

Research Project

Epidemiology of Representations: An Empirical Approach

Sébastien Lerique*

Advisor: Jean-Pierre Nadal[†]

Co-advisor: Camille Roth[‡]

1 Introduction

TODO: Faire attention à attacher aussi un peu dans le champ de l'anthropologie culturelle, pas que psycho-cog (pas du temps perdu car utile pour l'état de l'art de la thèse) si jamais les Gutachterinnen ont l'heur d'être SHS

Recent years have seen several attempts to bring cognitive science and social science together. “Social cognition”, “cognitive economy” and “cultural evolution” are examples of such fields having recently emerged, approached from multiple viewpoints and involving a variety of disciplines. The project we put forward focuses on a subject formalised by Dan Sperber in the mid-nineties: in a series of innovative articles gathered in Sperber (1996), the author suggests a research program called *cultural epidemiology*, which aims to provide the cognitive and social sciences with a common framework to address a long-standing anthropological question: how can societies be so diverse when the people that form them are all made of the same psychological substrate? The framework Sperber suggests starts from an ontology made only of “mental representations” (those from cognitive science) and their expressions in the outer world, “public representations”. The core principle is then to characterise societies by the distribution of public representations that circulate in them, and combine knowledge from cognitive science and anthropology to explain their evolution. As Sperber argues, this naturalistic approach builds on cognitive principles while at the same time allowing anthropological questions to be rephrased in terms of epidemiology of representations. For instance: what types of representations are only weakly transformed as they are interpreted and produced anew by successive people? Those representations, spreading wider than the others, become *cultural*. Are they attractors for the interpretation-reproduction process of representations? If so, which cognitive modules are involved in the stability of such representations?

This research program provides a framework to explore the interactions between evolution and culture (as exemplified in Baumann, André, & Sperber, 2013), and has recently seen several theoretical developments with genuine mathematical models (see for instance Claidière & Sperber, 2007). A number of other works also explore these questions with empirical approaches. Artificial transmission

*Centre d'Analyse et de Mathématique Sociales (CAMS, UMR 8557, CNRS-EHESS, Paris). Email: sebastien.lerique@normalesup.org.

[†]CAMS, and Laboratoire de Physique Statistique (LPS, UMR 8550, CNRS-ENS-UPMC-Univ. Paris Diderot, Paris). Email: nadal@lps.ens.fr

[‡]CAMS, Centre Marc Bloch (CMB, UMIFRE 14, CNRS-MAEE-HU, Berlin), and Sciences Po Paris. Email: roth@ehess.fr

chains in the laboratory, for instance, have been used extensively to study iterated language evolution, as Tamariz & Kirby (2016) review. That paradigm has also been used to study the evolution of short audio loops (MacCallum, Mauch, Burt, & Leroi, 2012), of risk perception (Moussaïd, Brighton, & Gaissmaier, 2015), and of abstract visual patterns transmitted by apes (Claidière, Smith, Kirby, & Fagot, 2014). Another approach is the meta-analysis of a large number of anthropological or historical works on a given subject to rebuild the evolution of a representation as it happened. This technique has been used by Morin (2013) in his study of how painted portraits change over the centuries, and by Miton, Claidière, & Mercier (2015) in their examination of the practice of bloodletting. A third approach crawls online content and communities in order to study the spread and propagation of representations in human networks. While earlier works were based on atomic propagation and exposure models where simple entities such as URLs and innovations were the central object (see J.-P. Cointet & Roth, 2007 for a discussion), this stream of research is increasingly modelling representations as deep objects with complexity of their own, improving on simpler virus-like models. Several works have now studied large quantities of meaningful sentences, showing that their propagation depends heavily on context (Bakshy, Karrer, & Adamic, 2009) and semantic content (Danescu-Niculescu-Mizil, Cheng, Kleinberg, & Lee, 2012), as well as on competition between items (Weng, Flammini, Vespignani, & Menczer, 2012).

The wide array of disciplines studying these complimentary questions, and the variety of techniques used in the process testify to a major obstacle: collecting relevant data in usable amounts to analyse cultural evolution is not easy. The works cited above develop several strategies to face the problem, yet invariably leave core aspects of the question aside: transmission chains operate on extremely simple representations; recompiling historical and anthropological works uncovers trends with many explanations competing for causality; models of online content propagation overlook cognitive levels of explanation by and large. It is possible, however, to combine the advantages of these techniques into new methods that significantly expand what empirical studies can tackle. By applying the tools of psycholinguistics to the study of online communities on one side, and enabling transmission chains to benefit from widespread computing power and internet infrastructure on the other side, we are able to collect massive amounts of usable data for the empirical and quantitative study of out-of-laboratory epidemiology of representations.

2 Scientific Goals

This research project, which, with the support of the DAAD, I will be able to conclude within my PhD, develops empirical approaches for the study of cultural evolution with the following two broad goals:

1. Dramatically increase the amounts of data collected in order to improve experiments' resolving power, and
2. Diversify data sources to open the study of out-of-laboratory quantitative questions on cultural evolution.

We accomplish this by bringing data mining and psycholinguistics together through the possibilities of modern browsers and computing platforms. A first approach builds on the vast amounts of digital traces available online: by applying data mining techniques to filter and extract relevant information from large data sets, it becomes possible to ask quantitative cognitive-cultural questions to traces of activity from everyday lives. Our second approach connects fields the other way around: by using the potential and ubiquity of modern browsers, we lift a number of constraints on cognitive science and cultural evolution experiments in the laboratory. In particular, we develop transmission chain experiments as real-time web applications in the browser. This lets us work on ecological content with a much shorter transmission cycle and a diversified and easily scalable subject pool, while still keeping full control over the complexity of cognitive tasks implemented. Alongside this empirical

work, we review the main critiques of the *epidemiology of representations* program, both from social anthropology and non-classical streams of cognitive science, which frame the questions of cultural evolution differently and pinpoint debated assumptions in need of further work.

Upon completion, this project will have shown a path to analyse the evolution of real linguistic content, and the links it has with overall cultural change, with a set of workable constraints compared to previous techniques. By fruitfully combining social science, psycholinguistics, and large scale modelling, it significantly widens the field of empirical possibilities and allows the various disciplines involved to address new and debated questions on the links between cognition and cultural evolution. As Greenwald (2012) has argued, “there is nothing so theoretical as a good method”.

3 Project Breakdown

TODO: Privilégier une présentation séquentielle : les fonds français m’ont permis de réaliser toutes ces étapes jusqu’ici, et d’en arriver à me dire que l’Allemagne pourrait mettre au pot pour finaliser cette dernière partie dont je vous parle là

3.1 Transformations of quotations in blogspace

On se propose de contribuer à l’exploration empirique de l’épidémiologie des représentations en partant d’un cas particulier : les citations (au sens de *quotations*). On choisit délibérément de se restreindre à ce cas particulier le plus simple possible afin de s’assurer que l’étude sera réalisable avec les moyens et les connaissances actuels, et pour pouvoir profiter du cadre intéressant offert par les citations. En effet :

- Alors que les citations ne devraient subir aucune transformation quand elles sont reproduites, l’observation empirique montre le contraire : il est très fréquent qu’une citation soit transformée (par rognage, ou par remplacement, insertion ou délétion de mots), ou modifiée grammaticalement. C’est bien connu à l’oral dans des tâches de rappel libre ou indicé (Potter & Lombardi, 1990) ; voir Simmons, Adamic, & Adar (2011) pour une illustration sur la blogosphère, et Lauf, Valette, & Khouas (2013) pour une typologie des transformations fréquemment observées.
- Les citations, généralement assez courtes, se prêtent bien à l’étude automatisée (Traitement Automatique du Langage), contrairement à d’autres types de représentations qu’on pourrait imaginer considérer.
- On dispose d’une observation de référence qui est proche de la situation qui pourra être simulée en expérience de chaîne de transmission : la dynamique des citations sur les blogs sur Internet (voir Leskovec, Backstrom, & Kleinberg, 2009 ; et Simmons et al., 2011). Cette situation semble unique lorsqu’on compare à d’autres représentations qu’on pourrait envisager d’étudier (comme des techniques, des images ou dessins, etc.).

Un tel projet comporte trois grandes parties complémentaires.

La première consiste à observer hors laboratoire, et analyser autant que les contraintes le permettent, le processus d’interprétation-reproduction par lequel les citations sont transformées. On confirme ainsi l’existence de biais cognitifs dans transmission des citations dans la société quotidienne, et on identifie les liens que ces biais peuvent avoir avec une attraction culturelle telle que proposée par l’épidémiologie des représentations.

La deuxième consiste à reproduire le processus et ces effets dans une série d’expériences de chaînes de transmission pour en examiner les déterminants en situation contrôlée. Cette partie permet d’examiner

les effets accumulés de la cognition sur l'évolution du contenu ainsi que les rétroactions que cette évolution génère sur le processus de transformation lui-même. Elle permet également de mettre en lumière le rôle de la tâche de reproduction dans le type d'évolution à long terme qu'on observe. Ce travail crée les conditions pour une modélisation réelle du processus dans des travaux futurs, intégrant les principes et rétroactions identifiés dans des simulations à comparer aux phénomènes réels, pour identifier ainsi le rôle des transformations dans l'évolution du paysage global.

La troisième étape consiste à faire l'examen critique des apports et des limites de ces approches, notamment dans le cadre du débat sur les façons de lier sciences sociales et cognition, dans lequel de nombreuses disciplines proposent des points de vue différents.

Par l'introduction de nouvelles méthodes, ce projet permet de lever progressivement certaines simplifications imposées par les outils actuels dans l'étude de l'évolution culturelle (voir Greenwald, 2012, pour plus de détails sur l'intérêt de développer de nouvelles méthodes) : on étudie du contenu plus complexe, dans des situations souvent plus écologiques. La suite de ce document expose le travail accompli dans chacune des trois grandes parties, puis détaille le travail restant pour compléter le projet.

Le premier volet du projet est l'exploration de certaines transformations que subissent les citations dans la blogosphère. De nombreux travaux ont déjà exploré l'évolution culturelle dans d'autres cas ; on peut citer la *Théorie de l'Hérédité Duale* (ou *Dual Inheritance Theory* en anglais) introduite par Boyd & Richerson (1985) et Cavalli-Sforza & Feldman (1981), les modélisations et travaux plus récents étudiant la notion d'attracteur culturel (voir par exemple Claidière & Sperber, 2007 ; ou Claidière, Scott-Phillips, & Sperber, 2014), les expériences de chaînes de transmission dont l'usage est discuté par Mesoudi & Whiten (2008) et Tamariz & Kirby (2016), ou encore les compilations de travaux anthropologiques et historiques telles que l'ont fait Morin (2013) et Miton et al. (2015). À notre connaissance cependant, aucun travail de recherche n'a examiné l'évolution de contenus linguistique hors du laboratoire.

Étant donné que les citations sur la blogosphère sont graduellement transformées, au fur et à mesure de leur propagation, par des biais cognitifs automatiques comme ceux que Sperber propose comme mécanisme pour l'épidémiologie des représentations, on cherche à savoir à quel point la théorie rend compte de l'évolution observée dans le jeu de données MemeTracker récolté par Leskovec et al. (2009). Notamment, il s'agit de tester si des attracteurs culturels apparaissent dans la dynamique des citations sur internet, et dans ce cas quel rôle jouent les biais cognitifs dans leur émergence.

Le travail de recherche mené pour répondre à ces questions est présenté dans l'article "The semantic drift of quotations in blogspace: a case study in short-term cultural evolution", soumis en août 2016 à la revue *Cognitive Science* en révision majeure. Il consiste à adopter une approche de *data mining* comme celle utilisée par de nombreux travaux sur les communautés en ligne (par exemple J. P. Cointet & Roth, 2009 ; Danescu-Niculescu-Mizil et al., 2012 ; ou encore Gruhl, Guha, Liben-Nowell, & Tomkins, 2004) et à la combiner avec les connaissances existantes en psycholinguistique (en particulier celles sur les effets de certaines caractéristiques des mots sur le rappel, par exemple Yonelinas, 2002 ; ou Zevin & Seidenberg, 2002) pour analyser les remplacements d'un mot par un autre dans le jeu de données MemeTracker. En caractérisant les mots remplacés et les mots apparaissant à leur place, on montre que la façon dont les auteurs substituent des mots dans les citations est fortement biaisée, en accord avec les effets connus en psycholinguistique, et compatible avec l'hypothèse d'attracteurs culturels.

Ce travail, qui complète la première partie du projet, montre qu'il est possible d'étudier l'épidémiologie des représentations sur du contenu linguistique complexe hors du laboratoire, en faisant l'équivalent d'une expérience de psycholinguistique sur des données de la blogosphère. Il montre également le coût d'une telle approche qui doit faire face à la complexité de données *in vivo*, que ce soit au travers du filtrage nécessaire de l'information ou de la reconstitution des données manquantes.

3.2 Semi-controlled transmission chains: Gistr experiment

ADD:

- Elements of the Gistr Whitepaper (not limited to quotations),
- Details on the first launch,
- Work remaining, which will be finished by November

3.3 Discussing empirical works in the context of current debates

ADD: Elements of the epistemological paper (shortcomings), and how we show that the previous works open these questions to scrutiny

Social anthropologists and unorthodox streams from cognitive science in particular (extended, embodied, enactive, and embedded cognitive science), debate the notion of *cognitive representations*, and criticize the boundaries it defines between biological, psychological, and cultural [see in particular ingold; as well as cuffari].

Important when you touch the relationship with context, which is what underlies meaning, and this surfaces in current experiments.

4 Planned work

ADD: Remind last steps for Gistr

ADD: Explain an outline for the manuscript

Bibliographie

Bakshy, E., Karrer, B., & Adamic, L. A. (2009). Social Influence and the Diffusion of User-created Content. In *Proceedings of the 10th ACM Conference on Electronic Commerce* (pp. 325–334). New York, NY, USA: ACM. <http://doi.org/10.1145/1566374.1566421>

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, 36(01), 59–78.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, N.J.: Princeton University Press.

Claidière, N., & Sperber, D. (2007). The role of attraction in cultural evolution. *Journal of Cognition and Culture*, 7(1), 89–111. <http://doi.org/10.1163/156853707X171829>

Claidière, N., Scott-Phillips, T. C., & Sperber, D. (2014). How Darwinian is cultural evolution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1642), 20130368. <http://doi.org/10.1098/rstb.2013.0368>

Claidière, N., Smith, K., Kirby, S., & Fagot, J. (2014). Cultural evolution of systematically structured behaviour in a non-human primate. *Proceedings of the Royal Society of London B: Biological Sciences*,

281(1797), 2014:1541. <http://doi.org/10.1098/rspb.2014.1541>

Cointet, J. P., & Roth, C. (2009). Socio-semantic Dynamics in a Blog Network. In *International Conference on Computational Science and Engineering, 2009. CSE '09* (Vol. 4, pp. 114–121). <http://doi.org/10.1109/CSE.2009.105>

Cointet, J.-P., & Roth, C. (2007). How Realistic Should Knowledge Diffusion Models Be? *Journal of Artificial Societies and Social Simulation*, 10(3), 5. Retrieved from <http://jasss.soc.surrey.ac.uk/10/3/5.html>

Danescu-Niculescu-Mizil, C., Cheng, J., Kleinberg, J., & Lee, L. (2012). You had me at hello: How phrasing affects memorability. *ArXiv:1203.6360 [Physics]*. Retrieved from <http://arxiv.org/abs/1203.6360>

Greenwald, A. G. (2012). There Is Nothing So Theoretical as a Good Method. *Perspectives on Psychological Science*, 7(2), 99–108. <http://doi.org/10.1177/1745691611434210>

Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Information Diffusion Through Blog-space. In *Proceedings of the 13th International Conference on World Wide Web* (pp. 491–501). New York, NY, USA: ACM. <http://doi.org/10.1145/988672.988739>

Lauf, A., Valette, M., & Khoulas, L. (2013). Analyzing Variation Patterns In Quotes Over Time. *Research in Computing Science*, 70, 223–232. Retrieved from http://www.micai.org/rscs/2013_70/Analyzing%20Variation%20Patterns%20In%20Quotes%20Over%20Time.html

Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). Meme-tracking and the Dynamics of the News Cycle. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 497–506). New York, NY, USA: ACM. <http://doi.org/10.1145/1557019.1557077>

MacCallum, R. M., Mauch, M., Burt, A., & Leroi, A. M. (2012). Evolution of music by public choice. *Proceedings of the National Academy of Sciences*, 109(30), 12081–12086. <http://doi.org/10.1073/pnas.1203182109>

Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1509), 3489–3501. <http://doi.org/10.1098/rstb.2008.0129>

Miton, H., Claidière, N., & Mercier, H. (2015). Universal cognitive mechanisms explain the cultural success of bloodletting. *Evolution and Human Behavior*, 36(4), 303–312. <http://doi.org/10.1016/j.evolhumbehav.2015.01.003>

Morin, O. (2013). How portraits turned their eyes upon us: Visual preferences and demographic change in cultural evolution. *Evolution and Human Behavior*, 34(3), 222–229. <http://doi.org/10.1016/j.evolhumbehav.2013.01.004>

Moussaïd, M., Brighton, H., & Gaissmaier, W. (2015). The amplification of risk in experimental diffusion chains. *Proceedings of the National Academy of Sciences*, 112(18), 5631–5636. <http://doi.org/10.1073/pnas.1421883112>

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, 29(6), 633–654. [http://doi.org/10.1016/0749-596X\(90\)90042-X](http://doi.org/10.1016/0749-596X(90)90042-X)

Simmons, M. P., Adamic, L. A., & Adar, E. (2011). Memes Online: Extracted, Subtracted, Injected, and Recollected. In *Fifth International AAAI Conference on Weblogs and Social Media*. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2836>

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford, UK; Cambridge, Mass.: Blackwell.

Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, 8,

37–43. <http://doi.org/10.1016/j.copsyc.2015.09.003>

Weng, L., Flammini, A., Vespignani, A., & Menczer, F. (2012). Competition among memes in a world with limited attention. *Scientific Reports*, 2. <http://doi.org/10.1038/srep00335>

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, 46(3), 441–517. <http://doi.org/10.1006/jmla.2002.2864>

Zevin, J. D., & Seidenberg, M. S. (2002). Age of Acquisition Effects in Word Reading and Other Tasks. *Journal of Memory and Language*, 47(1), 1–29. <http://doi.org/10.1006/jmla.2001.2834>