrivants seraient très volatiles.

Sur la liste Naissance, le taux de renouvellement est moins important et les nouveaux venus restent plus longtemps dans le groupe³¹.

D'où sortent ces nouveaux dominants qui surgissent chaque année ? Pourraient-ils être d'anciens de la liste ? L'impact du renouvellement des effectifs sur la vie de la liste serait alors a priori limité. L'examen des données montre que ce n'est pas le cas, comme illustré dans les graphiques suivants.

³¹ Le caractère atypique de l'année 2007 est dû au fait qu'il s'agit d'une année partielle (5 mois) et que les discussions autour de la fermeture de la liste ont fait redevenir actifs des locuteurs qui s'étaient mis en retrait.



Graphiques 10 : Date d'arrivée sur les listes des nouveaux dominants de chaque année

Les nouveaux dominants sont en large majorité (60 à 70%) de nouveaux arrivants sur la liste. Les comportements de novices décrits par Burke, Moira *et al.* (2010) ou Golder & Donath (2004) – période d'observation ou de faible activité le temps d'acquérir une connaissance suffisante du groupe avant de s'y risquer – ne semblent pas s'appliquer dans nos cas. Dans le cas de la liste Naissance, on peut même être sûr du contraire puisque l'organisation de la liste – en obligeant à une présentation au moment de l'inscription – fait coïncider la date d'inscription avec la première apparition sur le groupe. On ne peut donc pas considérer l'appartenance au groupe des dominants comme une sorte d'autopromotion qui interviendrait après une période d'apprentissage : certains locuteurs sont d'emblée très actifs, et s'ils obtiennent des retours qui leur conviennent, le restent pour une période de temps significative.

Quatrième remarque : les locuteurs dominants ont majoritairement des positions de dominants tout au long de leur durée d'activité

On peut aller plus loin par rapport au constat précédent, en analysant la « carrière » des locuteurs dominants, c'est-à-dire en regardant pour l'ensemble de la durée de leur activité combien d'années ils ont proportionnellement passé dans le groupe dominants.

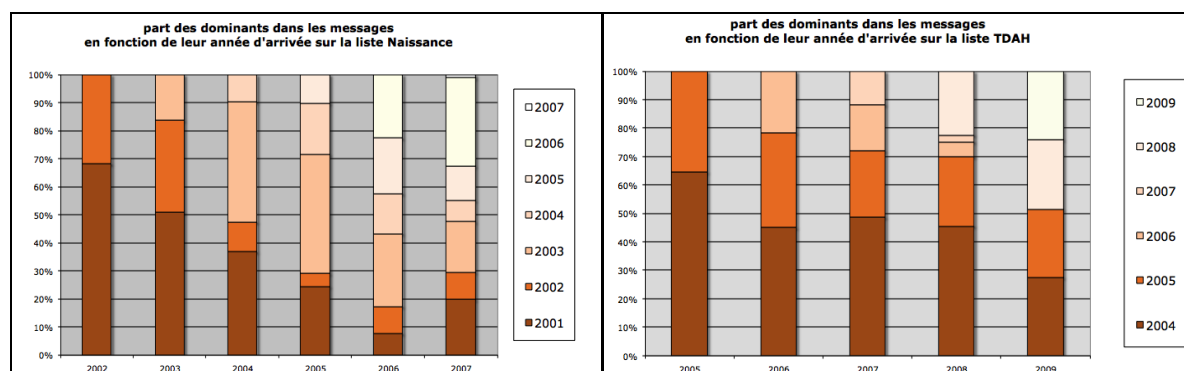
	Naissance	TDAH
ont toujours été dominants	40%	49%
DuréeDom ≥ DuréeNonDom	31%	34%
DuréeDom < DuréeNonDom	30%	17%

Tableau 3 : Répartition des dominants en fonction de leur durée d'activité en tant que dominants

Le constat est semblable sur les deux listes : les locuteurs qui n'appartiennent qu'une minorité de leur temps de participation au groupe des dominants représentent une part faible des locuteurs dominants (30% pour Naissance, 17% pour TDAH). Il y a donc ce que l'on pourrait appeler des carrières de dominants, qui sont marquées par un fort investissement des participants tout au long de leur durée d'activité au sein de la liste. Ceci valide dans une certaine mesure la méthode que nous avons utilisée pour construire le groupe des dominants : on constate que ces locuteurs présentent une relative homogénéité de comportement.

Cinquième remarque : le poids des locuteurs dominants est proportionnel à leur ancienneté sur la liste.

Quel est l'effet de cette dynamique des locuteurs sur les messages eux-mêmes ? Les répartitions en loi de puissance ont ceci de particulier qu'elles s'appliquent quelle que soit l'échelle d'observation, s'apparentant de ce fait aux fractales. Ceci signifie qu'à l'intérieur du groupe dominant, il existe une grande inégalité de répartition à l'image de celle qui prévaut pour la population entière : le tableau 2 des durées d'activité en donne d'ailleurs une idée puisqu'à l'intérieur du groupe dominant, figurent des locuteurs qui ont plus de 1000 jours de durée d'activité (voir même 2000), et d'autres qui se situent en dessous de 50. Le renouvellement des effectifs pourrait ne pas se traduire par une transformation très visible dans la boîte mail si les nouveaux entrants sont de « petits » locuteurs dominants.



Graphiques 11 : part des dominants dans les messages / année d'arrivée sur la liste

Comme on le perçoit dans ces graphiques, la part prise par les « anciens » dominants reste importante au fil du temps, notamment pour la liste TDAH où les locuteurs arrivés en 2004 et 2005 représentent toujours la moitié des messages des dominants en 2009, soit plus d'un tiers de tous les messages de la liste.

Pour rendre plus palpable cette disproportion entre la part de messages représentés par une catégorie de locuteurs et la part qu'ils représentent dans les locuteurs, nous avons calculé le rapport entre ces deux proportions, soit la **densité** du groupe dominant : un groupe qui représente 20% des effectifs de la liste et qui poste 40% des messages aura une densité de 2.

Nais.	2002	2003	2004	2005	2006	2007
2001	1,85	1,48	1,69	1,87	1,36	1,12
2002	0,50	1,05	1,22	1,05	1,74	1,66
2003		0,47	1,23	1,22	1,56	1,53
2004			0,28	1,05	2,58	0,64
2005				0,34	0,59	0,70
2006					0,68	1,07
2007						0,20

TDAH	2005	2006	2007	2008	2009
2004	1,57	1,85	1,46	1,87	2,84
2005	0,60	1,37	3,03	2,00	2,47
2006		0,42	0,80	0,72	
2007			0,31	0,24	
2008				0,48	1,09
2009					0,41

Tableau 4 : rapport entre part des messages et part des locuteurs

On constate ainsi que sur les deux listes, les derniers arrivants postent toujours de manière beaucoup moins intense que les plus anciens, puis l'intensité de leur participation semble augmenter au fil du temps. Ceci n'est pas dû principalement à un effet d'apprentissage : l'examen des contributions des dominants montre que certains postent davantage la deuxième année que la première mais l'inverse est vrai aussi ; cela semble plutôt à mettre en rapport avec le tassement des effectifs de chaque promotion au fil du temps, ceux qui restent le plus longtemps étant les plus actifs.

Conclusion

Dans cette partie, nous avons essayé de construire une catégorie, celle de locuteurs dominants, qui permette de prendre en compte l'inégalité de distribution des contributions dans les listes de discussion et d'analyser la dynamique des échanges.

Nous avons à cet effet élaboré une série d'indicateurs, parmi lesquels :

- Taille du groupe dominant : part représentée par les locuteurs dominants par rapport à l'ensemble des locuteurs
- Poids du groupe dominants : part représentée par les messages des locuteurs dominants par rapport à l'ensemble des messages
- Durée de présence des locuteurs dans le groupe dominant et part représentée par cette durée dans l'ensemble de leur durée d'activité sur la liste
- Répartition par ancienneté des locuteurs dominants et de leurs contributions à la liste

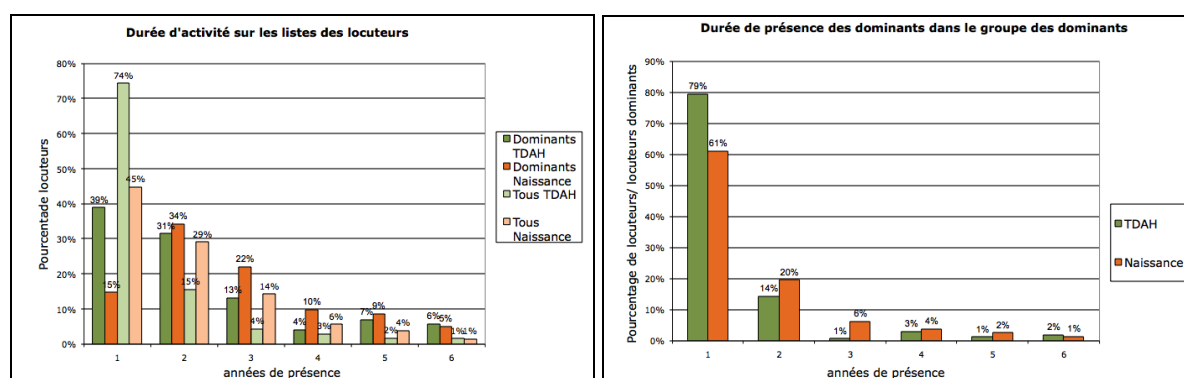
Ces locuteurs dominants constituent une catégorie distincte des autres locuteurs dans la mesure où une grande partie d'entre eux ont un comportement spécifique : ils ont un niveau d'activité élevé en permanence et une durée d'activité sur la liste longue.

Chaque année, le groupe des locuteurs dominants connaît un renouvellement important (50 à 60% de nouveaux sur la liste Naissance, 70% sur la liste TDAH). Cependant, ce renouvellement ne se traduit pas immédiatement par une transformation radicale dans

la répartition des auteurs de mail, car plus les locuteurs sont anciens et plus ils sont proportionnellement actifs. Ce qui conduit à faire l'hypothèse d'un noyau dur de locuteurs qui constituent le socle de la liste, à la fois par l'intensité de leur participation et par la durée de leur inscription sur la liste, noyau qui évolue lentement et qui est susceptible d'assurer une certaine continuité dans le contenu et la forme des échanges.

Les développements précédents ont aussi permis de mettre en évidence des traits distinctifs entre les deux listes :

- Une temporalité différente : sur la liste Naissance, les durées d'activité sont plus longues pour tous les locuteurs, ce qui se traduit par un investissement plus durable des locuteurs dominants.



Graphiques 12 : Comparaison des durées d'activité des locuteurs et des locuteurs dominants

- Une forte opposition sur la liste TDAH entre un petit noyau de locuteurs dominants installés sur la durée et contribuant de manière importante à la discussion et un flux de locuteurs dominants se renouvelant à l'inverse très rapidement.

Se dessinent ainsi deux formes de continuité de la liste un peu différentes : une continuité plus diffuse et portée par un large ensemble de locuteurs très impliqués – niveau d'activité moyen élevé – sur la liste Naissance ; une continuité incarnée par un petit groupe durable et actif sur la liste TDAH.

La dynamique des discussions

Les constats faits précédemment amènent à s'interroger sur les liens qui existent entre petits locuteurs et locuteurs dominants : jusqu'à quel point peut-on considérer que l'ensemble des locuteurs forment une « communauté » de discussion ? Existe-t-il ou non une segmentation dans les échanges entre différents types de locuteurs qui auraient des intérêts différenciés ? Comment qualifier les contributions des uns et des autres à la discussion ?

Plusieurs éléments, susceptibles d'intervenir sur le caractère plus ou moins inclusif et partagé des échanges, ont été mis en avant par les différents travaux dont nous avons rendu compte :

- Les réactions du groupe vis à vis des messages qui lancent une nouvelle discussion, et notamment lorsque ces messages émanent de locuteurs qui ne sont encore jamais exprimés (Arguello *et al.* 2006; Burke *et al.* 2007; Burke, Moira *et al.* 2010).
- La « structuration » de la discussion avec en particulier une opposition entre (i) un modèle orienté vers l'information qui suppose une prégnance de la forme Questions/Réponses et une division entre experts et non experts et (ii) un modèle orienté vers l'échange d'opinions (Himmelboim 2008 ; Turner *et al.* 2005; Welser *et al.* (2007), les groupes autour de problèmes de santé étant considérés a priori comme relevant de la première espèce. Cette configuration serait marquée par une certaine dissymétrie entre des questionneurs qui lancent des sujets de discussion et des répondeurs qui apportent leur expertise, ces derniers concentrant une très grande activité.

Nous ferons l'hypothèse, avec d'autres, que les conversations longues qui associent de multiples locuteurs sont emblématiques de la vie collective d'une liste : dans ce cadre, la composition des sous-groupes qui participent à ces conversations est un élément qui permet d'apprécier le caractère plus ou moins « communautaire » de la liste.

Nous nous proposons donc dans la suite de nous concentrer sur une série d'indicateurs qui permettent d'apprécier ces différents éléments :

- La répartition entre les différents locuteurs du travail de réponse aux nouveaux arrivants
- La répartition entre les différents locuteurs du travail de réponses aux messages qui débutent une nouvelle conversation
- La répartition des locuteurs dans le lancement des différents types de conversations (courtes/longues)
- La répartition des locuteurs dans la réponse aux différents types de conversation (courtes/longues, type de lanceurs)

En préalable, il est nécessaire de préciser la manière dont la différence entre conversations courtes et conversations longues est construite.

Nous appelons conversation l'ensemble des messages dont le « sujet », tel qu'il apparaît dans la boîte mail de l'utilisateur, est identique. On appellera ici « fil de discussion », l'ensemble des messages qui sont liés entre eux par une chaîne d'actions de « réponse ».

Nous avons vu précédemment qu'une grande partie des travaux sur la structuration des fils de discussion s'appuie sur les fils de discussion et non sur les conversations. Les difficultés que pose une telle position ont aussi été commentées. Notre choix en faveur de la conversation plutôt que du fil est motivé par deux considérations principales en dehors des réserves déjà exprimées :

- le sujet – qui définit la conversation – est un élément organisateur pour l’abonné puisque c’est celui qui apparaît dans la boîte mail et qui permet éventuellement des tris. Dans les premières années des listes de discussion, il était fréquent que les participants à une liste de discussion soient rappelés à l’ordre quand, à partir d’une première discussion, ils embrayaient sur une autre discussion sans changer le sujet apparent, preuve s’il en est besoin que cette information est considérée comme importante pour l’utilisateur.
- l’indication du « parent » d’un message (c’est-à-dire l’identifiant du message auquel il répond) est une donnée de qualité moyenne : elle est notamment manquante de manière systématique pour certains utilisateurs, ce qui est assez problématique.

Nous adopterons un mode de définition pour les conversations longues analogue à celui que nous avons adopté pour la définition des locuteurs dominants. En effet, la répartition des conversations par nombre de messages / par nombre de locuteurs différents obéit aussi à une loi de puissance.

Les inégalités qui résultent de cette distribution sont perceptibles de la même manière par l’utilisateur devant la table des matières de sa boîte mail qui lui indique les émetteurs des messages comme leurs sujets. Si l’on applique strictement ce principe, cela revient à considérer les conversations qui comportent plus de deux fois plus de messages que la moyenne.

Nous avons cependant apporté un correctif à cette définition. En effet, dans les deux listes étudiées, on constate une certaine redondance dans les auteurs des conversations comme l’indiquent ces tableaux : les locuteurs interviennent souvent plusieurs fois dans la même conversation, ce qui a pour conséquence que par exemple, une conversation comprenant 6 messages comporte en moyenne 3 à 4 locuteurs différents.

NAIS.	Nb locuteurs différents /conv.	Nb messages /conv.	Redondance
2001	3,0	4,5	1,5
2002	3,9	6,9	1,8
2003	3,6	6,2	1,7
2004	3,8	6,9	1,8
2005	3,7	6,4	1,7
2006	4,0	8,0	2,0
2007	4,5	8,0	1,8

TDAH	Nb Locuteurs différents /conv.	Nb messages /conv.	Redondance
2004	2,8	4,2	1,5
2005	2,9	4,1	1,4
2006	2,6	3,6	1,4
2007	2,5	3,3	1,3
2008	2,8	4,1	1,5
2009	3,2	4,8	1,5

Tableau 5 : moyennes du nombre de locuteurs et du nombre de messages par conversation

L’indicateur de **redondance** est une moyenne et à ce titre, il agrège des configurations variées : certaines conversations sont extrêmement redondantes, alors que d’autres sont composées de contributions uniques de chacun des locuteurs. Si l’on prend bien en considération l’utilisateur-lecteur en définissant les conversations longues par un seuil en nombre de messages, il nous semble qu’il faut aussi tenir compte de l’utilisateur-

contributeur qui, par l'intérêt qu'il manifeste pour un sujet, produit une autre mesure de la visibilité. C'est la raison pour laquelle nous avons choisi de prendre en compte les conversations longues et/ou collectives, c'est-à-dire les conversations qui admettent soit un nombre de messages double de la moyenne, soit un nombre de locuteurs différents double de la moyenne.

Première remarque : l'accueil des nouveaux arrivants est pris en charge par les locuteurs dominants

Les travaux que nous avons évoqués précédemment sur ces questions (Burke, M. *et al.* 2007; Joyce & Kraut 2006) reposent sur deux hypothèses principales : (i) le fait qu'un nouvel arrivant dans un groupe suscite ou non des réponses à sa première intervention a un effet déterminant sur son retour au sein du groupe ; (ii) la capacité à recevoir une réponse est liée au contenu formel du message. Comme nous l'avons vu, ces hypothèses ne sont que partiellement validées par les analyses statistiques et les expérimentations menées par les chercheurs.

Nous allons prendre ici un parti inversé par rapport à ces travaux et nous intéresser davantage à la réception des premiers messages qu'à leur conception. À l'appui de ce choix, signalons que les taux de réponse observés sur nos deux listes sont plus importants que ceux indiqués dans la littérature (57% chez Burke, M. *et al.* (2007); 74% chez Arguello *et al.* (2006) : ils atteignent 91% sur la liste Naissance et 80% sur la liste TDAH.³² Du coup, on peut légitimement se demander si le seul message conditionne sa réponse.

Par ailleurs, les nouveaux arrivants sur les listes Naissance et TDAH ont plus souvent tendance, en guise d'introduction, à envoyer un message de réponse dans une conversation existante plutôt que de démarrer une nouvelle conversation : cela se produit dans 25% des cas sur la liste TDAH et 44% sur la liste Naissance. Cette dernière fonctionne de manière spécifique puisqu'elle a mis en place après quelque temps d'existence une procédure d'inscription : la personne qui veut s'abonner à la liste envoie au propriétaire une présentation que celui-ci fait suivre aux autres membres avant validation de la demande. Les « premiers » messages sont donc la plupart du temps précédés d'une forme d'introduction, ce qui permet peut-être au novice d'entrer plus directement dans la discussion et expliquerait l'importance des premiers messages de type « réponse ».

Quoi qu'il en soit, on observe sur les deux listes une distribution atypique des premières réponses (premier message qui est envoyé en réponse au message du nouvel arrivant.) aux premiers messages :

³² Est-ce lié à la différence entre groupes Usenet et listes de discussion ? Ou à la méthode permettant de déterminer les réponses ? Ou encore simplement aux spécificités de chaque groupe ? Nos chiffres sont sous estimés (erreurs, réponse avec changement de sujet etc.)

NAIS.	Réponses dominants	Messages Dom. Liste	Ecart: convivialité
2002	81%	61%	20%
2003	92%	73%	19%
2004	78%	71%	7%
2005	84%	54%	30%
2006	82%	57%	25%

TDAH	Réponses dominants	Messages Dom. Liste	Ecart: convivialité
2005	75%	71%	4%
2006	75%	68%	7%
2007	80%	70%	10%
2008	84%	71%	13%
2009	76%	67%	9%

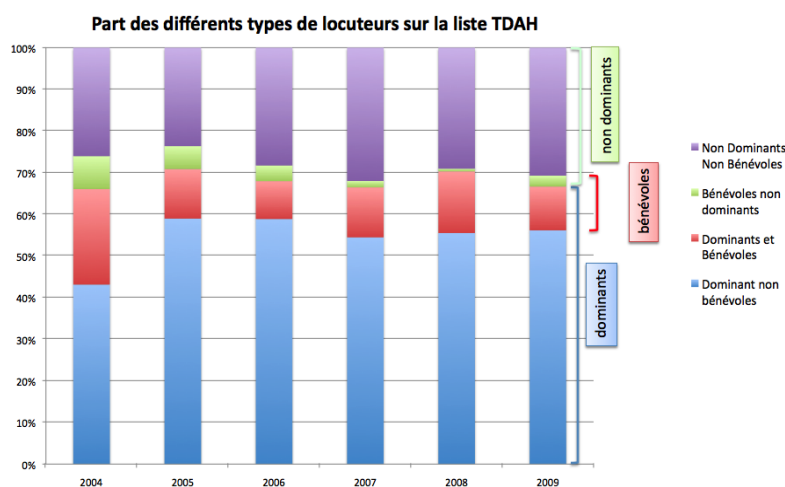
Tableau 6 : part des dominants dans les premières réponses aux arrivants et dans la liste

Nous avons calculé un indicateur de **convivialité** du groupe dominant comme l'écart entre leur part dans les premières réponses aux messages des nouveaux arrivants et leur part dans les messages. On constate que, sur les deux listes, les locuteurs dominants font preuve d'une convivialité élevée. Cette convivialité est cependant moins marquée sur la liste TDAH. Ceci étant, si l'on considère l'ensemble formé par les locuteurs dominants et les bénévoles non dominants³³ présents sur la liste TDAH, on obtient une figure plus proche de celle de la liste Naissance.

Année	Réponses des dominants et bénévoles	Messages des dominants et bénévoles dans la liste	Ecart: convivialité
2005	84%	76%	8%
2006	84%	72%	12%
2007	86%	68%	18%
2008	90%	71%	19%
2009	80%	69%	10%

Tableau 7 : part des dominants et bénévoles aux premières réponses aux arrivants et dans la liste

³³ Un certain nombre de locuteurs signent en indiquant leur statut de bénévoles et a minima leur département ou leur région de sorte à faciliter les prises de contact éventuelles : notre définition des bénévoles est uniquement basée sur cette déclaration. Il est possible que d'autres bénévoles soient présents sur la liste sans s'être déclarés en tant que tels. Certains bénévoles appartiennent aussi au groupe des dominants, notamment la présidente de l'association.



Graphique 13 : part des différents types de locuteurs sur la liste TDAH

Une analyse plus précise des locuteurs impliqués dans ces réponses permet d'affiner la description : dans les deux listes, environ un locuteur sur 10 a produit au moins une première réponse à un nouvel arrivant ; cependant sur la liste Naissance, le fondateur de la liste représente à lui tout seul de 43% de ces réponses soit plus de la moitié des réponses faites par les locuteurs dominants, alors que sur la liste TDAH, une grande partie du « travail » est réparti entre 9 locuteurs différents (dont 3 bénévoles) qui totalisent près de 60% des réponses.

Sans qu'il n'y ait de règle explicite, on constate qu'une grande attention est portée collectivement par les locuteurs dominants et les bénévoles à l'accueil des nouveaux arrivants. Parmi ces locuteurs, certains occupent des positions remarquables : le propriétaire de la liste sur Naissance ; les bénévoles de l'association sur la liste TDAH.

Tout se passe donc comme si, pour les locuteurs dominants, la liste constituait un véritable espace collectif dont ils se sentent responsables. ils apportent ainsi une réponse à la question de la « communauté » : par leur comportement, ils tendent à performer cette communauté.

Notons par ailleurs que ce comportement semble avoir des effets ; de ce point de vue, nos constatations rejoignent celles de Burke et ses collègues : le fait de recevoir une réponse à un premier message est lié à la probabilité de poster à nouveau sur la liste.

NAISSANCE	ont posté à nouveau	n'ont pas posté à nouveau	TDAH	ont posté à nouveau	n'ont pas posté à nouveau
ont reçu une réponse	99%	1%	ont reçu une réponse	62%	38%
n'ont pas reçu une réponse	86%	14%	n'ont pas reçu de réponse	50%	50%
premier message = réponse	92%	8%	premier message = réponse	64%	36%

Tableau 8 : comportement des nouveaux arrivants en fonction de leur premier message

Une fois encore, constatons à quel point il est compliqué d'aller plus avant dans l'interprétation : il est possible qu'un message ne reçoive pas de réponse parce qu'il est inapproprié par rapport à l'objet de la liste ; du coup, la probabilité que la personne ayant posté s'investisse dans la liste est faible qu'elle reçoive ou non une réponse.

Deuxième remarque : le travail de « réponse » aux initiatives des participants (lancement de nouveaux sujets) est inégalement réparti selon les listes.

De manière plus générale, le taux de messages sans réponse peut être considéré comme un indicateur de la plus ou moins grande interactivité du groupe et donc de son caractère proprement collectif. Il est normal qu'un certain nombre de messages restent isolés, car ils peuvent consister en une information qui n'appelle pas de réponse. Sur la liste TDAH, 30% des conversations lancées n'ont pas reçu de réponse, alors que ce

chiffre est de 22% pour la liste Naissance, proportions de toute façon inférieures aux données que l'on trouve dans la littérature.

À l'image de ce que nous avons constaté pour les messages des nouveaux arrivants, on peut se demander qui amorce le dialogue, ou plus exactement le multilogue tel que le rebaptise justement Ekeblad (2001). Nous avons calculé un indicateur de **réactivité** qui est défini comme le rapport entre la part des dominants dans les seconds messages divisée par la part des dominants dans la totalité des messages. Les deux listes présentent de ce point de vue des profils contrastées : alors que sur la liste Naissance, la répartition de ces premiers répondants est strictement à l'image de la répartition des messages sur la liste, sur la liste TDAH, on retrouve une certaine surreprésentation des locuteurs dominants et des bénévoles.

NAIS.	seconds messages des dominants	messages dominants	réactivité
2001	68%	68%	1
2002	63%	61%	1,04
2003	76%	73%	1,04
2004	74%	71%	1,04
2005	56%	54%	1,04
2006	59%	57%	1,04
2007	67%	65%	1,03

TDAH	seconds messages des dominants et bénévoles	messages des dominants et bénévoles	réactivité
2004	82%	74%	1,11
2005	84%	76%	1,11
2006	80%	72%	1,11
2007	80%	68%	1,18
2008	79%	71%	1,11
2009	77%	69%	1,12

Tableau 9 : part des dominants dans les seconds messages

Pourtant, malgré cet effort particulier des dominants sur la liste TDAH, la part des messages sans réponse est plus importante sur cette liste.

Nous voyons ici se dessiner deux modes de fonctionnement différents :

- Sur la liste Naissance, passée la phase d'introduction, les comportements des petits locuteurs apparaissent homologues de ceux des locuteurs dominants : ils s'autorisent à/ trouvent la motivation pour engager la conversation sur la base d'un premier message envoyé par n'importe lequel des locuteurs. Le groupe apparaît donc « soudé » au sens où une multitude de liens sont construits entre les différents locuteurs.
- Sur la liste TDAH, l'asymétrie entre les dominants et les bénévoles de l'association d'un côté et les autres locuteurs de l'autre paraît plus forte. L'initiative de la réponse aux premiers messages de nouvelles conversations est davantage entre les mains des dominants et des bénévoles. La plus faible réactivité des petits locuteurs est sans doute à mettre en relation avec le taux plus élevé de premiers messages sans réponse.

Troisième remarque : les petits locuteurs sont proportionnellement plus impliqués dans le lancement de nouveaux sujets

Où la dynamique d'échanges s'alimente-t-elle? Quelles sont les initiatives qui permettent d'enclencher des discussions plus longues et/ou plus collectives? Les différentes catégories de locuteurs ont-elles des rôles particuliers de ce point de vue ?

Nous avons calculé la part des locuteurs dominants dans le lancement de sujet, et rapporté cette part à la part des messages, construisant ainsi un **indicateur d'animation**. On constate une asymétrie forte dans le lancement des sujets sur la liste TDAH, puisqu'en proportion les locuteurs dominants lancent de 11 à 24% moins de sujets qu'ils n'envoient de messages, alors qu'en ce concerne la liste Naissance, la différence est beaucoup moins marquée.

Naissance	% messages dominants	% lancements sujets	animation
2001	68%	60%	0,88
2002	61%	55%	0,90
2003	73%	71%	0,97
2004	71%	63%	0,89
2005	54%	55%	1,02
2006	57%	51%	0,89
2007	65%	66%	1,02

TDAH	% messages dominants	% lancements sujets	animation
2004	66%	49%	0,74
2005	71%	58%	0,82
2006	68%	53%	0,78
2007	70%	59%	0,84
2008	71%	47%	0,66
2009	67%	51%	0,76

Tableau 10 : part des dominants dans les lancements des sujets

On pourrait penser que le groupe des locuteurs dominants agit globalement comme un « répondeur » (Turner, T. C. *et al.* 2005; Viegas, F. B. & Smith 2004; Welser, H. *et al.* 2007). Cependant cette notion suppose, dans les analyses dans lesquelles elle a été mise en avant, un pendant, le « questionneur » et s'inscrit dans un scénario dans lequel l'échange est pensé sous la forme d'une alternance Questions / Réponses, souvent soutendue par une distribution asymétrique de l'expertise. Or, comme le suggère le cas des seconds messages sur la liste Naissance et comme nous le verrons dans la suite, les petits locuteurs sont aussi des répondeurs. De plus, la part des sujets lancés par les locuteurs dominants est très importante et disproportionnée par rapport à la taille du groupe. Toutes ces raisons nous poussent à conserver un langage de description prudent n'impliquant pas une interprétation qui ne serait pas rigoureusement établie : les données réunies ici suggèrent que les petits locuteurs sont un moteur non négligeable dans la dynamique de la liste. D'autres éléments nous permettent d'aller un peu plus loin dans l'interprétation.

Quatrième remarque : les petits locuteurs participent davantage à la dynamique collective par le lancement de nouveaux sujets / les locuteurs dominants sont plus impliqués dans l'animation de la discussion

Si l'asymétrie entre lancement et participation est moins marquée en général pour la liste Naissance, elle apparaît de manière contrastée dès lors que l'on s'intéresse aux conversations longues et/ ou collectives.

Signalons que ces conversations représentent 17% des conversations sur la liste Naissance, et 16% sur la liste TDAH, pour 60% des messages sur la liste Naissance et 50% sur la liste TDAH .

Cela signifie qu'en pratique, lorsque l'abonné parcourt sa boîte mail, il voit une bonne moitié de ces messages qui affichent deux ou trois sujets, et d'autres messages plus disparates dont le sujet disparaît rapidement, ensevelis par de nouveaux messages et de nouveaux sujets, ce qui produit un effet de fixation thématique et rend visible d'une certaine manière le groupe comme collectif de débat. Autrement, cette concentration participe aussi à la constitution de la « communauté », qui n'est autre que le nom donné aux effets produits par cette accumulation/ combinaison d'action individuelles, mais qui en même temps, produit en retour des effets sur les individus.

naissance	% <i>messages dominants</i>	% sujets longs	% sujets courts	tdah	% <i>messages dominants</i>	% sujets longs	% sujets courts
2001	68%	53%	60%	2004	66%	51%	48%
2002	61%	46%	56%	2005	71%	68%	56%
2003	73%	65%	71%	2006	68%	48%	54%
2004	71%	51%	65%	2007	70%	50%	60%
2005	54%	34%	59%	2008	71%	54%	46%
2006	57%	43%	52%	2009	67%	65%	49%
2007	65%	58%	67%				

Tableau 11 : part des dominants dans le lancement des sujets longs/ courts

La liste Naissance se caractérise par l'importance relative des petits locuteurs comme « lanceurs de sujets longs »³⁴ – alors que le lancement de sujets courts suit d'assez près la répartition des messages. En revanche, si, sur la liste TDAH, les petits locuteurs apparaissent comme des lanceurs proportionnellement importants, la répartition de cette activité entre sujets longs et sujets courts est irrégulière.

³⁴ Bien évidemment, le lancement de sujets longs n'est pas une activité en soi : ce sont les autres qui, par leur action, qualifient rétrospectivement le lancement. En même temps, on ne peut supposer non plus que la longueur du fil soit totalement sans rapport avec le message. Attribuer à une entité collective constituée par le chercheur « les petits locuteurs » la qualité de « lanceurs de sujets longs » est donc une opération criticable et qui ne doit être prise dans un sens littéral, c'est-à-dire comme l'expression d'une disposition individuelle.

Les sujets lancés par les petits locuteurs semblent donc nourrir fortement la dynamique d'échange, avec des modalités différentes : sur la liste Naissance, cet effet se porte préférentiellement sur les conversations longues, alors que sur la liste TDAH, il concerne aussi bien les conversations longues que les conversations courtes.

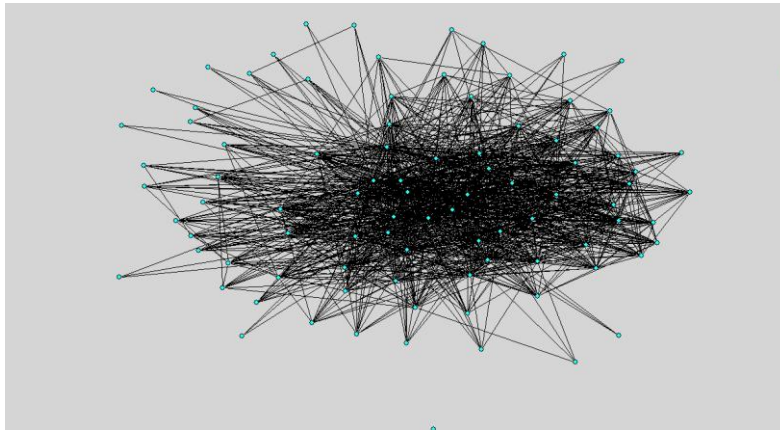
Une autre manière de décrire les inégalités constatées consisterait à dire que les locuteurs dominants sont davantage investis dans la participation à des conversations existantes que dans le lancement de nouveaux sujets. Ce qui est juste aussi.

Cette double « lecture » possible d'une même irrégularité statistique attire notre attention sur l'impossibilité qu'il y a à séparer description et interprétation : le matériau statistique n'est d'ailleurs pas plus « objectif » en cela que n'importe quel autre type de matériau, car tout énoncé à propos de ce matériau contient de fait déjà une définition des agents et une imputation causale.

Cinquième remarque : il existe une certaine segmentation des locuteurs dans les conversations collectives

Le tableau que nous avons brossé précédemment n'est pas dénué d'ambiguïtés : d'un côté, il montre que les locuteurs dominants déploient un certain nombre d'efforts pour faire exister le groupe en tant que tel ; de l'autre, il permet de voir aussi que les comportements des uns et des autres se différencient sur certains aspects par d'autres éléments que le volume de messages postés. D'où la question que l'on peut se poser de l'existence d'une segmentation des locuteurs : jusqu'à quel point le multilogue fait-il exister de fait un espace plus ou moins continu ?

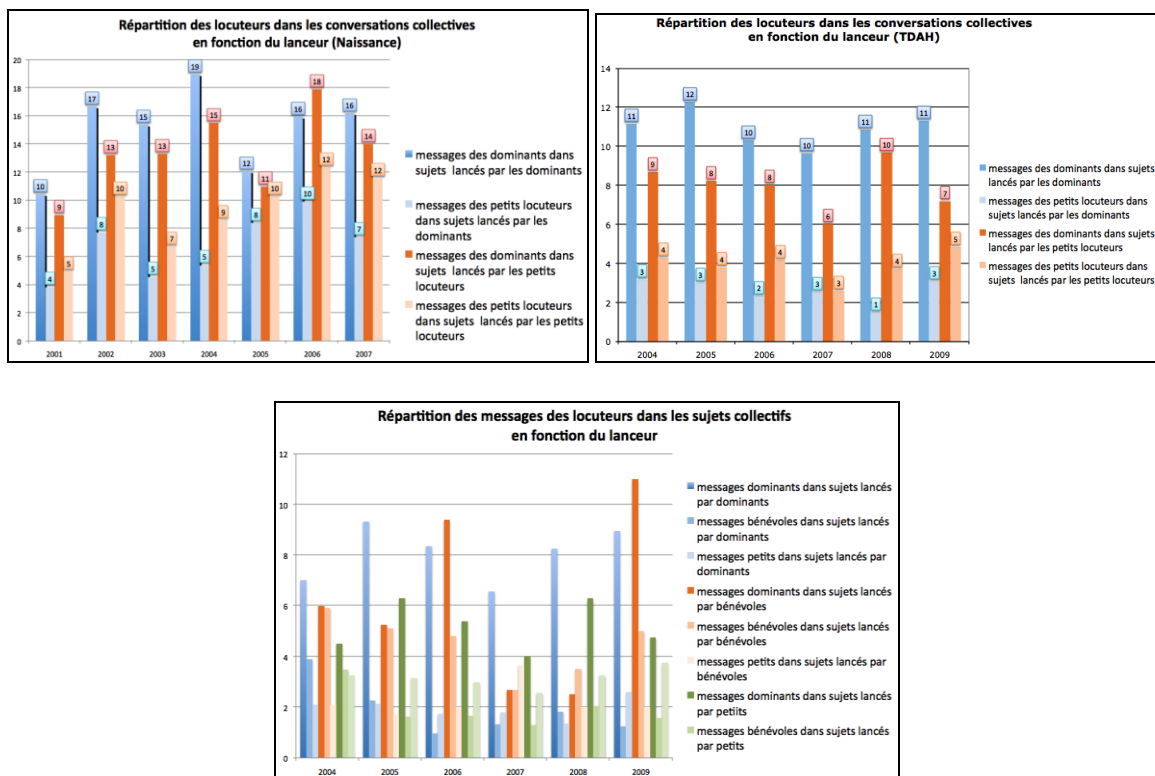
On pourrait penser au vu des travaux présentés plus haut qu'une analyse de réseaux constitue une approche possible pour traiter de cette question : en considérant que la participation à une conversation révèle des intérêts partagés, on pourrait tracer les clusters qui émergeraient alors de l'analyse de ces liens. Disons-le tout de suite : sauf à se limiter à une période de temps très courte, à l'instar des auteurs que nous avons cités – et alors se pose la question du caractère significatif des relations mises en évidence – cette piste ne mène nulle part. La raison en est que les listes sont, en raison de l'extrême activité dont elles sont le siège et de l'inégalité des contributions, saturées de liens. Ce qui, accessoirement, ouvre une autre perspective sur la question de la communauté : l'impossibilité de faire passer des démarcations très tranchées entre différents groupes de participants est une manière de décrire l'existence d'un « commun » autour duquel se définit le groupe. Les tentatives que nous avons faites de tracer les réseaux créés par les échanges débouchent sur des représentations en pelote très emmêlée du type de celle-ci :



Graphique 14 : réseau des liens entre les locuteurs de la liste Naissance en 2001, obtenu à partir d'une analyse de cooccurrence dans les conversations

Ce type de représentation ne fait au fond que redire sous une autre forme ce que nous avons montré jusqu'à présent, à savoir qu'il y a un sous-groupe fort actif – ici au centre – et d'autres locuteurs moins actifs, à la périphérie, les liens étant assez directement liés à l'activité. Les analystes des réseaux ont été peut-être aveuglés par la théorie de l'attachement préférentiel au point d'en oublier un fait de base, à savoir d'inégale contribution des membres d'une liste : de ce point de vue, la répartition en loi de puissance ne ferait que rendre compte de la variabilité des individus du point de vue de leurs goûts ou intérêts, de même que la loi normale rend compte de la variabilité des individus du point de vue de la taille par exemple. Peut-être est-il possible que l'on puisse ensuite modéliser cette répartition par un modèle d'attachement préférentiel – mais à quoi ? Ceci étant, la répartition des liens entre les différents locuteurs nous semble en grande partie déterminée par la répartition des contributions : en clair, si l'on envoie peu de messages sur une liste, on sera a priori lié à peu de contributeurs, quelle que soit la manière dont on calcule les liens (liens de message à message par la fonction *Reply* ou liens par l'appartenance à une même conversation), l'inverse étant tout aussi prévisible.

Nous avons suivi une stratégie plus simple qui consiste à regarder comment se répartissent, dans les conversations longues et/ou collectives, les contributions des différents locuteurs en fonction de l'identité du lanceur. Dans les graphiques suivants, on a représenté le nombre moyen de messages envoyés par chaque catégorie de locuteur dans les conversations lancées par les locuteurs dominants et dans les conversations lancées par les petits locuteurs :



Graphiques 15 : Répartition des locuteurs dans les conversations collectives en fonction du lanceur
 Deux remarques principales peuvent être faites :

- Dans les deux listes, on constate une constante inégalité de répartition des locuteurs dans les conversations qui dépend du type de lanceur ; bien sûr, compte tenu de la place occupée par les locuteurs dominants dans les messages, il est normal qu'il y ait une présence plus importante de ces locuteurs dans toutes les conversations. Cependant, on observe que les conversations lancées par les locuteurs dominants attirent davantage de locuteurs dominants qu'elles ne le devraient en théorie, et les conversations lancées par les petits locuteurs recrutent plus de petits locuteurs qu'attendu. Cette inégalité est sensiblement plus marquée dans la liste TDAH que dans la liste Naissance : les locuteurs dominants de Naissance postent environ 12% de messages en plus qu'attendu dans les conversations lancées par leurs pairs, alors que ce pourcentage est de l'ordre de 20% pour la liste TDAH, ce qui conduit à des écarts deux fois plus importants entre les conversations lancées par les dominants et celles lancées par les petits locuteurs. En ce qui concerne la liste TDAH, cet effet de « préférence » est aussi visible lorsque l'on prend en considération la catégorie des bénévoles : bien que les conversations lancées par les bénévoles soient plus courtes (à l'exception de 2009) que celles lancées par les dominants, le nombre moyen de messages envoyés par les bénévoles dans les premières est toujours nettement supérieur au nombre moyen de messages envoyés dans les secondes. En revanche, la participation des autres locuteurs est plus variable, mais ceci peut être mis en relation avec le fait que la part des bénévoles dans l'activité de la liste étant plus faible (graphique 13), le nombre de sujets longs lancés est faible (51 sur les 6 années

et 2 ou 3 par an seulement depuis 2007, contre 385 pour les dominants et 346 pour les petits locuteurs), ce qui limite l'effet lissant des grands nombres.

- La longueur moyenne des conversations qu'elles soient lancées par les locuteurs dominants ou les petits locuteurs est identique sur la liste Naissance, à l'exception de l'année 2006, alors que, sur la liste TDAH, les conversations lancées par les locuteurs dominants sont presque toujours plus longues que les conversations lancées par les petits locuteurs et les bénévoles.

Comment interpréter ces différences ? Deux types d'explication sont possibles : la première, de type réseau social, fera l'hypothèse que les individus sont attachés entre eux de manière différentielle, que l'implication quotidienne dans des échanges soutenus avec certaines personnes créent de liens entre ces personnes, d'où la plus grande réactivité des locuteurs dominants à l'égard de leurs pairs ; la seconde sera axée sur les contenus et supposera que les deux groupes ont des intérêts différenciés ce qui explique leur tropisme vers des sujets lancés par leurs pairs. Ces deux types d'explication n'ont en fait pas à être opposées : quiconque a fréquenté des listes de discussion ou même simplement utilisé son mail sait bien qu'un certain nombre de mails n'apportent rien du point de vue du contenu de la discussion, mais sont là en tant que témoins / performateurs de certaines relations (amitié, estime, complicité, inimitié etc.) entre les personnes. En même temps, on ne peut exclure que la nature même des discussions ait à voir avec ces différences. Nous allons nous engager dans la suite dans une analyse thématique afin d'être en mesure d'apporter des éléments de réponse à cette question.

Conclusion

La définition d'un certain nombre d'indicateurs centrés sur l'identification des contributions de chaque type de locuteurs à la dynamique d'échanges nous a permis d'avancer dans la spécification du fonctionnement des listes.

Indicateurs considérés :

- Convivialité et Réactivité : Part des différentes catégories de locuteurs dans les premières réponses aux nouveaux arrivants et les premières réponses aux nouveaux sujets
- Animation : part des différentes catégories de locuteurs dans le lancement de nouveaux sujets, dans le lancement de sujets qui donnent lieu à des conversations courtes/longues
- Part des différentes catégories de locuteurs dans les conversations longues en fonction du lanceur

La liste Naissance apparaît comme un groupe relativement homogène : passée la phase introductive, dans laquelle les locuteurs dominants et notamment le propriétaire de la liste jouent un rôle important d'intégration, les comportements des uns et des autres semblent assez similaires, à l'intensité de participation près. Cependant, les conversations longues et collectives prennent appui plus souvent qu'attendu sur les

messages envoyés par les petits locuteurs, et on note une légère propension dans chacun des deux groupes à privilégier les sujets lancés par leurs pairs.

La liste TDAH présente une physionomie caractérisée par un contraste marqué entre locuteurs dominants et petits locuteurs. Non seulement les locuteurs dominants « accueillent » les nouveaux arrivants davantage que les petits locuteurs, mais ils « animent » les échanges en répondant assez systématiquement aux messages qui lancent de nouveaux sujets. Les conversations lancées par les petits locuteurs sont en proportion sensiblement plus importantes en volume qu'attendu au vu de la répartition des messages, sans être à l'inverse plus longues que celles lancées par les locuteurs dominants. En revanche, le tropisme de chaque type de locuteurs en faveur des conversations lancées par leurs pairs est particulièrement sensible, renforçant l'idée d'un différentiel entre ces deux types de locuteurs.

La dynamique des débats

Nous avons conclu la partie précédente sur une question, celle de la relation possible entre la préférence marquée par chaque catégorie de locuteurs pour des conversations lancées par leurs pairs et une orientation spécifique que pourraient avoir ces conversations. Comment décrire et qualifier le contenu des échanges ? Nous avons vu dans le chapitre précédent que quelques auteurs avaient essayé d'apporter des réponses à cette question. Aucune n'est totalement satisfaisante à l'échelle d'une liste en raison des caractéristiques du matériau : confus, hétérogène, redondant, volumineux. Par ailleurs, il pose des problèmes de délimitation. Rien à voir par exemple avec l'univers des blogs, dans lequel chaque article forme une entité en tant que tel, est généralement écrit dans un style en accord avec les règles usuels de grammaire et d'orthographe, déroule un argument ou une description du début à la fin et admet un auteur identifié. Un message dans un groupe de discussion électronique est un élément qui ne prend sens que dans un macro-texte dont on ne peut définir aisément les limites ni les auteurs. Les « conversations » électroniques n'ont pas non plus la linéarité des conversations en coprésence. Comme Conein & Latapy (2008), Ekeblad (2001) et Barcellini (2005) l'ont montré et même traduit visuellement, une conversation se développe en fait plutôt en différentes grappes qui peuvent se mélanger à nouveau à tout moment. Aucune méthode ne semble capable de capturer complètement cette complexité.

À l'échelle d'une liste, c'est-à-dire d'un ensemble comprenant plusieurs dizaines de milliers de messages, une première stratégie consiste à se saisir le dispositif même qui a été mis en place pour opérer une simplification et donner une idée du contenu : le sujet. Le sujet joue un rôle un peu du même ordre que le titre d'un article, en offrant au lecteur une prise synthétique sur un contenu. Bien sûr, compte tenu des méandres que suivent les conversations, les sujets ne peuvent prétendre à représenter l'ensemble de ce qui s'échange, mais au moins donnent-ils une idée des « projets » qu'ont en tête ceux qui lancent de nouvelles conversations. In fine, on peut sans doute les considérer comme rendant compte des thèmes qui intéressent la liste. Mais les seuls sujets représentent

une quantité de données importante, car, sur la période étudiée, on compte près de 5000 sujets pour la liste TDAH et de 11 000 pour la liste Naissance.

Nous avons donc décidé de procéder de la manière suivante : nous avons établi un répertoire des mots figurant dans l'ensemble de ces sujets, calculé les occurrences de ces mots, puis construit un certain nombre de catégories/ thèmes à partir des mots extraits, thèmes dans lesquels nous avons en même temps placé les mots en commençant par les plus fréquents.

Dans un second temps, les thèmes ont été associés aux sujets, par l'intermédiaire des mots qu'ils contiennent ; on a recherché les contextes d'utilisation pour les mots ambigus ou polysémiques et seuls ont été catégorisés ceux pour lesquels un seul contexte bien déterminé était retrouvé. Les thèmes ont été pondérés : chaque sujet ne peut compter que pour un thème, donc s'il était lié à 2 thèmes différents, un poids de 0,5 a été affecté à chacun des thèmes.

L'index comportait pour la liste TDAH 1467 mots (3155 pour Naissance), 538 (781 pour Naissance) ont été affectés à 16 (19) thèmes, ce qui représentait 70% (65%) des mots non uniques se trouvant dans les sujets. Au total 75% des sujets de TDAH et 76% des sujets de Naissance ont été thématiques.

Les catégories ainsi construites sont en partie communes aux deux listes ; on trouve ainsi les catégories suivantes :

- besoin : renvoie à l'expression d'une demande de la part de la personne
- relation : comprend les termes qui comporte une forme d'adresse que ce soit par l'intermédiaire d'un prénom ou d'un nom, d'une salutation...
- documentation : renvoie à des formes élaborées de publications
- information : renvoie aux media de masse
- état mental : comprend les termes qui qualifient l'état psychologique ou émotionnel de la personne
- expérience : renvoie à l'expérience personnelle, notamment par l'emploi de la première personne du singulier
- professionnel : désignation des professionnels de santé
- famille : qualifie les différents membres de l'entourage

Thèmes des sujets de TDAH, suivis des premiers mots associés	
association	abonnement, adhésion, afdepa, aftcc, AG, association, benevole, echange, fichier, formulaire, forum, genedys, hypers, hypersupers, inscription, liste, membre...
besoin	action, adresse, aide, alternative, appel, assistance, avis, besoin, cause, comprehension, conseil, contact, coordonnees, demande, eclairer, explication, interrogation, probleme... sos, svp...
relation	accueil, adieu, bonjour, bonsoir, desabonnement, desinscription, desole, message, pardon, « prenom », presentation, remerciement, retour, salutation, secours, soutien, tchat...
documentation	article, document, documentation, livre, livret, pdf, publication, recherches, these, traduction
information	antenne, colloque, communication, conference, congres, connaissance, Delarue, doctissimo, emission, figaro, film, information, internet, journal, lien, magazine, majordomo...
etat mental	angoisse, anxiete, apprehension, blessure, bonheur, calme, cauchemar, colere, concerne, confiance, consterne, courage, craquer, deboussole, deception, decouragement, defoulement, degout, depression, deroute, desarroi, desespero...
experience	1ere pers sing, anecdote, experience, histoire, parcours, sondage, temoignage, vecu
manifestation	absence, agitation, agressivite, attention, bagarre, comportement, concentration, conduite, demon, diable, faineantise, hyperemotivite, impulsivite, incontrolable, insolence, insomnie, insupportable, intenable, irritabilite, kleptomanie, monstre, mouvement, obeissance, oubli...
trouble	abandon, adhd, alcoolisme, anemie, aphasique, asperger, associe, autisme, bipolaire, borderline, cannabis, caryotype, cerveau, choree, chromosome, circadin, coeliaque, cognition, comorbidite, comportementalisme, debilite, deficiencie, deficit, dys, dyscalculie...
diagnostic	annonce, bilan, conners, connors, consultation, depistage, diagnostic, rdv, reconnaissance, reperage, suspicion, test
medicament	atarax, atomoxetine, automedication, comprime, concerta, delivrance, depkine, dopage, dosage, drogues, medicament, medication, methadone, mg, micropakine, millepertuis, modafinil, nicotine, pharmacien, pharmacogenetique, posologie, prescripteur...
traitement	accompagnement, acupuncture, adaptation, aménagement, antidouleurs, approches, Barkley, biokinergie, caseine, cmp, cmpp, cure, debre, ergotherapie, erikson, fer, gestion, gluten, guidance, homeopathie, hospitalisation, hypnose, kinesiologie, Lecendreux, macrokinesie...
professionnel	cardiologie, enterologue, equipe, ergotherapeute, generaliste, graphotherapeute, hopital, infirmiere, medecin, neurologue, neuropediatre, neuropsy, orthophoniste, orthoptiste, pediatre, pedopsychiatre, posturologue, professionnel, psychiatre, psychiatrique, psychologue
institution	administration, allocation, amendement, circulaire, commission, deutes, droits, gendarmerie, institution, jugement, jurisprudence, justice, loi, mdph, mediateur, ministere, oms, procedure...
ecole	apprentissage, auxiliaire, avs, bac, bulletin, cartable, classe, cliss, college, collegien, cours, cp, descolarisation, devoirs, ecole, educateur, eduscol...
famille	adolescence, adoption, colonie, education, enfant, famille, femmes, fils, fratrie, frere, garcon, gosse, hyperpapas, maman, mari, mere, papa, parent, pere, soeur

Les catégories spécifiques à cette liste concernent l'association liée à la liste, la qualification des comportements associés au TDAH (manifestation), les troubles qui lui sont associés (trouble), le processus de diagnostic, les médicaments spécifiques au TDAH et les controverses associées (médicament), les autres formes de traitement du trouble (traitement), l'encadrement juridique et l'environnement institutionnel (institution), l'école.

Thèmes des sujets de Naissance, suivis des premiers mots associés	
liste	2eme pers plur, abonne, abonnement, adhesion, archives, deconnexion, depart, desabonnement, desinscription, inscription, liste de discussion, message, moderation, presentation, yahoo
besoin	aide, avis, besoin, conseil, demande, explication, interrogation, question, questionnement, recherche, renseignement, reponse, secours, sos,
relation	2eme pers sing, adieu, bienvenue, bise, bonjour, bonsoir, bravo, chat, coucou, debat, desaccord, digression, discussion, echange, email, humour, liste, nom, reflexion, remerciement
documentation	article, bibliographie, chapitre, colloque, conference, document, documentaire, editeur, edition, enquete, etude, fichier, lecture, litterature, livre, ouvrage, parution, publication, reference, science, seminaire, statistique, symposium, texte, these, traduction, txt
information	arte, atelier, chiffres, citation, doctissimo, emission, film, forum, fraternet, http, infobebes, information, internet, interview, journal, journaliste, lactaliste, lien, livret, magazine, magicmaman, medias, mediatisation, net, news, page, portail, presse, référencement, reportage, site, television, video, web, webmaster, wikipedia, www
etat mental	acceptation, amour, angoisse, attente, audace, bienveillance, blues, bonheur, cauchemar, colere, confiance, confusion, conscience, courage, crainte, craque, croyance, culpabilite, deception, degout, delire, depression, desarroi, desespoir, desillusion, desir, desolation, deuil, doute, emotion, enervement, enfer, envie, fantasme, fatigue, foi, folie, gentillesse, harcèlement, ...
experience	1ere pers sing, anecdote, confiance, deces, drame, echec, ego, epilogue, episode, experience, fiasco, histoire, mort, projet, questionnaire, recit, sensation, solitude, sondage, temoignage, temoin, vecu
grossesse	avortement, conception, enceinte, fecondation, grossesse, haptonomie, prenatal, preparation, reeducation, trimestre
accouchement	accouchement, cordon, delivrance, enfantement, expulsion, naissance, postpartum, relevailles
corps	abdominal, accroupi, acne, alimentation, anal, aphte, appetit, articulation, bedaine, boire, bouchon, bronche, brulure, calcium, cerveau, cheveux, cicatrice, circoncision, colostrum, constipation, contraction, corps, cou, crane, cul, decalottage, déchirure, defecation, degueuler...
enfant	adolescence, allaitement, bambin, bebe, biberon, cododo, complement, dent, diversification, echarpe, ecole, education, enfant, fessée, homeschooling, infantile, instruction, jumeaux, lactation, lait, lange, nourrice, nourrisson, portage, precocite, sevrage, succion, tete, tetine
famille	aine, compagne, compagnon, conjoint, couple, entourage, famille, fille, foyer, frere, mari, maternage, mere, parent, parentage, paternite, pere, proches, relation
medecine	abcès, allergie, analyse, anemie, anomalie, antibiotique, anticorps, artere, arthrose, asthme, autisme, bacterie, bronchite, calcification, cancer, candida, chirurgie, colique, congenital, conjonctivite, consultation, convulsion, courbe, crohn, cystite, depistage, diabete, diagnostic...
obstetrique	accreta, agglutinines, albicans, albumine, amniocentese, amniotique, analgesie, anesthesie, aspiration, bassin, bilirubine, cerclage, cervical, cesarienne, cholestase, circulaire, clampage, cytotec, declenchement, decollement, desobstruction, echographie, eclampsie, effacement, episiotomie, examen, expression, fertilité, foetoscope, foetus, forceps, gestation, ...
professionnel	acupuncture, docteur, generaliste, gynecologue, gynobs, kinesitherapie, medecin, obsteticien, osteopathe, osteopathie, pediatre, pediatrie, praticien, professionnel, psychanalyse, psychiatrie, puericulture, sage-femme, sfmp, soignant, sophrologie
systeme de soins	ambulatoire, clinique, hopital, maternite, minguettes, usager
economie	assurance, codification, commerce, consommateur, consommation, cout, depassement, economie, fric, mutuelle, nomenclature, paiement, prestation, prix, remboursement, secu, tarif
droits	abus, affaire, association, avocat, certificat, cesarine, charte, citoyennete, civil, code, collectif, constitution, crime, delegation, democratie, dictature, EGN, electif, election, ENA, gouvernement, grassrootsnews, ideologie, inegalite, injustice, interassociatif, interdiction, jugement, juridique, jurisprudence, justification, legalite, legislation, legitime, loi...
revendication	Accouchement à domicile, accompagnement, accouchement non accompagne, alternative, aquatique, autonomie, avac, choix, combat, confidentialite, dignite, domicile, doula, droit, eau, feminisme, intime, justice, liberte, maison de naissance, obligation, petition ...

Les catégories spécifiques à cette liste concernent la liste elle-même, la qualification non médicale de la grossesse et de l'accouchement, la description « ordinaire » de phénomènes corporels (corps), l'enfant et les soins qu'on lui prodigue (enfant), les termes médicaux a priori non spécifiques de l'obstétrique (médecine), les termes médicaux de l'obstétrique (obstétrique), le système de soins, l'économie, les termes en rapport avec les droits des personnes et les moyens utilisés pour les faire respecter (droits), un certain nombre de revendications.

En première instance, le travail de catégorisation permet de faire une cartographie des sujets sur lesquels les participants aux listes discutent. Comme nous l'avons suggéré plus haut, c'est une méthode relativement efficace et précise.

Nous avons testé parallèlement l'utilisation du logiciel Alceste qui segmente le matériau en fonction des cooccurrences de termes : reste à voir si cette segmentation peut ensuite faire l'objet d'une catégorisation de la part de l'analyste. Un prélèvement au hasard d'un message sur 10 a été fait dans chacune des listes. Les résultats sont contrastés. En ce qui concerne la liste TDAH, le logiciel aboutit à 5 classes d'énoncés qui correspondent grosso modo au trouble, à l'association et à la liste, à la famille, à l'école, et au médicament : globalement, cette catégorisation fait sens, mais laisse de côté des aspects (par exemple état mental) qui se caractérisent par la très grande variété des termes qui y font référence. En revanche, pour la liste Naissance, les résultats sont inutilisables : le logiciel s'avère incapable de construire des classes stables sur deux années, et, le reste du temps, débouche sur des classes qui ne sont pas pertinentes, car elles reposent sur des mots très génériques. Ceci illustre bien les difficultés que pose le matériau des listes, très peu « cadré » et qui comprend des formes d'expression extrêmement variables. La particulière inefficacité du logiciel sur la liste Naissance nous apprend paradoxalement quelque chose sur la nature des échanges, à savoir le caractère intriqué des différents « répertoires » à l'intérieur des messages.

Première remarque : les thèmes de discussion sont relativement stables au cours du temps.

À partir de cette thématisation des sujets, il est possible d'en suivre l'évolution au cours du temps : les tableaux présentés ci dessous pour les deux listes montrent une grande stabilité dans la répartition des thèmes, la part de chacun d'entre eux variant pour la plupart de 15 à 20% autour de la moyenne.

	2001	2002	2003	2004	2005	2006	2007	Moyenne	Ecart-type
Naissance									
liste	4,1%	3,9%	3,2%	4,2%	3,3%	3,5%	3,4%	3,6%	0,4%
besoin	3,4%	2,6%	4,7%	4,7%	3,1%	5,2%	3,9%	4,0%	1,0%
relation	16,2%	17,8%	17,1%	16,2%	17,3%	16,9%	18,8%	17,4%	0,9%
Documentation	5,3%	2,9%	3,7%	3,5%	3,9%	3,0%	1,8%	3,1%	0,7%
information	6,7%	4,5%	3,7%	5,9%	5,2%	3,6%	4,5%	4,6%	0,9%
etat mental	4,6%	5,7%	7,1%	5,7%	3,2%	3,7%	5,8%	5,2%	1,5%
experience	3,4%	4,5%	5,7%	4,2%	5,4%	6,4%	8,0%	5,7%	1,4%
grossesse	1,6%	2,5%	2,1%	2,8%	3,4%	3,0%	3,0%	2,8%	0,5%
accouchement	8,3%	8,5%	7,4%	9,1%	10,6%	11,8%	11,6%	9,8%	1,8%
corps	7,4%	9,9%	9,4%	8,1%	8,5%	8,2%	5,8%	8,3%	1,4%
enfant	5,5%	7,7%	8,3%	5,1%	6,0%	7,2%	7,0%	6,9%	1,1%
famille	3,5%	3,2%	3,0%	3,1%	2,8%	2,1%	2,2%	2,8%	0,5%
medecine	6,7%	8,4%	5,6%	6,8%	7,1%	6,1%	5,8%	6,6%	1,0%
obstetrique	11,7%	9,6%	9,1%	12,1%	11,7%	9,9%	9,6%	10,4%	1,3%
Professionnel	3,6%	1,5%	2,0%	2,2%	2,3%	1,4%	0,5%	1,6%	0,7%
Systeme soins	0,9%	1,1%	1,4%	0,7%	1,1%	2,1%	1,0%	1,2%	0,5%
economie	1,1%	0,5%	0,7%	0,9%	0,6%	0,8%	0,8%	0,7%	0,2%
droits	2,4%	2,1%	2,3%	1,5%	1,9%	2,2%	1,7%	1,9%	0,3%
Revendication	3,9%	3,2%	3,5%	3,2%	2,5%	2,9%	4,8%	3,4%	0,8%

TDAH	2004	2005	2006	2007	2008	2009	Moyenne	Ecart-type
association	7,6%	6,0%	8,3%	8,5%	5,9%	6,8%	7,2%	1,1%
besoin	12,7%	12,1%	9,4%	10,2%	12,4%	9,9%	11,1%	1,4%
relation	14,7%	16,9%	15,1%	11,8%	15,6%	12,9%	14,5%	1,8%
documentation	3,3%	2,6%	2,6%	3,0%	1,2%	1,8%	2,4%	0,8%
information	9,8%	5,9%	6,0%	7,5%	5,1%	6,1%	6,7%	1,7%
etat mental	3,6%	4,9%	5,6%	3,9%	3,6%	4,1%	4,3%	0,8%
experience	4,2%	5,0%	4,2%	4,5%	3,5%	3,8%	4,2%	0,5%
manifestation	2,2%	2,7%	3,2%	1,5%	2,2%	2,7%	2,4%	0,6%
trouble	14,0%	11,1%	13,7%	12,8%	11,5%	11,2%	12,4%	1,3%
diagnostic	2,9%	3,3%	2,5%	3,7%	1,9%	5,0%	3,2%	1,1%
medicament	8,1%	5,2%	7,8%	8,1%	10,9%	11,0%	8,5%	2,2%
traitement	2,5%	4,6%	3,8%	2,0%	3,7%	3,5%	3,4%	0,9%
professionnel	2,2%	2,7%	2,1%	2,9%	4,7%	3,1%	3,0%	0,9%
institution	0,6%	1,8%	0,8%	1,5%	1,1%	1,2%	1,2%	0,4%
ecole	4,8%	6,4%	7,2%	9,3%	9,5%	10,3%	7,9%	2,2%
famille	6,8%	8,9%	7,5%	8,7%	7,0%	6,6%	7,6%	1,0%

Tableaux 12 : Évolution des thèmes au fil du temps

Cela ne signifie pas pour autant que ce qui circule sur la liste est stable, mais plutôt que la nature des préoccupations partagées par les membres de la liste l'est : ce travail de catégorisation permet justement d'identifier le cœur de ces préoccupations communes.

La comparaison entre la répartition des thèmes de toutes les conversations et des thèmes des conversations longues et/ou collectives fait apparaître des différences, révélant ainsi ce qui fait débat au sein du groupe.

Naissance	Toutes conv.	Conv. longues	Longues/ Toutes
liste	3,6%	2,4%	0,65
besoin	4,0%	3,9%	0,97
relation	17,1%	12,5%	0,73
medecine	6,7%	6,1%	0,92
documentation	3,6%	1,9%	0,52
information	4,7%	2,7%	0,57
etat subjectif	5,6%	6,1%	1,09
experience	5,1%	6,0%	1,19
grossesse	2,5%	3,1%	1,26
accouchement	8,8%	10,1%	1,15
corps	8,8%	10,8%	1,24
enfant	7,1%	9,7%	1,37
famille	3,0%	3,7%	1,22
obstetrique	10,3%	12,6%	1,23
professionnel	2,0%	1,4%	0,70
systeme de soins	1,2%	1,1%	0,93
economie	0,8%	0,8%	1,13
droits	2,1%	1,4%	0,66
revendication	3,4%	3,7%	1,11

TDAH	Toutes conv.	Conv. longues	Longues/ Toutes
association	7,1%	2,6%	0,37
besoin	11,3%	12,1%	1,07
relation	14,9%	9,8%	0,66
documentation	2,5%	2,1%	0,85
information	6,8%	3,9%	0,57
etat mental	4,3%	6,1%	1,41
experience	4,3%	5,2%	1,22
manifestation	2,5%	2,9%	1,17
trouble	12,4%	14,9%	1,20
diagnostic	3,1%	3,3%	1,05
medicament	8,2%	12,7%	1,56
traitement	3,5%	3,5%	1,00
professionnel	2,9%	3,0%	1,06
institution	1,2%	1,2%	0,99
ecole	7,4%	9,4%	1,27
famille	7,6%	7,2%	0,96

Tableaux 13 : comparaison des thèmes des conversations longues avec l'ensemble des conversations

De manière générale, les références au groupe (liste, association), aux relations entre les participants (relation), à l'information et à la documentation créent relativement peu de discussions que ce soit dans l'une ou l'autre liste. Dans la liste Naissance, le thème « enfant » est à l'origine d'un surplus de discussion principalement autour de la question de l'allaitement, alors que sur la liste TDAH, le médicament est un sujet majeur de débat – de fait, cette question est controversée même en dehors de la liste – ainsi que les sujets qui font référence à l'état mental sont aussi surreprésentés : ces sujets expriment presque toujours une situation de grand mal-être (désespoir, stress, désarroi, angoisse, enfer, énervement etc. sont les termes récurrents de ces sujets) face à laquelle les participants à la liste réagissent et essaient d'apporter du réconfort.

Deuxième remarque : la « spécialisation » des différentes catégories de locuteurs dépend des listes mais reste limitée

Les différentes catégories de locuteurs se différencient-elles du point de vue de leurs thèmes de prédilection ? Ceci permet-il de comprendre le tropisme des locuteurs vers les sujets lancés par leurs pairs ? Dans les tableaux suivants, nous avons représenté la répartition des thèmes des sujets lancés par les différents types de locuteurs.

Naissance	Dominant	Non Dominant
liste	4%	2%
besoin	4%	5%
Documentation	3%	1%
information	4%	2%
etat mental	7%	6%
experience	7%	7%
grossesse	3%	4%
accouchement	10%	13%
corps	12%	12%
enfant	9%	12%
famille	4%	4%
medecine	6%	8%
obstetrique	14%	14%
professionnel	2%	1%
systeme soins	1%	1%
économie	1%	1%
droits	2%	1%
revendication	5%	3%

TDAH	bénévole	dominant	non dominant
association	9%	2%	2%
besoin	11%	10%	14%
relation	4%	4%	3%
documentation	7%	2%	2%
information	4%	11%	8%
etat mental	6%	7%	5%
experience	9%	4%	6%
manifestation	6%	7%	8%
trouble	4%	5%	3%
diagnostic	5%	1%	0%
medicament	7%	3%	2%
traitement	6%	14%	13%
professionnel	2%	4%	2%
institution	5%	7%	13%
ecole	2%	4%	3%
famille	15%	14%	16%

Tableaux 14 : comparaison des thèmes des conversations longues lancées par les différents types de locuteurs

Plusieurs remarques peuvent être faites :

- Nous avons vu précédemment qu'une conversation longue attire proportionnellement davantage de locuteurs appartenant à la même catégorie que le lanceur. Cet effet était sensiblement plus fort sur la liste TDAH que sur la liste Naissance. Si l'on regarde les conversations longues lancées par les différentes catégories de locuteurs, on constate que les différences thématiques sont ici plus marquées sur la liste TDAH. Ceci donne quelque crédibilité à l'hypothèse d'une préférence thématique comme un des éléments permettant d'expliquer le tropisme des locuteurs à l'égard des conversations lancées par leurs pairs.
- Sur la liste Naissance, les locuteurs dominants manifestent un intérêt relatif un peu plus important à l'égard de la liste elle-même, de la documentation et de l'information, ainsi que des revendications. Les non-dominants sont eux plus centrés sur l'accouchement et l'enfant. Ceci dessine un espace légèrement polarisé dans lequel les locuteurs dominants seraient investis dans une réflexion plus politisée et

se situant à un niveau plus général que les locuteurs non dominants davantage centrés sur leur propre expérience.

- Sur la liste TDAH, le contraste le plus fort oppose les bénévoles et les autres types de locuteurs : dans les thèmes de prédilection des bénévoles, on voit se dessiner plus spécifiquement leurs domaines d'expertise.
- Globalement, cependant, on constate une assez forte similarité entre les profils des différents types de locuteurs : il ne semble pas y avoir de césure thématique entre ceux qui sont très actifs et ceux qui le sont moins ; ils s'intéressent grosso modo aux mêmes thématiques, ce qui ne veut évidemment pas dire que leurs contributions sont analogues.

Ceci étant, ce type d'analyse, s'il donne quelques indications utiles, reste assez rudimentaire. Au delà de ces chiffres bruts, il serait nécessaire de développer un traitement qui prenne en compte la multiplicité des thèmes pour un sujet. En effet, les différences entre les locuteurs se manifestent de manière plus complexe : par exemple, les conversations longues lancés par les bénévoles et qui se rattachent au thème « expérience » sont en fait toutes des appels à témoignages de manière à recueillir des informations capables d'étayer l'action associative, alors que celles qui sont lancées sur le même thème par les petits locuteurs sont toujours des appels à l'aide qui partent d'un énoncé décrivant un problème expérimenté par le locuteur.

Troisième remarque : les différences interindividuelles sont plus marquées que les différences entre catégories de locuteurs

La répartition des thèmes des sujets lancés et des sujets des messages dans lesquels un locuteur a posté un message constitue un profil thématique du locuteur en question. Bien entendu, un tel profil n'a de sens que pour des contributeurs importants. Qu'elle que soit la liste, on constate des différences importantes dans les profils de lancement, beaucoup moins sensible dans les profils de participation aux discussions. L'homogénéité constatée dans l'attachement des différents types de locuteurs aux thèmes de la liste résulte donc d'une agrégation de différences marquées entre des individus.

On pourrait imaginer sur la base de ces profils d'établir des classes de locuteurs qui seraient en fait transverses au découpage que nous avons opéré en distinguant les locuteurs en fonction de l'intensité de leur participation. Cependant ces classes ne permettent pas de faire d'autres différences que celles déjà constatées : la variété des positions et des intérêts par rapport aux différents thèmes centraux de la liste ne se traduit pas par une segmentation de la liste en sous forums dans lesquels viendraient débattre les locuteurs particulièrement intéressés par tel ou tel thème. En effet, les locuteurs les plus actifs sur chacune des listes « se mêlent » de tout et interviennent sur une grande variété de sujets, manifestant par là ce que l'on pourrait qualifier de « souci

de la communauté », de même qu'ils manifestent collectivement ce souci en prenant soin de répondre aux messages qui lancent de nouveaux sujets de conversation.

Conclusion

L'analyse de la dynamique des débats à l'échelle de la liste pose des difficultés techniques importantes. Pour contourner ces difficultés, nous avons proposé de partir de l'analyse des sujets. Ce travail, trop limité de notre point de vue et qui appellerait d'autres développements permet néanmoins de montrer que les thèmes de discussion ainsi identifiés sont, sur les deux listes, stables dans le temps. De manière à première vue paradoxale, c'est sans doute l'arrivée continue de nouveaux participants, confrontés aux mêmes types de difficultés ou d'interrogations, qui contribue à assurer cette stabilité. La spécialisation des catégories de locuteurs n'est pas très marquée, alors qu'à l'inverse, les individus manifestent au travers du lancement de sujets des intérêts fortement différenciés. Cependant, la participation beaucoup plus répartie des contributions aux conversations construit un maillage serré des thèmes et des contributeurs de sorte qu'il n'existe pas de segmentation des listes en sous-réseaux spécialisés.

Plusieurs points peuvent être soulignés :

- La « communauté » est activement performée par les participants, notamment les plus importants, en même temps qu'ils construisent collectivement l'espace des questions et des problèmes qui la définissent.
- La stabilité de cette communauté est non pas liée à une homogénéité des participants, mais à l'inverse à une diversité de ces participants. Parce que cette diversité traverse la stratification des locuteurs par rapport à leur niveau d'activité, elle contribue à la cohésion de l'ensemble. De ce point de vue, les listes illustrent à merveille la position de Tarde pour lequel la totalité émerge de l'agrégation d'entités individuelles différenciées et ne les explique pas (Latour 2010; Tarde 2005 (1898)).
- La stabilité de cette communauté est aussi liée à son permanent renouvellement qui réactive et remet sur le tapis les questions qui sont au cœur des préoccupations des personnes intéressées au sujet général de la liste.

La dynamique d'une question

La description que nous venons de faire pourrait laisser penser que les listes de discussion sont des espaces statiques dans lesquels se rejouent en permanence les mêmes échanges, qu'il s'agisse de conforter des personnes dans des situations plus ou moins difficiles, d'apporter des informations, ou de débattre collectivement des problèmes.

Il y a certainement une part de récurrence et de répétition dans le fonctionnement des listes ; cependant, l'impression que l'on peut avoir vis à vis du statisme/ dynamisme des

discussions, dépend selon nous de l'échelle à laquelle on se positionne. Tant que l'on reste au niveau de la liste entière – au moins sur nos deux listes – la stabilité prédomine pour les raisons que l'on a énumérées plus haut. Si l'on veut pouvoir apprécier l'évolution de la liste, il faut se situer à une échelle intermédiaire, autour d'un problème particulier, et rentrer dans les contenus. C'est ce que nous nous proposons de faire ici en prenant l'exemple de la thématique du handicap sur la liste TDAH.

Le choix de cette thématique est liée à l'activité qu'a déployée l'association à laquelle la liste est liée à partir du milieu des années 2000. Cette étude de cas sera donc l'occasion de se demander aussi ce que fait l'association à la liste et ce que la liste fait à l'association.

Après une période assez longue de distance vis à vis de la notion de handicap, l'association a pris en effet un virage important en participant par le biais de sa présidente au groupe « handicap cognitif », dans le cadre du comité interministériel chargé de la mise en œuvre de la loi de 2005 sur le Handicap.

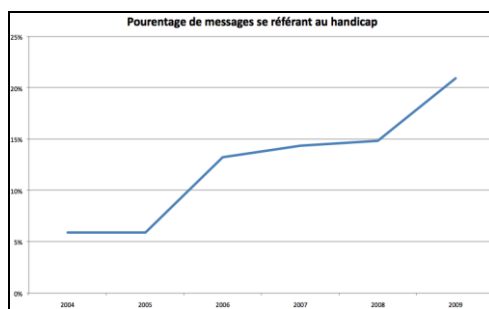
La notion de handicap cognitif vise, au plan des connaissances et de la prise en charge, à soustraire le trouble du tout neurologique et du tout psychiatrique, et à faire reconnaître :

- que le trouble est un trouble du développement avec une dimension neurologique claire, mais une dimension qui ne concerne pas seulement les processus d'apprentissage mais aussi les processus d'attention, de mémorisation, de remémoration.
- que le trouble est multifactoriel, et notamment que ses manifestations dépendent aussi de l'environnement dans lequel l'enfant va grandir. Certains parlent d'une inadaptation sociale, qui provoque des limitations de participation. Compte tenu de la plasticité du cerveau, notamment dans les premières années de la vie, si l'environnement est adapté, alors l'enfant aura plus de chances d'évoluer correctement. La notion de handicap vient ici appuyer un argumentaire qui doit, entre autres, convaincre l'école de mettre en place des aménagements pédagogiques.
- que la prise en charge du trouble est nécessairement multimodale. Le médicament seul ne suffit pas. La psychanalyse, dans certaines situations, peut être utile, car elle permet de contenir un jeune sur de longues périodes, ce que ne savent pas nécessairement faire d'autres praticiens. Les thérapies familiales, l'accompagnement des parents sont à développer. Surtout, il faut ajuster la prise en charge à chaque jeune, au moment présent.

La notion de « handicap cognitif » se trouve placée finalement en position centrale dans la stratégie de l'association. Elle permet 1) d'asseoir une définition du trouble et de se positionner par rapport à des définitions concurrentes, 2) d'inscrire le TDAH dans la politique du handicap, c'est-à-dire d'ouvrir la porte à une démarche multimodale qui ait pour objectif de prendre en compte le trouble dans l'ensemble de ses aspects, en cherchant non seulement à le traiter et le corriger, mais à pallier les déficits qui lui sont associés.

La stratégie de l'association se traduit-elle sur la liste et si oui comment ? La liste joue-t-elle un rôle dans l'évolution de l'association ?

Premier constat, le thème du handicap a subi une croissance forte de 2004 à 2009 comme on peut le voir sur le graphique suivant. Il est important de noter que les pourcentages figurant sur le graphique ont été obtenus en sélectionnant tous les messages contenant la chaîne « handicap », que ce soit dans la partie du message spécifiquement rédigée par l'expéditeur ou dans les messages précédents repris dans son propre message.



Graphique 16 : Pourcentage de messages se référant au handicap

Face, une fois de plus, à l'abondance du matériau, nous avons choisi la stratégie suivante : nous avons sélectionné tous les messages contenant la chaîne « handicap » dans les 1000 premiers caractères, afin de limiter la redondance, ce qui nous a fourni un échantillon représentant de 20 à 30% des messages en rapport avec le handicap.

Nous avons parcouru l'ensemble du matériau et avons fait une première analyse de contenu qui met en évidence deux grandes modalités du terme handicap ou de ses dérivés :

- Le handicap apparaît comme une catégorie d'autodescription ambivalente : plusieurs discussions se concentrent sur la question du vocabulaire adéquat pour décrire le TDAH et communiquer avec les personnes extérieures au problème.
- Le handicap apparaît aussi comme une catégorie administrative sophistiquée qui peut être utilisée afin de faire reconnaître le TDAH et le faire prendre en compte par le système scolaire.

Ces deux modalités sont temporellement distinctes et nous allons donc les présenter dans l'ordre de leur apparition dans les discussions.

Vivre avec le TDAH : le handicap comme mode d'auto-description ?

Le mot « handicap » tel qu'il apparaît en 2004-2005 dans les échanges est porteur d'une grande ambiguïté et d'une grande ambivalence. Il est utilisé dans un sens euphémisé pour décrire les conséquences du TDAH : dans cette perspective le TDAH est une pathologie dont les manifestations peuvent créer des difficultés spécifiques, mais ces

difficultés peuvent être transitoires, et un certain nombre de thérapies peuvent les alléger.³⁵

Mais le terme est aussi utilisé dans un sens plus fort pour désigner quelque chose susceptible d'imprimer sa marque sur une vie entière ; c'est une des raisons qui poussent les parents à en refuser l'utilisation. L'acceptation du handicap devient l'acceptation d'un destin sur lequel on ne peut plus influencer³⁶. Une mère raconte que, malgré les difficultés qu'elle rencontre, elle ne peut se résoudre à commencer de remplir le formulaire qui lui permettrait d'engager une procédure pour bénéficier d'une reconnaissance du handicap de son fils, car cela reviendrait à reconnaître le caractère permanent de ses difficultés.

Dans ces deux configurations, le mot « handicap » est très souvent mis en relation avec les troubles associés au TDAH plus qu'avec le TDAH lui-même : beaucoup d'enfants souffrent de dyslexie, dysgraphie, dyspraxie etc., toutes difficultés qui ont été officiellement reconnues comme des handicaps depuis déjà un certain temps.

Parfois, les parents racontent qu'à leur grande surprise, leur enfant revendique d'être un handicapé. Ce qui provoque des réactions fortes des parents qui rejettent l'idée que leur enfant puisse penser sa propre vie en ces termes. L'extrait suivant provient d'une conversation unique en 2005 dont le sujet était justement « handicap » :

³⁵ « Mon fils travaille très lentement, ce qui l'handicape. »

« Mon fils aussi souffre de dysgraphie) ; tu ne devrais pas trop t'inquiéter) ; avec un bon suivi, c'est OK. Mais pour être capable de progresser à l'école, il a besoin d'une adaptation à son handicap »

« Mon fils a est un petit bonhomme de 7 ans qui était en échec scolaire parce qu'il était incapable de produire ce qu'il savait. Le déficit d'attention l'handicapait sans ses apprentissages. Il était également très mal dans sa peau de se sentir en difficulté sans comprendre pourquoi. Nous avons accepté la mise sous Rita. Ça n'a pas changé fondamentalement notre petit bonhomme, mais ça lui permet de se concentrer et quelle différence pour lui !!

Dans certains cas même, le handicap est perçu comme étant causé par le traitement : « Changement de traitement à la rentrée, et là tout est devenu catastrophique, de jour en jour des tics apparaissaient, toujours plus nombreux et handicapants car à l'école surtout cela se voyait beaucoup et les autres enfants commençaient à se moquer. »

³⁶ « Pour les enfants présentant d'autre troubles (dysgraphie, dyspraxie, problème de vue...) le médicament ne fait pas de miracle. N'en déplaise à ses détracteurs, la R... permet juste de mettre de l'ordre dans la tête... Pour le reste, c'est comme pour tous les autres enfants, il y a la rééducation, la volonté et parfois, malheureusement, l'acceptation du handicap. »

« Quant au redoublement tu as eu raison de le refuser car de toute façon le handicap est là, il faut mieux l'aider et ne pas démotiver l'enfant en le faisant redoubler. »

« Moi quand mon fils m'a parlé de handicap, je lui ai expliqué qu'il était handicapable, que malgré ses difficultés dans certains domaines, il en avait d'autres qu'il pouvait développer et qui lui permettrait de faire des choses que les autres enfants ne pourraient pas... »

« L'autre jour, j'ai été surprise d'entendre mon fils de 8 ans dire : « je suis un handicapé » sans la moindre connotation négative. OK, il n'est pas dans un fauteuil roulant, mais je ne peux pas dire qu'il n'a pas de handicap : il est complètement déscolarisé et les troubles lui pourrissent la vie (...) C'est délicat car je crains que + tard il se sente diminué par cet état de fait (même si je garde espoir quand à une nette amélioration au fil du temps).. »

« J'ai un agenda 2004 qui est couvert de RDV avec des médecins et des thérapeutes (pour mon fils et moi-même) et début 2005, j'ai décidé qu'il fallait que cela change ! Il y a un mois, il m'a dit en rentrant de l'école « tu sais maman, le remplaçant, il ne savait pas que je suis handicapé » ; j'ai répondu « moi non plus ! » ; il a insisté : « tu sais, le TDAH, la précocité, l'hypermétropie, la dysgraphie, la dyslexie... je sais que je ne suis pas comme les autres ». Je suis restée sereine, mais je me suis dit qu'il était urgent qu'il soit « comme les autres. Ne vous méprenez pas : je suis lucide, je sais qu'il sera toujours un enfant particulier (d'abord parce qu'il est le plus radieux des soleils). Cela fait 7 ans qu'il va à l'école et que je me bats pour le faire rentrer dans le moule. Il est temps d'admettre que c'est un échec complet. À partir d'aujourd'hui, je vais prendre les choses dans l'autre sens : je vais prendre en considération les besoins et les capacités de mon fils et on va adapter le système. »

Nous observons donc une variété de réactions, même si elles convergent toutes dans le refus de l'étiquette « handicap ». La première met en avant la valeur positive de la différence. La seconde refuse le fait que le handicap puisse devenir constitutif de l'identité de son fils. La troisième – en dehors du fait qu'elle valorise la différence (*le plus radieux des soleils*) comme la première et qu'elle refuse le handicap comme la seconde – reproduit de manière instinctive le renversement opéré par les mouvements du domaine : ils considèrent que le handicap n'est pas un attribut des individus mais se constitue au croisement entre les individus et la société ; en conséquence, une société qui a fait de l'égalité d'accès un principe fort doit créer les conditions pour rendre cette égalité réelle pour tous ses membres. D'un certain point de vue, les parents qui utilisent le mot « handicap » pour décrire les éléments dans l'environnement qui le mettent en difficulté s'inscrivent dans la même perspective : « l'écriture le handicape » ; « redoubler une classe est un handicap, car le redoublant est censé mieux réussir que ceux qui « n'ont pas eu la chance » d'explorer le programme une première fois. ». Nous verrons dans la suite comment ce genre d'arguments s'est répandu dans le groupe.

Les parents doivent affronter aussi le regard des autres et il y a un certain nombre de circonstances dans lesquelles le mot « handicap » prend la forme de l'insulte, de la répulsion, de l'exclusion. Une mère explique comment elle s'est sentie rejetée par quelqu'un qu'elle considérait comme une amie : alors qu'elle l'invitait à l'accompagner à une réunion de l'association, son « amie » lui a répondu que « ça ne lui disait rien de voir un groupe de mongoliens errant comme des âmes en peine ».

Le risque de stigmatisation a conduit un certain nombre de participants au groupe à se demander comment ils doivent décrire le problème auquel leur enfant est confronté. AU départ, quelqu'un sur la liste raconte avoir posé la question au Dr Lecendreau qui, lors d'une assemblée générale de l'association en 2004, avait fait un exposé sur la médiatisation du TDAH. Celui-ci répondit que « le mot *maladie* était le plus puissant sur le plan médiatique, mais moins approprié du point de vue scientifique, que le mot *trouble* ne devait pas être accompagné par l'adjectif « neurologique » afin de ne pas anticiper sur le résultat des recherches, et qu'il fallait éviter le mot *handicap* qui représente un stade plus évolué du trouble ». Il y a donc à l'époque une convergence entre la réticence des parents à utiliser le mot « handicap » et la position prise par le médecin. Ce n'est qu'un peu plus tard que la question du vocabulaire va émerger dans la liste de discussion comme un problème, à un moment où l'association préconise de rechercher la reconnaissance du « handicap » comme un moyen permettant de négocier l'intégration des enfants dans le système scolaire.

Le handicap comme outil de négociation : la loi de 2005 sur le handicap et son impact

Curieusement, bien que l'association ait été impliquée dans la préparation de la loi de 2005, cet épisode est pratiquement invisible en tant que tel sur la liste de discussion³⁷. Mais comme nous allons le voir, cette loi a eu un effet massif sur la manière dont les difficultés rencontrées par les parents ont été traitées sur la liste.

Sur la liste, les discussions autour de la loi de 2005 sur le handicap ont concerné principalement un amendement déposé par des sénateurs à la fin de 2004 cet amendement était supposé donner à l'administration la possibilité de ne pas prendre en compte la volonté des parents de maintenir leur enfant dans le système scolaire ordinaire « au cas où ce maintien pourrait affecter durablement la communauté des élèves. » La liste de discussion a servi alors comme moyen de mobilisation afin que des personnes concernées envoient des lettres à leurs députés. Mais bien que cet épisode ait duré près de deux mois sur la liste, il n'y a pas eu d'annonce triomphale une fois la loi votée sans l'amendement en question.

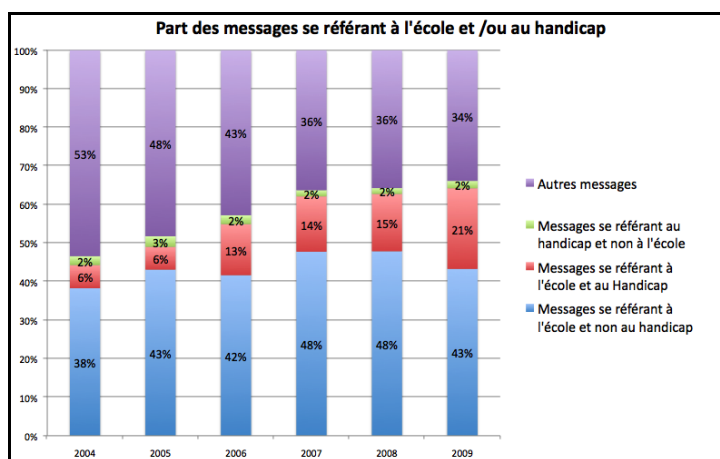
Cependant, si l'on regarde ce qui se passe sur la liste, on observe une intrication de plus en plus forte entre les questions scolaires et le handicap. Comme on peut le voir sur ce graphique³⁸, les questions relatives à l'école et/ou au handicap ont toujours occupé une

³⁷ Ce constat illustre la position qu'occupe la liste vis à vis de l'association : elle représente un moyen que se donne l'association d'atteindre les personnes concernées, de leur venir en aide et, dans un second temps, de recruter des adhérents et des bénévoles. Elle n'est pas du tout un espace dans lequel la politique de l'association serait discutée.

³⁸ Les messages au sujet de « l'école » ont été sélectionnés s'ils contiennent au moins l'une des chaînes de caractères suivantes : scola, école, maîtr, AVS (auxiliaire de vie scolaire), PPS (plan personnalisé de scolarisation). Les messages au sujet du « handicap » contiennent la chaîne de caractères handicap qui peut figurer dans des formes substantive, adjective et verbale ainsi que

grande part des messages, mais cette proportion s'est accrue pendant la période étudiée : alors qu'elles étaient présentes dans 47% des messages en 2004, on peut les trouver dans 66% des messages en 2009.

Pendant la même période, les questions relatives au handicap ont crû de manière encore plus marquée : elles étaient présentes dans 8% des messages en 2004 et dans 23% en 2009.



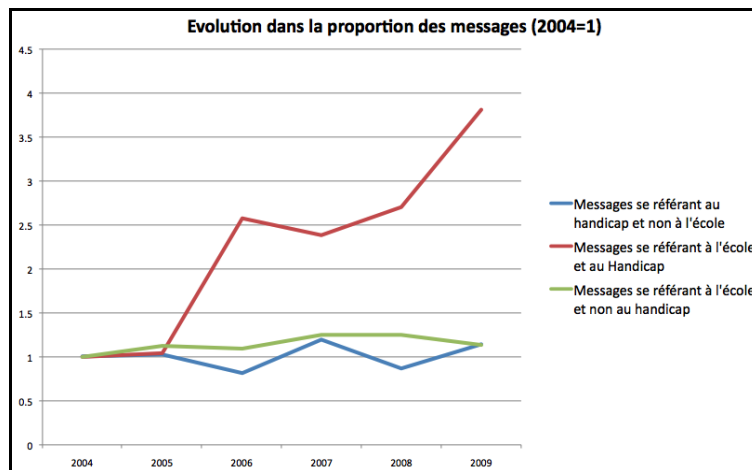
Graphique 17 : part des messages contenant des références à l'école et/ou au handicap

Mais alors que la proportion des messages contenant des références au handicap sans référence à l'école est restée stable tout au long de la période, le recouvrement de plus en plus grand des thèmes de l'école et du handicap est très frappant : il y avait seulement 5% de messages contenant des références aux deux thématiques en 2004-2005 contre 19% en 2009. En fait, l'essentiel de la croissance de la catégorie « école » a été faite au travail de l'expansion de la catégorie « école et handicap ».

En résumé, l'augmentation de l'intérêt pour le handicap sur la liste a à voir avec les questions relatives à l'école et réciproquement.

Dans le graphique suivant, on peut voir que les taux de croissance de la catégorie « école » sans handicap et de la catégorie « handicap » sans école sont du même ordre, environ 20%, alors que le taux de croissance combiné « école et handicap » est de près de 300%. Le basculement entre la situation initiale et la nouvelle situation s'est opéré en 2006, soit l'année qui a suivi le vote de la loi sur le handicap.

les acronymes MDPH (maison départementale des personnes handicapées) et COTOREP qui est en quelque sorte l'ancêtre de la MDPH.



Graphique 18 : Évolution dans la proportion des messages

Qui a amené le thème du handicap dans les discussions ? Globalement 5 contributeurs (sur 175 concernés) représentent 38% des messages contenant une référence au handicap, alors que ces 5 locuteurs comptent pour 20% de l'activité totale de la liste.

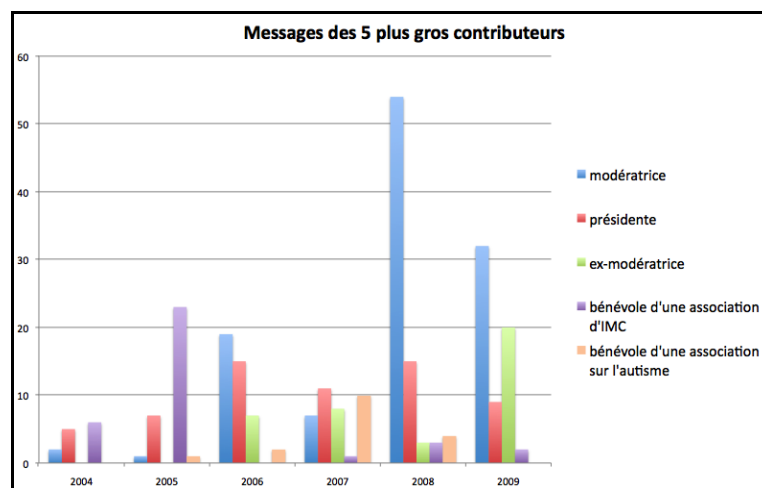
- La participante la plus prolixe (16% des messages) est aussi la première contributrice de la liste (7% des messages) : elle n'était pas une bénévole de l'association, mais compte tenu de son investissement dans le groupe, elle a été « promue » en tant que modératrice.
- La seconde participante est la présidente de l'association.
- La troisième a un profil analogue à la première et a été aussi quelque temps modératrice.
- La quatrième est une bénévole dans une association de soutien aux parents d'enfants avec une infirmité motrice cérébrale
- Enfin la dernière est une bénévole de l'association.

À l'intérieur du groupe formé par les 20 contributeurs les plus importants sur ce thème, on trouve 6 bénévoles de l'association, 2 bénévoles d'associations d'autiste, 1 bénévole de l'association d'IMC³⁹. Il y a donc une forte représentation de bénévoles y compris d'autres associations dans le groupe leader sur ces questions de handicap ; il faut noter cependant que la distribution des messages entre les bénévoles, dominants, petits locuteurs est à peu près conforme à celle de la liste en général ? Quel rôle jouent ces bénévoles ?

³⁹ La liste est ouverte à tous ceux qui se sentent concernés par le TDAH, qu'ils soient ou non membres de l'association. De plus le TDAH est souvent associé à d'autres pathologies, quelques participants sont membres d'autres associations.

Les messages postés par les 5 premiers contributeurs peuvent être groupés en deux catégories : (i) des actualités telles la publication d'un article de journal, ou une nouvelle règle administrative par exemple, (ii) des réponses à des messages envoyés par des parents ordinaires à la recherche d'informations, de soutien, d'avis. Dans cette dernière configuration, de loin la plus commune, le message initial ne contient souvent aucune référence directe au handicap : nous observons que ces premiers contributeurs jouent un rôle important dans la traduction des préoccupations des parents – principalement en matière scolaire – dans le langage du handicap, leur donnant ainsi accès aux outils spécifiques qui lui sont associés.

Le niveau d'expertise accumulé sur ces questions s'est accru au fil du temps, et ceci est particulièrement vrai pour la première locutrice qui tient une place considérable dans ces échanges comme on le voit sur le graphique suivant.



Graphique 19 : Messages des 5 premiers contributeurs aux échanges sur le handicap

Il est intéressant de regarder d'un peu plus près ses messages et leur progression ; au début de ses interventions, elle est en position de recherche :

« j'aimerais bien savoir en fait si officiellement le TDAH fait partie de la liste des handicaps. On a parlé de tiers temps supplémentaire et de programme allégé : qui en décide, comment cela se passe-t-il? Pour l'instant je ne pense pas demander cela pour Simon mais j'aimerais me renseigner pour savoir le cas échéant » 2004

« Depuis quelques semaines Simon se rend de plus en plus compte de son handicap. Il a échoué à l'apprentissage de la multiplication et toute la classe se moque de lui, le traite de débile voire de fou (à l'interclasse de la cantine il est sans ritaline entre midi et 13h, donc comme je le connais il doit tourner comme une toupie sans but, incapable de participer à quoi que ce soit). Comment expliquer calmement et de manière rassurante à mon fils que en plus de l'hyperactivité, du déficit de l'attention, maintenant il va falloir faire avec des difficultés d'apprentissage additionnelles. Comment lui parler de la dyspraxie visuo-spatiale, de son avenir à l'école? Comment expliquer aux autres enfants de sa classe qu'il faut l'aider et pas l'enfoncer? » 2006

Dans un second temps, elle raconte un certain nombre d'épisodes de ses démarches en vue d'obtenir une reconnaissance de handicap, demande des conseils ou des avis, puis peu à peu se positionne quasiment en personne référente sur ces questions ; pendant la période 2008-2009, elle transmet régulièrement les dépêches ou les articles sur le sujet qui seraient susceptibles d'intéresser les autres parents ; ses messages deviennent de plus en plus directs et donnent des conseils précis et concis :

« Le taux de handicap est forcément situé entre 50 et 79% et on ne précise pas combien, c'est la fourchette qui est précisée je pense As tu demandé un AEEH (allocation pour enfant handicapé) pour les thérapies si tu en as à payer? »

« C., si vous obtenez une reconnaissance MDPH, ils ont TORT! en effet ce n'est pas à l'enfant qui a un handicap de s'adapter mais à l'école et aux enseignants de s'adapter à LUI. Donc la liste des démarches que je t'ai faite reste valable, mais auras tu le temps de franchir toutes les étapes? Appelle l'enseignant référent pour le handicap (ERH) dont le nom doit être obligatoirement posté sur le tableau d'affichage de l'école, elle pourra intervenir TRES rapidement à ta demande. Allez, on y va :). »

« Qu'il oublie ses affaires, c'est normal, c'est un TDAH. Maintenant il faut trouver une solution, tu as les choix suivants: faire accepter par la maîtresse avec une lettre du médecin d'apporter une aide au cartable; si ça ne marche pas tu contactes l'enseignant référent pour le handicap de ton coin (ses coordonnées doivent être affichées devant l'école) et tu lui expliques que ton enfant a besoin d'une aide au cartable (et peut être d'autres adaptations), consulte pour cela le livret de l'association Le Tdah et l'école et note les aménagements qui seraient utiles à ton fils. Tu lui demandes si elle peut mettre ça en place avec ou sans PPS (ou avec un PAI). Si PPS, elle convoquera une réunion de l'équipe éducative en vue de la signature du PPS et tu pourras expliquer les problèmes résultant du TDAH nécessitant l'application des dispositions qui y seront inscrites. Le PPS s'impose à toutes les parties. »

« Tu fais bien de nous écrire car je pense que ce qui se trame pourrait mettre ton enfant en danger. D'abord, la réunion de l'équipe éducative doit se clore par un compte rendu que tu dois VOIR de tes yeux et signer, et bien entendu que tu devrais en avoir une copie. Qui était présent? Est ce que tu as fait appel à l'ERH (enseignant référent pour le handicap, qui doit avoir une bonne connaissance du TDAH) ? D'autre part, le PPRE pour un enfant TDAH qui ne sait pas lire, ça peut être assez dangereux car c'est la première étape pour autoriser ensuite soit un redoublement, voire une exclusion de l'école si l'enfant ne remplit pas le contrat. Qu'est ce qui a été mis dans ce PPRE? quelles aides concrètes aura l'enfant pour progresser? Bénéficie-t-il d'un traitement et quelle est cette histoire de cachet ??? La ritaline n'est pas un médicament qui donne la science infuse, donc s'il a pris du retard, il faut l'aider à le rattraper ! »

« Pour obtenir 50% [de handicap] normalement il faut une gêne majeure dans les actes de la vie quotidienne. C'est pourquoi les dyslexiques en général ne l'obtiennent pas, ni les TDAH bien traités. Par contre avec un second trouble dys ça peut être jouable, notamment la dyspraxie (je répondais ça surtout pour M. car son fils pourrait y prétendre) En tous cas, ma réponse est valable selon le guide barème que j'ai utilisé pour la demande de reconnaissance de handicap et l'AEEH à la MDPH. Cela dit, pour avoir droit à un PPS il n'y a pas besoin d'avoir 50%, n'importe quel diagnostic fiable suffit pour avoir droit à des adaptations scolaires. »

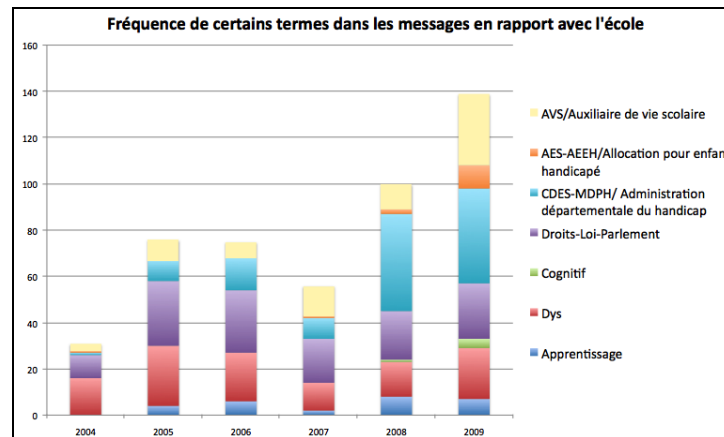
Depuis que la loi de 2005 a été adoptée, les leaders de la liste de discussion – qui sont aussi pour partie des bénévoles de l'association – se sont appropriés son contenu. Comme on le voit sur ces différents extraits, une expertise considérable a été accumulée, notamment au travers de l'expérience directe qu'ont fait les parents des dispositifs mis en place par la loi : de ce point de vue, on constate que les bénévoles ne sont pas d'emblée en position surplombante maîtrisant des contenus qui échapperaient aux parents ordinaires et dispensant, depuis cette position, informations et conseils. Malgré des asymétries évidentes, il y a une véritable continuité entre les différents participants qui devient visible dès lors que le monde extérieur confronte les uns et les autres à une situation nouvelle. Du coup, la liste devient l'un des lieux dans lequel peut se reconstituer une expertise qui est, dans un second temps, mobilisée dans les échanges pour venir en aide aux parents. Cette expertise concerne de multiples aspects et intègre différents types de savoir : connaissances juridiques, connaissances sur le fonctionnement de l'administration, connaissances tirées de l'expérience sur la meilleure manière de présenter son dossier, sur l'attitude la plus efficace, sur les alliances qui peuvent être construites avec d'autres parents, avec l'enseignant, comme on peut le voir sur ce message envoyé par une bénévole :

« Si votre enfant est diagnostiqué, alors demandez un rendez-vous au médecin scolaire et parlez lui de vos difficultés, notamment des petits soucis que vous avez avec l'enseignant, de la mise en place des aménagements scolaires possibles ; bien souvent, l'instit ou le directeur écoute le médecin mais pas les parents parce qu'il a du poids et une crédibilité que vous n'avez pas ; en outre à chaque problème, prenez le comme intermédiaire cela évitera de vous énerver (Ne jamais dire du mal d'un enseignant). Le médecin scolaire est formé, il y a de fortes chances pour qu'il explique que le trouble TDAH est un vrai handicap et que votre enfant ne profite pas de la situation. Si cela vient du médecin, l'instit ne pourra s'y opposer. Mais en situation générale TRAITÉZ VOS ENSEIGNANTS COMME VOUS VOUDRIEZ QU'ILS TRAITENT VOS ENFANTS, AVEC PATIENCE ET EN LES VALORISANT. N'oubliez pas que le TDAH met vos enseignants en échec, vous pouvez penser d'eux qu'ils sont nuls et incompetents, mais vous pouvez aussi les encourager à persévérer dans leur effort de compréhension du TDAH ; si vous commencez votre entretien en les remerciant de leur patience pour votre enfant qui n'est pas de tout repos, si vous lui demandez de l'aide, en un mot si vous êtes capable de faire vous-même ce que vous exigez de vos enseignants, à savoir CHANGER LE REGARD QUE VOUS PORTEZ SUR EUX, il y a de fortes chances pour qu'un miracle se produise. »

Le contenu des discussions portant sur l'école et le handicap a évolué au fil du temps, comme on peut l'observer sur le graphique suivant. Celui-ci est basé sur un décompte des messages qui contiennent des termes spécifiques : certains messages en contiennent plus qu'un, ce qui induit une légère redondance (de l'ordre de 20% des messages), mais ne bouleverse pas significativement l'équilibre général. Ces termes appartiennent à trois catégories :

- Ils qualifient des problèmes cognitifs fréquemment associés au TDAH (le spectre des dys : -lexie, -graphie, -praxie, -orthographe, -calculie) ; mots associés : dys, apprentissage, cognitif

- Ils sont en rapport avec les aspects juridiques de la question ; mots associés : droits, loi, parlement
- Ils concernent la « gestion » du handicap ; mots (acronymes) associés : AVS, AES, ADDH, CDES, MDPH.



Graphique 20 : Fréquence de certains termes dans les messages en rapport avec l'école

En 2004, les discussions se concentraient principalement sur les pathologies « dys » et sur des questions en rapport avec la préparation de la loi de 2005. Durant les années suivantes, on assiste à l'émergence de préoccupations en lien avec la reconnaissance du handicap et les aménagements scolaire. En 2008-2009, ces questions « explosent » et deviennent le centre des discussions : grâce à la loi sur le handicap, au lien effectué par l'association entre TDAH et Handicap, et à l'expertise accumulée peu à peu par les parents, le handicap est devenu le moyen par lequel les parents peuvent enfin avoir prise sur la vie scolaire de leur enfant. Non seulement le passage par le handicap permet d'ériger des dispositifs tiers entre les parents et l'école, détournant ainsi la violence qui résultait souvent de leur confrontation directe, mais il offre toute une panoplie de ressources organisationnelles, pédagogiques, thérapeutiques, financières susceptibles de soutenir les parents et leurs enfants dans leur combat. Dans ce contexte, la liste joue un rôle majeur de traduction et d'acculturation des parents au monde du handicap ; de surcroît, elle permet de maintenir un flux de circulation des informations et des expériences qui contribue à la mise à jour des connaissances et au renforcement de l'expertise collective sur ces questions.

Conclusion

Dans cette partie, nous avons voulu montrer comment il était possible de mobiliser une pluralité de méthodes afin d'analyser la dynamique d'une question, à la fois sur le plan du contenu et de son importance relative. Nous avons combiné des analyses de contenu classiques qui ne peuvent se déployer que un échantillon relativement limité, avec des analyses statistiques (i) permettant de suivre l'évolution de la part prise par certains termes et par leur combinaison, et (ii) qui relie la production de messages à leurs

auteurs et permettent de positionner les auteurs les uns vis à vis des autres dans cette discussion et plus globalement dans l'activité de la liste.

Cette combinaison d'approche permet de décrire quelque chose comme l'émergence d'une « action collective » qui n'est pas le produit d'une décision délibérée enclenchant une cascade d'actions pour la œuvre, mais le résultat de l'agrégation de micro-actions individuelles qui se répondent les unes aux autres, et qui, ensemble, finissent par construire une dynamique et des effets observables à un niveau global.

Pourquoi la combinaison des méthodes est-elle cruciale dans cette affaire ? Pourquoi ne pas s'en tenir simplement à une analyse de contenu classique ? D'abord, on l'a vu, le matériau est trop abondant pour que cela soit possible : on peut toujours prélever un échantillon comme nous l'avons fait. Cependant, en dehors d'un travail qui permette de resituer ce qui s'observe au niveau des messages dans le contexte global de la liste, et de comprendre comment s'articulent les interventions des différents intervenants, cela reste une analyse de textes qui se répondent les uns aux autres et qui déploient des représentations éventuellement contrastées. Pour parvenir à cerner ce en quoi ces échanges forment une action collective et ce en quoi la liste devient un agent dans un processus de transformation, il est nécessaire d'en passer par ce travail d'articulation entre les différentes dimensions, au sens presque littéral, des échanges, d'en prendre la mesure en un sens.

Nous avons développé de manière très exploratoire une série d'indicateurs, de méthodes, d'outils pour essayer de rendre compte de la dynamique des listes de discussion. L'ambition était avant tout descriptive et vise à donner des prises sur un type d'espace qui se présente à première vue comme le degré zéro de l'organisation sociale. L'analyste n'a en effet à sa disposition aucune des catégories commodes qui permettent, ailleurs, de planter un décor et de dérouler une intrigue à partir des bribes qui lui proviennent de ses entretiens ou documents. Il n'a paradoxalement qu'un trop plein de données entretenant l'illusion d'une possible connaissance totale et objective.

Nous avons été constamment guidé dans notre démarche par plusieurs exigences :

- essayer de construire un point de vue qui puisse s'articuler avec celui de l'utilisateur, d'où l'importance que nous avons accordée à un certain nombre de paramètres qui organisent sa perception, la récurrence des noms ou des sujets dans la boîte mail par exemple.
- prendre en compte la temporalité et redonner aux listes une dimension historique, ce qui implique notamment de s'écarter des approches structuralistes,
- ne pas inférer de modèle causal à partir d'une simple corrélation : il n'y a pas à proprement parler « d'explication » de la dynamique des listes, au sens où des variables données a priori permettraient de reconstituer quelque chose qui serait de l'ordre de la trajectoire ; la notion de boucle immergence/ émergence proposée par

Cointet (2009) est de ce point de vue beaucoup plus appropriée, car elle saisit la double nature de la liste, à la fois agrégation d'actions individuelles obéissant à des logiques variées et entité dotée de certaines propriétés plus ou moins durables, l'une et l'autre étant dans une relation de codétermination.

- de tenir ensemble les deux faces du message, la relation et le contenu.

La démarche conduit à deux types de résultats :

- Le développement de deux notions clé – locuteurs dominants, conversations longues – à partir desquels il est possible de construire un certain nombre d'indicateurs (taille et poids du groupe dominant, densité du groupe dominant, redondance des conversations, convivialité, réactivité, animation) qui permettent de décrire grossièrement la manière dont le groupe fonctionne et son évolution.
- Le développement de méthodes ad hoc qui, combinées aux analyses plus génériques, permettent d'avancer dans l'analyse d'une question donnée, en permettant une articulation entre la dynamique des contenus et la dynamique des relations.

L'ensemble de ce travail nous conduit à plaider pour le développement de « petits outils » ajustables par le chercheur au plus près des questions auxquelles il veut répondre. Deux raisons principales motivent cette position : (i) plus un outil est sophistiqué et plus il fait d'hypothèses fortes à la fois sur l'objet qu'il prétend analyser et sur les questions qui peuvent faire l'objet de l'analyse ; (ii) plus il est complexe et plus il conduit le chercheur à développer une interprétation plutôt qu'une description, c'est-à-dire plus il l'engage à construire des raisonnements de type causal et structural afin de combler le fossé existant entre les indicateurs produits et l'expérience de ce qu'est un groupe de discussion.

Bibliographie

- Adamic, L., & Huberman, B. (2002). Zipf's law and the Internet. *Glottometrics*, 3, 143-150.
- Adamic, L. A., & Huberman, B. A. (2001). The Web's hidden order. *Commun. ACM*, 44(9), 55-60.
- Akrich, M., & Méadel, C. (2001). Listes de discussion électroniques et nouveaux collectifs autour des problèmes de santé. Vers une caractérisation des collectifs et débats. (p. 102). Paris: CNRS - CSI.
- Akrich, M., & Méadel, C. (2002). Prendre ses médicaments/prendre la parole : les usages des médicaments par les patients dans les listes de discussion électroniques. *Sciences Sociales & Santé*, 20(1), 89-116.
- Akrich, M., & Méadel, C. (2004). Problématiser la question des usages. Introduction au n°spécial "Les technologies de l'information à l'épreuve des pratiques". *Sciences Sociales et Santé*, 22(1), 5-20.

- Akrich, M., & Méadel, C. (2009). Les échanges entre patients sur l'internet. *La presse médicale*(38), 1484-1493.
- Arguello, J., Butler, B., Joyce, E., Kraut, R., Ling, K., Rosé, C., et al. (2006). Talk to me: foundations for successful individual-group interactions in online communities. CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM Press. 959-968.
- Barabasi, A.-L., & Réka, A. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509-512.
- Barcellini, F., Détienne, F., Burkhardt, J.-M., & Sack, W. (2005). A study of on-line discussions in Open-Source Software Community: Reconstructing thematic coherence and argumentation from citation practices. *Communities and Technologies*. Milano.
- Barcellini, F., Détienne, F., Burkhardt, J.-M., & Sack, W. (2008). A socio-cognitive analysis of online design discussions in an Open Source Software community. *Interacting with Computers*, 20(1), 141-165.
- Barker, K. K. (2008). Electronic Support Groups, Patient-Consumers, and Medicalization: The Case of Contested Illness. *Journal of Health & Social Behavior*, 49(1), 20-36.
- Baym, N. (1996). Agreements and Disagreements in a Computer-Mediated Discussion. *Research on Language & Social Interaction*, 29(4), 315.
- Baym, N. (1997). Interpreting soap operas and creating community: inside an electronic fan culture. In S. Kiesler (Ed.), *Culture of the Internet* pp. 103-120). Mahwah: Lawrence Erlbaum Associates.
- Baym, N. (1998). The emergence of on-line community. In S. G. Jones (Ed.), *CyberSociety 2.0: Revisiting computer-mediated communication and community* pp. 35-68). Thousand Oaks CA: Sage.
- Baym, N. (2000a). *Tune In, Log On. Soaps, Fandom, and Online Community*. London: Sage.
- Baym, N. K. (2000b). *Tune In, Log On: Soaps, Fandom, and Online Community*. London: Sage.
- Beaudouin, V., & Velkovska, J. (1999). Présentation de soi et structures d'une communauté sur Internet. Actes du 2ème Colloque International sur les Usages et Services des Télécommunications. Bordeaux. 502-513.
- Benvegna, N., & Brugidou, M. (2008). Prendre la parole sur Internet. Des dispositifs sociotechniques aux grammaires de la discussion. *Réseaux*, 150(4), 51-82.
- Blanchard, A. (2004). Virtual Behavior Settings: An Application of Behavior Setting Theories to Virtual Communities. *Journal of Computer-Mediated Communication*, 9(2), 00-00.

- Brush, A. J. B., Wang, X., Turner, T. C., & Smith, M. A. (2005). Assessing differential usage of usenet social accounting meta-data. *Proceedings of the SIGCHI conference on Human factors in computing systems* pp. 889-898). Portland, Oregon, USA: ACM.
- Burke, M., Joyce, E., Kim, T., Anand, V., & Kraut, R. (2007). Introductions and Requests: Rhetorical Strategies That Elicit Response in Online Communities. In C. Steinfield, B. T. Pentland, M. Ackerman & N. Contractor (Eds.), *Communities and Technologies 2007* pp. 21-39). London: Springer.
- Burke, M., & Kraut, R. (2008). Mind your Ps and Qs: the impact of politeness and rudeness in online communities. *Proceedings of the 2008 ACM conference on Computer supported cooperative work* pp. 281-284). San Diego, CA, USA: ACM.
- Burke, M., Kraut, R., & Joyce, E. (2010). Membership Claims and Requests: Conversation-Level Newcomer Socialization Strategies in Online Groups. *Small Group Research*, 41(1), 4-40.
- Burnett, G., & Bonnici, L. (2003). Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library & Information Science Research*, 25(3), 333-351.
- Burnett, G., & Buerkle, H. (2004). Information Exchange in Virtual Communities: A Comparative Study. *Journal of Computer-Mediated Communication*, 9(2), 00-00.
- Butler, B., Sproull, L., Kiesler, S., & Kraut, R. (2007). Community Effort in Online Groups: Who Does the Work and Why? *Human-Computer Interaction Institute*. Paper 90. <http://repository.cmu.edu/hcii/90>. Accessed on May 4th, 2012.
- Butler, B. S. (2001). Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Information Systems Research*, 12(4), 346-362.
- Callon, M., & Ferrary, M. (2006). Les réseaux sociaux à l'aune de la théorie de l'acteur-réseau. *Sociologies Pratiques*, 2(13), 37-44.
- Caviale, O., & Bruillard, R. (2009). Les jeux d'acteurs sur des listes de discussion institutionnelles d'enseignants. *Réseaux*, 155(3), 137-176.
- Cointet, J.-P. (2009). Dynamiques sociales et sémantiques dans les communautés de savoirs. Morphogenèse et diffusion. (p. 265). Paris: Ecole Polytechnique.
- Conein, B., & Latapy, M. (2008). Les usages épistémiques des réseaux de communication électronique†: Le cas de l'Open-Source. *Sociologie du Travail*, 50(3), 331-352.
- Coursaris, C. K., & Liu, M. (2009). An analysis of social support exchanges in online HIV/AIDS self-help groups. *Computers in Human Behavior*, 25(4), 911-918.
- D'Souza, R. M., Borgs, C., Chayes, J. T., Berger, N., & Kleinberg, R. D. (2007). Emergence of tempered preferential attachment from optimization. *Proceedings of the National Academy of Sciences*, 104(15), 6112-6117.
- Donath, J. (1999). Identity and Deception in the Virtual Community. In M. A. Smith & P. Kollock (Eds.), *Communities in Cyberspac*. New York: Routledge.
- Donath, J. (2001). Mediated Faces. *Lecture Notes In Computer Science*, 2117, 373 - 390.

- Dorat, R., Latapy, M., Conein, B., & Auray, N. (2007). Multi-level analysis of an interaction network between individuals in a mailing-list. *Annals of telecommunications*, 62, 320-344.
- Edwards, S. L., & Huff, M. T. (2003). Anonymity in Electronic Discussions Groups: Effect on Diversity Discussion. *Journal of Baccalaureate Social Work*, 9(1), 131-147.
- Ekeblad, E. (2001). The Emergence and Decay of Multilogue. Self-Regulation of a Scholarly Mailing List. . *revised version of a paper presented in the symposium: Time and coordination in a virtual community of learners. European Association for Research on Learning and Instruction conference (EARLI 99): Advancing Learning Communities In The New Millennium, Göteborg, Sweden, August 24-28 1999.*, <http://lchc.ucsd.edu/mca/Paper/eva/The%20Dynamics%20of%20Multilogue.htm>.
- Fiore, A. T., Tiernan, S. L., & Smith, M. A. (2002). Observed behavior and perceived value of authors in usenet newsgroups: bridging the gap. Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves pp. 323-330). Minneapolis, Minnesota, USA: ACM.
- Fisher, D. (2005). Using Egocentric Networks to Understand Communication. IEEE Internet Computing pp. 20-28).
- Fisher, D., Smith, M., & Welser, H. (2006). You Are Who You Talk To: Detecting Roles in Usenet Newsgroups. HICSS '06: Proceedings of the 39th Annual Hawaii International Conference on System Sciences. IEEE Computer Society. 3, 1-10.
- Flichy, P. (2008). Internet et le débat démocratique. *Réseaux*, 150(4), 159-185.
- Flichy, P. (2010). La démocratie 2.0. *Etudes*, 412(5), 617-626.
- Gleave, E., Welser, H., Lento, T., & Smith, M. (2009). A Conceptual and Operational Definition of 'Social Role' in Online Community. 42nd Hawaii International Conference on System Sciences. Waikoloa, HI.
- Golder, S. A., & Donath, J. (2004). Social roles in electronic communities. *Internet Research*(5).
- Goos, M., & Bennison, A. (2008). Developing a communal identity as beginning teachers of mathematics: Emergence of an online community of practice. *Journal of Mathematics Teacher Education*, 11(1), 41-60.
- Groth, R. E. (2005-2006). Analysis of an Online Case Discussion about Teaching Stochastics. *Mathematics Teacher Education and Development*, 7, 53-71.
- Hansen, D., Shneiderman, B., & Smith, M. (2010). Visualizing Threaded Conversation Networks: Mining Message Boards and Email Lists for Actionable Insights. In A. An, P. Lingras, S. Petty & R. Huang (Eds.), *Active Media Technology* pp. 47-62): Springer Berlin / Heidelberg.
- Hara, N., Bonk, C., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, 28(2), 115-152.

- Harry, I., Gagnayre, R., & d'Ivernois, J.-F. (2008). Analyse des échanges écrits entre patients diabétiques sur les forums de discussion. Intérêt pour l'éducation thérapeutique du patient. *Qualitative Health Research*, 6(3), 393-412.
- Himmelboim, I. (2008). Reply distribution in online discussions: A comparative network analysis of political and health newsgroups. *Journal of Computer-Mediated Communication*, 14(1), 156-177.
- Himmelboim, I., Gleave, E., & Smith, M. (2009a). Discussion catalysts in online political discussions: Content importers and conversation starters. *Journal of Computer-Mediated Communication*, 14, 771-789.
- Himmelboim, I., Lavelle, S. M., & Kafri, R. (2009b). The Structure of an Online Discussion - Using Network Analysis to Study a Political Usenet Group. Annual meeting of the International Communication Association. Sheraton New York, New York City, NY: <http://www.allacademic.com/meta/p13870_index.html>.
- Huberman, B., & Adamic, L. (1999). Growth dynamics of the World-Wide Web. *Nature*, 401(6749), 131-131.
- Jensen, M. J., Danziger, J. N., & Venkatesh, A. (2007). Civil Society and Cyber Society: The Role of the Internet in Community Associations and Democratic Politics. *Information Society*, 23(1), 39-50.
- Jeong, A. C. (2003). The Sequential Analysis of Group Interaction and Critical Thinking in Online Threaded discussions. *American Journal of Distance Education*, 17(1), 25 - 43.
- Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration. *Information Systems Research*, 15(2), 194-210.
- Jouët, J. (1987). La sociabilité télématique. *Communication et langages*(72), 78-87.
- Joyce, E., & Kraut, R. (2006). Predicting Continued Participation in Newsgroups. *Journal of Computer-Mediated Communication*, 11(3), 723-747.
- Kalman, Y. M., Ravid, G., Raban, D. R., & Rafaeli, S. (2006). Pauses and Response Latencies: A Chronemic Analysis of Asynchronous CMC. *Journal of Computer-Mediated Communication*, 12, 1-23.
- Kavanaugh, A., Carroll, J. M., Rosson, M. B., Reese, D. D., & Zin, T. T. (2005). Participating in civil society: the case of networked communities. *Interacting with Computers*, 17(1), 9-33.
- Kelly, J., Fisher, D., & Smith, M. (2009). Friends, Foes, and Fringe: Norms and Structure in Political Discussion Networks. In T. Davies and S. Peña Gangadharan (eds.), *Online Deliberation: Design, Research, and Practice* pp. 83-93: CSLI Publications.
- Kitchin, H. A. (2002). Alcoholics Anonymous discourse and members' resistance in a virtual community: exploring tensions between theory and practice. *Contemporary Drug Problems*, 29(4), 749.

- Kling, R., & Courtright, C. (2003). Group Behavior and Learning in Electronic Forums: A Sociotechnical Approach. *Information Society*, 19(3), 221.
- Kollock, P., & Smith, M. (1999). Communities in Cyberspace. In M. Smith & P. Kollock (Eds.), *Communities in Cyberspace* pp. 3-26: Routledge.
- Latour, B. (2010). Tarde's idea of quantification. In M. Candea (Ed.), *The Social After Gabriel Tarde: Debates and Assessments* pp. 145-162). London: Routledge.
- Lave, J., & Wenger, E. (1991). *Situated Learning. Legitimate peripheral participation.* New York: Cambridge University Press.
- Liu, C.-C., & Tsai, C.-C. (2008). An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. *Comput. Educ.*, 50(3), 627-639.
- Loader, B. D., Muncer, S., Burrows, R., Pleace, N., & Nettleton, S. (2002). Medicine on the line? Computer-mediated social support and advice for people with diabetes. *International Journal of Social Welfare*, 11(1), 53-65.
- Mahoney, D. F. (1998). A content analysis of an Alzheimer family caregivers virtual focus group, *American Journal of Alzheimer's Disease and Other Dementias*, 13(6), 309-316.
- Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Trans. Comput.-Hum. Interact.*, 12(2), 201-232.
- Marra, R., Moore, J., & Klimczak, A. (2004). Content Analysis of Online Discussion Forums: A Comparative Analysis of Protocols. *Educational Technology Research and Development*, 52(2), 23-40.
- Matzat, U. (2004). Academic communication and Internet Discussion Groups: transfer of information or creation of social contacts? *Social Networks*, 26(3), 221-255.
- McKenna, K., Green, A., & Gleason, M. (2002). Relationship formation on the Internet: What's the big attraction? *Journal of Social Issues*, 58(1), 9-31.
- Mondada, L. (1999). Formes de séquentialité dans les courriels et les forums de discussion. Une approche conversationnelle de l'inteaction sur internet. *Apprentissage des langues et système d'information et de communication*, 2(1), 3-25.
- Monnoyer-Smith, L. (2006). Etre créatif sous la contrainte. *Une analyse des formes nouvelles de la délibération publique. Le cas DUCSAI*, 75(3), 75-101.
- Nonnecke, B., & Preece, J. (2000). Lurker demographics: counting the silent. CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems. The Hague, The Netherlands: ACM. 73-80.
- Papadakis, M. C. (2003). *People Can Create a Sense of Community in Cyberspace.* Arlington, VA: SRI International (International's Science and Technology Policy Program).

- Paravel, V., & Rosental, C. (2003). Les reseaux, des objets relationnels non identifiés ? Le cas de la communication électronique dans la recherche. *Réseaux*, 118, 237-270.
- Pfaffenberger, B. (1996). "If I Want It, It's OK": Usenet and the (Outer) Limits of Free Speech. *The Information society*, 12(4), 365-386.
- Preece, J., Nonnecke, B., & Andrews, D. (2004). The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior. The Compass of Human-Computer Interaction*, 20(2), 201-223.
- Proulx, S. (2006). Les communautés virtuelles: ce qui fait lien. In S. Proulx, L. Poissant & M. Sénécal (Eds.), *Communautés virtuelles. Penser et agir en réseau*. pp. 13-26): Presses de l'université Laval.
- Raban, D. R., & Rabin, E. (2009). Statistical inference from power law distributed web-based social interactions. *Internet Research*, 19(3), 266-278.
- Radin, P. (2006). "To me, it's my life" : Medical communication, trust, and activism in cyberspace. *Social Science & Medicine*, 62, 591-601.
- Ridings, C. M., & Gefen, D. (2004). Virtual Community Attraction: Why People Hang Out Online. *Journal of Computer-Mediated Communication*, 10(1), 00-00.
- Roth, C., & Cointet, J.-P. (2010). Social and semantic coevolution in knowledge networks. *Social Networks*, 32(1), 16-29.
- Sack, W. (2002). Conversation Map: A Content-based Usenet Newsgroup Browser. In C. Lueg & D. Fisher (Eds.), *From Usenet to CoWebs: Interacting with Social Information Spaces* pp. 92-109). New York: Springer.
- Sandaunet, A.-G. (2008). The challenge of fitting in: non-participation and withdrawal from an online self-help group for breast cancer patients. *Sociology of Health & Illness*, 30(1), 131-144.
- Shirky, C. (2003). Power Laws, Weblogs, and Inequality. Clay Shirky's Writings About the Internet. Economics & Culture, Media & Community, Open Source.
- Shumar, W., & Smith, M. (2004). Using Netscan To Study Identity And Interaction In A Virtual Community. Annual meeting of the American Sociological Association. Hilton San Francisco.
- Smith, C. B., McLaughlin, M., & Osborne, K. K. (1998). From Terminal Ineptitude to Virtual Sociopathy: How Conduct Is Regulated on Usenet. In M. McLaughlin & S. Rafaeli (Eds.), *Network and Netplay: Virtual Groups on the Internet* pp. 95-112). Menlo Park, CA: American Association for Artificial Intelligence Press.
- Smith, M. (1999). Invisible Crowds in Cyberspace: Mapping the Social Structure of the Usenet. *Communities in Cyberspace: Perspectives on New Forms of Social Organization*: Routledge Press.
- Smith, M. A., & Fiore, A. T. (2001). Visualization components for persistent conversations. Proceedings of the SIGCHI conference on Human factors in computing systems. Seattle, Washington, United States: ACM.

- Stromer-Galley, J. (2007). Measuring Deliberation's Content: A Coding Scheme. *Journal of Public Deliberation*, 3(1), <http://services.bepress.com/jpd/vol3/iss1/art12>.
- Tarde, G. (1993 [1890]). *Les lois de l'imitation*: Editions Kimé.
- Tarde, G. (2005 [1898]). *Les lois sociales. Esquisse d'une sociologie: Les classiques des sciences sociales*. Edition numérique. http://classiques.uqac.ca/classiques/tarde_gabriel/les_lois_sociales/les_lois_sociales.pdf
- Turner, J. W., Grube, J. A., & Meyers, J. (2001). Developing an optimal match within online communities: an exploration of CMC support communities and traditional support. *Journal of Communication*, 51(2), 231-251.
- Turner, T. C., & Fisher, K. E. (2006). The Impact of Social Types within Information Communities: Findings from Technical Newsgroups. 6, 135b-135b.
- Turner, T. C., Smith, M. A., Fisher, D., & Welser, H. T. (2005). Picturing Usenet: Mapping computer-mediated collective action. *Journal of Computer-Mediated Communication*, 10(4).
- Viegas, F. B., Golder, S., & Donath, J. (2006). Visualizing email content: portraying relationships from conversational histories. Proceedings of the SIGCHI conference on Human Factors in computing systems. Montréal, Québec, Canada: ACM.
- Viegas, F. B., & Smith, M. (2004). Newsgroup Crowds and AuthorLines: visualizing the activity of individuals in conversational cyberspaces. System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on. 10 pp.
- Waseleski, C. (2006). Gender and the Use of Exclamation Points in Computer-Mediated Communication: An Analysis of Exclamations Posted to Two Electronic Discussion Lists. *Journal of Computer-Mediated Communication*, 11(4).
- Wellman, B., & Gulia, M. (1999). Net surfers don't ride alone: virtual communities as communities. In B. Wellman (Ed.), *Networks in the Global Village*. Boulder, CO: Westview.
- Welser, H., Gleave, E., Fisher, D., & Smith, M. (2007). Visualizing the Signatures of Social Roles in Online Discussion Groups. *Journal of Social Structure*, 8(2).
- Welser, H. T., Gleave, E., Barash, V., Smith, M., & Meckes, J. (2009). Whither the Experts? Social Affordances and the Cultivation of Experts in Community Q&A Systems. Computational Science and Engineering, 2009. CSE '09. International Conference on. 4, 450-455.
- Wesemann, D., & Grunwald, M. (2008). Online discussion groups for bulimia nervosa: An inductive approach to Internet-based communication between patients. *International Journal of Eating Disorders*, 41(6), 527-534.
- Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). The dynamics of mass interaction. Proceedings of the 1998 ACM conference on Computer supported cooperative work. ACM Press New York, NY, USA. 257-264.

- Wikgren, M. (2001). Health discussions on the Internet: A study of knowledge communication through citations. *Library & Information Science Research*, 23(4), 305-317.
- Wilson, S. M., & Peterson, L. C. (2002). The Anthropology of Online Communities. *Annual Review of Anthropology*, 31(1), 449-467.
- Wright, S., & Street, J. (2007). Democracy, deliberation and design: the case of online discussion forums. *New Media & Society*, 9(5), 849-869.
- Zhang, J., Ackerman, M. S., & Adamic, L. (2007). Expertise networks in online communities: structure and algorithms. Proceedings of the 16th international conference on World Wide Web. Banff, Alberta, Canada: ACM.
- Ziebland, S., Chapple, A., Dumelow, C., Evans, J., Prinjha, S., & Rozmovits, L. (2004). How the internet affects patients' experience of cancer: a qualitative study. *BMJ*, 328(564).
- Zrebiec, J. F. (2005). Internet Communities: Do They Improve Coping With Diabetes? *The Diabetes Educator*, 31(6), 825-836.