



HAL
open science

Impact de la transition numérique sur la production de ressentis

Pauline Soutrenon

► **To cite this version:**

Pauline Soutrenon. Impact de la transition numérique sur la production de ressentis. Colloque international des Etudiant×e×s chercheur×se×s en DIIdactique des langues et Linguistique, CEDIL'18, May 2018, Grenoble, France. pp.123-128. hal-02648851

HAL Id: hal-02648851

<https://hal.univ-grenoble-alpes.fr/hal-02648851>

Submitted on 9 Jun 2020

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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

IMPACT DE LA TRANSITION NUMERIQUE SUR LA PRODUCTION DE RESSENTIS

Pauline SOUTRENON^{a, b}

psoutrenon@comongo.fr

^aSAS Comongo, Grenoble, France

^bUMR Litt&Arts (5316), Grenoble, France

1. Introduction

Dans le cadre d'une thèse CIFRE avec la startup Comongo, nous présentons un travail de recherche en sciences du langage et humanités numériques sur l'analyse automatique de production de ressentis. Nous avons deux objectifs : définir et évaluer l'image de marque d'un client et l'aider à anticiper les ressentis et les attentes de communication de publics cibles en vue d'améliorer cette image de marque. Dans notre démarche, nous employons le terme d'analyse de ressentis qui se distingue de l'analyse de sentiments ou d'opinions. Nous travaillons sur des productions écrites (des réponses à des questions ouvertes) qui nous permettent de comparer une identité souhaitée d'un client avec une image perçue. Nous cherchons à extraire les ressentis et à effectuer une représentation synthétique de ceux-ci.

Pour cela, nous bénéficions d'un cadre industriel au sein de la startup Comongo qui souhaite effectuer une transition numérique. Un des aspects qui nous intéresse particulièrement est de prendre en compte et d'analyser, dans les pratiques de Comongo, l'impact de la transition numérique sur la production de ressentis par des groupes de personnes.

Nous posons l'hypothèse que la transition à une pratique numérique, en présentiel comme en distanciel, modifie les contenus rédigés par les participants. Identifier et caractériser ces changements est par conséquent essentiel au bon fonctionnement de l'entreprise et au bon usage des outils de traitement de langues qui accompagnent l'analyse humaine des contenus.

2. Constitution des données

La startup Comongo a mis en place une méthode permettant d'évaluer et d'améliorer l'image de marque d'une personne physique ou morale. Cette méthode se déroule en quatre étapes comme le montre la Figure 1 ci-dessous. Tout d'abord, un sujet est défini par un client et une identité souhaitée est éventuellement constituée avec celui-ci. Dans un deuxième temps, le client définit un panel de publics cibles à interroger. Puis, les réponses sont recueillies via une interface numérique et ces réponses constituent l'image perçue (étape 3). Enfin, la dernière étape est la comparaison entre l'identité souhaitée et l'image perçue. Cette comparaison de texte fait ressortir les lignes convergentes et divergentes pour ainsi émettre des préconisations en termes d'axes de communication avec un accompagnement humain de la démarche du client dans l'amélioration et l'évaluation de son image de marque.



Figure 12 - La méthode de Comongo

Les données recueillies lors de la troisième étape sont des réponses à des questions orientées de différents types :

- Les évocations : ce que le sujet évoque ;
- Les polarités positives et négatives : ce qui est apprécié ou non vis-à-vis du sujet ;
- L'unicité : ce qui rend unique le sujet ;
- Les attentes et les manques : ce qui est attendu.

2.1. Formes d'énoncés

Lorsqu'ils répondent sur l'interface numérique, nous demandons aux audités de remplir une idée par champ. Une idée est une unité d'information telle que conçue par les personnes remplissant le questionnaire, ce qui implique une variabilité de la forme et de la structure. Des exemples sont donnés dans le Tableau 1 ci-dessous, il s'agit de réponses au sujet d'une ville. Chaque cellule correspond à une idée qui a été donnée par une personne. Les données recueillies peuvent être de formes différentes : nous pouvons trouver des mots, des groupes de mots ou encore des phrases.

Mots	Groupes de mots	Phrases
Montagne	Pistes cyclables	Une obligation de travail mais pas une envie de rester y vivre
Sport	Une ville à taille humaine	Il manque des entreprises de taille moyenne
Dynamisme	Un espace de rencontre	
Pollution	La proximité de la montagne	
Chaleur	Un cadre de vie idéal pour les sportifs	

Tableau 9 - Les données recueillies

2.2. Volumétrie

Nous constatons que la volumétrie est variable selon les études. En effet, selon la personne répondant au questionnaire et ses motivations vis-à-vis du sujet, elle peut être plus ou moins productive en termes de nombre d'idées.

Le Tableau 2 ci-dessous montre la volumétrie des études. Nous pouvons constater que les études réalisées en présentiel numérique (études β , γ et δ) ont tendance à générer un nombre moyen d'idées par audité plus faible que l'étude α réalisée en présentiel papier.

Etude	Nombre d'audités	Nombre d'idées	Nombre moyen d'idées par audité	Pratique
Etude α	29	693	23,9	Présentiel papier
Etude β	17	219	12,9	Présentiel numérique
Etude γ	13	260	20,0	Présentiel numérique
Etude δ	28	516	18,4	Présentiel numérique

Tableau 10 - Volumétrie des études

Quelque soit le nombre d'idées par étude, le but est de savoir si les mêmes idées sont répétées chez les différents audités. Pour cela, nous souhaitons automatiser partiellement le processus d'analyse en utilisant des ressources fondées sur l'intelligence humaine et permettant de repérer :

- Les mots dans les énoncés ayant des lemmes communs avec le Lexique des Formes Fléchies du Français (Clément et *al.*, 2004)
- Les idées qui appartiennent au même concept sémantique ou à concept similaire avec Diko (Lafourcade, 2007)
- La polarité, afin d'identifier si une idée est plutôt positive ou négative : LikeIt (Lafourcade et *al.*, 2015) et French Expanded Emotion Lexicon (Abdaoui et *al.*, 2016)

Cette analyse automatique est ensuite vérifiée et corrigée si besoin par le consultant. Il s'agit d'une approche incrémentale des ressources (voir Figure 2 ci-dessous) : les ressources

permettent l'annotation du corpus, chaque idée est associée à un concept. Une fois le corpus analysé, il nous sert à améliorer les ressources. Cet enrichissement est réalisé par le consultant qui peut corriger l'annotation. Ainsi, chaque nouvelle étude est enrichie de la précédente.

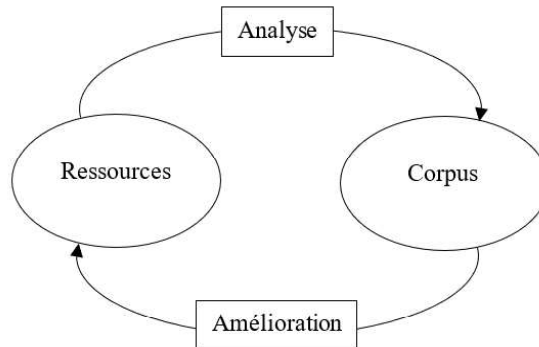


Figure 13 - Enrichissement des ressources

3. Transition numérique et distancielle

Nous sommes partis d'une pratique traditionnelle de Comongo d'évaluation et d'amélioration d'image de marque en présentiel papier. Le consultant Comongo aide le client dans la rédaction de son identité souhaitée et il accompagne les audités lors d'entretiens individuels ou en *focus groups* (petits groupes de 6 à 10 personnes) pour établir un diagnostic d'image perçue. Pendant ces entretiens, le consultant recueille des données qualitatives écrites sur papier par les audités. La comparaison entre l'identité souhaitée et l'image perçue est ensuite effectuée manuellement par le consultant.

Afin d'effectuer la transition numérique visée par l'entreprise, où l'utilisateur se retrouvera seul face à l'outil (pratique distancielle numérique), nous passons par une pratique transitionnelle en présentiel numérique. Au lieu de répondre sur papier, les audités saisissent leurs réponses directement sur l'interface numérique de recueil d'opinions et le consultant est présent pour accompagner la démarche. Les réponses peuvent être revues et modifiées si besoin par les audités avant de les envoyer, ce qui permet de développer chaque idée plus facilement que sur papier.

La mise en place de la pratique transitionnelle nous permet de comparer deux états de la donnée (présentiel papier et présentiel numérique) et d'identifier les impacts du numérique sur la méthode et les données. Lors du passage en distancielle, nous évaluerons les changements dus à l'absence du consultant (passage en distancielle). La comparaison des trois pratiques est présentée sur la Figure 3 ci-dessous. Les changements non exhaustifs entre les pratiques ont été mis en couleurs.

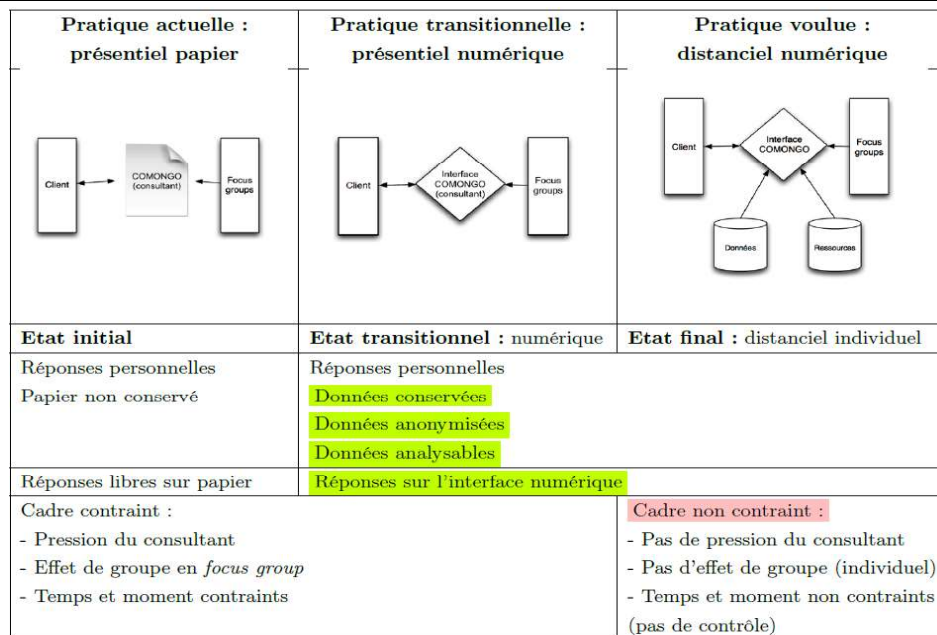


Figure 14 - Comparaison des trois pratiques

Les deux changements majeurs de la pratique actuelle à la pratique transitionnelle apparaissent en vert. Premièrement, le changement de format par le biais de la plateforme numérique risque d'avoir un impact sur la qualité des réponses. La lecture demande de l'attention (Lachaux, 2011) et, sur la plateforme, les audités sont amenés à lire un tutoriel et les questions auxquelles ils doivent répondre. Carr (2012) évoque un problème de lecture avec l'utilisation d'Internet : plus Internet est utilisé et plus il est difficile de se concentrer sur une seule chose et d'être attentifs sur de longs textes. L'attention est une ressource rare, cela est notamment dû à la surabondance informationnelle à laquelle nous sommes confrontés chaque jour (Citton, 2014). Il existe une quantité importante de questionnaires en ligne, les audités seront donc peut-être moins attentifs et cela aura un impact sur la donnée qui risque d'être moins qualitative. De plus, les usagers de la plateforme n'auront pas le même rapport aux outils numériques, il faut donc aussi considérer la question de la littératie numérique. Selon l'OCDE (2000), la littératie est :

« l'aptitude à comprendre et à utiliser l'information écrite dans la vie courante, à la maison, au travail et dans la collectivité en vue d'atteindre des buts personnels et d'étendre ses connaissances et ses capacités ».

La littératie numérique peut donc être définie comme la compétence de chaque usager à utiliser un outil numérique.

Deuxièmement, le fait que les réponses soient enregistrées dans une base de données les inscrit de manière plus ou moins pérenne dans un système informatique. Il ne s'agit donc plus d'une donnée volatile, ce changement de mode de communication peut être la cause de freins psychologiques chez certains usagers. A contrario, le « bouclier » que représente l'écran peut permettre à d'autres de s'exprimer plus librement, sans crainte d'être jugé par le groupe.

En rose, dans la Figure 14 ci-dessus, apparaissent les changements majeurs de la pratique transitionnelle vers la pratique voulue, à savoir : le cadre de la production qui n'est plus contraint, il n'y aura plus de pression du consultant ni d'effet de groupe et nous n'aurons plus

de contrôle sur les conditions dans lesquelles les audités répondront au questionnaire (le temps passé à répondre et le moment pendant lequel ils répondent).

4. Impact du numérique

Les premières observations de notre corpus par rapport à l'impact du numérique montrent qu'il existe des variations dues au cadre technologique.

La transition numérique pouvant avoir un impact sur les réponses des audités, une des hypothèses de départ était que nous risquions d'avoir moins de données en numérique que sur papier. En effet, dans la pratique transitionnelle de Comongo les audités notent leurs réponses directement dans l'outil qui les enregistre dans une base de données. Cela peut constituer un frein psychologique pour les audités qui pourraient s'autocensurer.

Cependant, nous avons constaté que le fait de répondre sur une interface numérique sans le biais du consultant rend l'expression plus libre. Bien que l'étude α réalisée en présentiel papier comporte le plus grand nombre d'idées et de tokens par rapport aux trois études réalisées en présentiel numérique (section 1.2). Nous avons remarqué que les audités argumentent plus sur l'interface numérique. Ils ont tendance pour chaque idée à faire des phrases plus longues en termes de nombre de caractères et de nombre de tokens (Tableau 3 ci-dessous) : nous avons pris les productions les plus longues (nombre de caractères) de l'étude α en présentiel papier et de l'étude γ en présentiel numérique.

Etude	Pratique	Nombre de caractères	Nombre de tokens	Production
Etude α	Présentiel papier	130	25	« Me fournir des alertes sur les sites web dont le business model est d'aller chercher des données et de les exploiter à mes dépends »
Etude γ	Présentiel numérique	285	48	« Un avis trop tranché parfois : sur certains projets ou nous avons des impératifs ou choisis tel ou tel format, il faudrait que X comprenne que nous devons avancer avec ce support et pas proposer absolument autre chose (se recentrer davantage sur le besoin client et mieux l'écouter) »

Tableau 11 - Comparaison du nombre de caractères

L'exemple de production dans le Tableau 4 ci-dessous est une réponse d'un audité à une question de type rejet (ce qui n'est pas apprécié vis-à-vis du sujet) pour laquelle il est demandé de renseigner une idée par champ dans l'interface (chaque champ est représenté dans le tableau par une cellule). La première idée répond à la question posée, puis l'audité complète et va plus loin dans la réflexion en donnant « quelques pistes d'amélioration ». Les audités précisent davantage leur pensée sur l'interface que sur papier.

Production
« Rien que je n'apprécie pas
En revanche , quelques pistes d'amélioration... sur deux points essentiels
Premier point : mieux définir les critères d'une stratégie de GO/NO GO sur certains projets. XXX n'a peut-être pas vocation à accepter toutes les missions, qui risqueraient de désorganiser/perturber/mettre trop de pression sur son équipe
Deuxième point : peut-être passer plus de temps à mieux cerner le besoin client pour être sûrs de bien définir une réponse adaptée, chiffrée au plus près »

Tableau 12 - Exemple de production

5. Conclusions et perspectives

Les outils du Traitement Automatique des Langues permettent d'accompagner et de faciliter les tâches du consultant, que ce soit pour le recueil des données, l'analyse de celles-ci ou leur visualisation. Notre système est fondé sur une approche centrée utilisateur : la machine

accompagne dans la réflexion et donne d'autres perspectives sur l'information, le consultant garde la maîtrise de l'interprétation et sous-traite à la machine l'organisation de l'information.

La transition numérique modifie les contenus rédigés par les audités. Nous avons constaté une baisse du nombre d'idées entre l'étude α réalisée en présentiel papier et les études réalisées en présentiel numérique. Il est prévu de réaliser d'autres études papier une fois la pratique distancielle numérique mise en place. Cela nous permettra de vérifier nos hypothèses et d'avoir une meilleure représentativité. Nous avons également remarqué que les réponses recueillies sur l'interface sont plus longues et plus complètes que celles recueillies sur papier en termes de nombre de caractères et de tokens. Dans les travaux à venir, il serait intéressant de mener une analyse génétique sur les études en traçant les ratures, les déplacements ou les insertions de texte en numérique et en papier.

Actuellement, nous commençons à nous diriger vers le distanciel numérique et nous avons déjà remarqué que les audités préfèrent répondre sans la présence du consultant. Nous nous apercevons également que sans la contrainte temps en distanciel, les audités ne vont pas au bout de leur pensée, mais reviennent dans l'outil plus tard pour compléter, ce qui enlève de la spontanéité.

Les altérations des contenus existent, qu'elles soient dues à la transition numérique ou à la transition de la modalité présenteielle à distancielle, il est essentiel d'en tenir compte dans la pratique d'entreprise mais aussi dans l'usage même des corpus. Il est primordial d'identifier les conditions de productions qui vont faire varier les ressources nécessaires aux outils.

Références bibliographiques

- ABDAOUI, Amine, AZE, Jérôme, BRINGAY, Sandra & PONCELET, Pascal (2016). FEEL : a French Expanded Emotion Lexicon. *Language Resources and Evaluation*, (pp. 1–23).
- CARR, Nicolas (2012). *Internet nous rend-il bête ? Réapprendre à lire et à penser dans un monde fragmenté*. Paris : Robert Laffont.
- CLÉMENT, Lionel, LANG, Bernard, & SAGOT, Benoît (2004). Morphology based automatic acquisition of large-coverage lexica. In *LREC 04* (pp. 1841–1844). Lisbonne, Portugal.
- CITTON, Yves (2014). *Introduction à L'économie de l'attention. Nouvel horizon du capitalisme ?* Paris : La découverte.
- LACHAUX, Jean-Philippe (2011). *Le Cerveau attentif*. Paris : Odile Jacob.
- LAFOURCADE, Mathieu (2007). Making people play for Lexical Acquisition with the JeuxDeMots prototype. In *SNLP'07 : 7th International Symposium on Natural Language Processing* (pp.7). Pattaya, Chonburi, Thailand.
- LAFOURCADE, Mathieu, LE BRUN, Nathalie & JOUBERT, Alain (2015). Vous aimez ?...ou pas ? LikeIt, un jeu pour construire une ressource lexicale de polarité. In *TALN : Traitement Automatique des Langues Naturelles*. Caen, France.
- OCDE (2000). *La littératie à l'ère du numérique*. Technical report, Organisation de coopération et de Développement Numérique.