

UNIVERSITÉ PARIS, SCIENCES & LETTRES

Camille Perault

Diplômée d'HEC Paris

Diplômée du Master 2 Sciences Cognitives de l'ENS Ulm - PSL

DETECTER ET QUANTIFIER L'AGENTIVITE DANS LA LANGUE

Une perspective évolutionnaire

Mémoire de deuxième année du master

« Humanités Numériques »

2023

Résumé

Ce projet interroge la possibilité de détecter de façon cohérente et systématique, dans le langage écrit, une des grandes dimensions psychologiques humaines : celle de l'« agentivité ». En s'intéressant aux ouvrages fictionnels, en particulier les pièces de théâtre pour leur forme stylistique relativement standardisée et leurs personnages types, il entend *in fine* faire le pont entre connotation agentive d'un texte et le degré d'agentivité de son auteur ainsi que de ses contemporains, ce dans une perspective évolutionnaire. Etant donné le caractère très exploratoire de ce travail, et l'ambiguïté de la notion-même d'agentivité étudiée ici, la majeure partie de la recherche a consisté, en amont, à définir plus finement la dimension psychologique en question, et, sur cette base, les marqueurs linguistiques théoriquement à même de refléter le niveau d'agentivité de leurs locuteurs. Ceci explique le relatif déséquilibre entre les chapitres théoriques et ceux ayant trait aux résultats, une grande partie des analyses statistiques restant à mener.

Mots-clés : agentivité ; psycholinguistique ; anthropologie évolutive ; théâtre classique ; humanités numériques ; lecture distante ; traitement automatique de la langue

Informations bibliographiques : Camille Perault, *Détecter et quantifier l'agentivité dans la langue : une perspective évolutionnaire*, mémoire de master 2 « Humanités Numériques », dir. [Nicolas Baumard, Jean-Baptiste Camps], Université Paris, Sciences & Lettres, 2023.

Abstract

This project investigates the possibility of consistently and systematically detecting, in written language, one of the great human psychological dimensions : that of "agency". By focusing on fictional works, particularly plays for their relatively standardized stylistic form and their typical characters, it ultimately intends to bridge the gap between the agentive connotation of a text and the degree of agency of its author and his or her contemporaries, this in an evolutionary perspective. Given the highly exploratory nature of this work, and the ambiguity of the very notion of agency studied here, the bulk of the research has consisted, upstream, in defining more precisely the psychological dimension in question, and, on this basis, the linguistic markers theoretically capable of reflecting the level of agency of their speakers. This explains the relative imbalance between the theoretical chapters and those dealing with results, with a large proportion of the statistical analyses still to be carried out.

Keywords : agency ; psycholinguistics ; evolutionary anthropology ; classical theater ; digital humanities ; remote reading ; automatic language processing

Informations bibliographiques : Camille Perault, *Detecting and quantifying agency in language : an evolutionary perspective*, Master 2 thesis « Digital Humanities », dir. [Nicolas Baumard, Jean-Baptiste Camps], Université Paris, Sciences & Lettres, 2023.

Remerciements

Je tiens à remercier chaleureusement mon directeur de mémoire Monsieur Nicolas Baumard, pour son enthousiasme pour le projet, sa grande disponibilité, et son aide précieuse tout au long de sa conduite, à commencer par la transmission d'un corpus de théâtre déjà traité qui m'a permise de m'appliquer plus longuement à la définition du cadre théorique.

Je tiens également à remercier Valentin Thouzeau, Edgar Dubourg et Jeanne Bollee, qui ont activement participé à la réflexion autour de ce projet.

J'aimerais enfin remercier Monsieur Jean-Baptiste Camps pour ses conseils et l'intérêt qu'il a su montrer pour ce travail, notamment lors du séminaire de Master.

Table des matières

Résumé	ii
Abstract	ii
Remerciements	v
Table des matières	vi
Introduction	3
I Agentivité et langage	7
1 Une définition de l’agentivité	9
1.1 <i>The Big Two</i> : Agentivité et Communion	9
1.2 La théorie de l’identification de l’action, ou l’existence d’un continuum agentif	12
2 Le langage comme reflet de la psychologie humaine	15
2.1 Transitivité et <i>foregrounding</i> : des indices corrélés de l’efficacité de l’action	15
2.2 Individualisme et pronoms : même méthode, nouveau raisonnement	18
II Questions de recherche et méthodologie	23
3 L’évolution de nos préférences psychologiques, un paradigme évolutionnaire	25
3.1 Les fondements de la psychologie évolutive	25
3.2 Les fictions narratives, des « produits dérivés » de l’esprit humain	27
4 Construction d’un indice d’agentivité global : une approche par faisceau d’indices	29
4.1 Validation interne : le choix d’un corpus théâtral	29
4.2 Validation externe : plusieurs propositions	30

4.3 Outils utilisés	33
III Détail des composants linguistiques retenus	35
5 Actionnalité, volition et locus de contrôle : une escapade sémantique	37
5.1 Les limites d'une analyse par mot-clé	37
5.2 Les adverbes intentionnels, statut hybride?	40
5.3 La modalité : entre sémantique et pragmatique	40
6 Les proto-marqueurs agentifs	49
6.1 Le verbe, un indicateur ambiguë	49
6.2 Pronoms et déictiques : la référence au soi	49
6.3 Analyse des dépendances syntaxiques : agentivité et voix grammaticales, verbes statifs, transitivité	50
6.3.1 Agentivité grammaticale	50
6.3.2 Voix active vs. passive	52
6.3.3 Verbes statifs	52
6.3.4 Transitivité	53
7 Traverser les distances psychologiques : l'agentivité et la planification	55
7.1 L'analyse aspectuelle de la télélicité	56
7.1.1 Les modes de l'irréel et la pensée hypothétique	63
7.1.2 Les marqueurs corrélés d'abstraction	63
8 Le langage de la personnalité : indices d'extraversion et d'agréabilité	65
IV Résultats et discussion	69
9 Des résultats prometteurs	71
10 Discussion	77
10.1 Limites de la présente recherche	77
10.2 Analyses successives	78
Conclusion	85
Annexes	89
.1 Disponibilité des données du mémoire	89
.2 Annexe psycho-linguistique	89

Table des matières	ix
.3 Résultats détaillés des modèles de régression	90
Bibliographie	93

Introduction

J'entendais tout cela, moi ; car nous autres esclaves, nous sommes doués contre nos maîtres d'une pénétration.

Cléanthis, L'île des esclaves, 1725

Au cours du dernier siècle notamment, s'est cristallisée l'idée que l'on assistait à une montée de l'« individualisme », tendance culturelle censée privilégier la valeur et les droits de l'individu par rapport à ceux de la société. De fait, l'individualisme devient dès la seconde moitié du XIX^e siècle un prisme d'analyse méthodologique à part entière pour la sociologie naissante. De là, le terme semble peu à peu acquérir une connotation négative : il n'est aujourd'hui pas rare qu'il soit apparenté à la dénonciation d'une focalisation excessive sur le soi, un égocentrisme contemporain qui proliférerait au détriment de notre capacité à prêter attention aux autres^{1 2}.

Or, cette notion portée aux nues par la sociologie-même semble ne pouvoir révéler de manière plus limpide les limites des sciences humaines et sociales « traditionnelles » (du moins dans leur penchant holiste) s'agissant de la compréhension de tels mouvements culturels. Comment, de fait, la tendance à donner la priorité aux individus sur le groupe, contre sa propre cohésion, pourrait-elle s'expliquer en référence à ce même groupe ? Pourquoi le capitalisme, et son pendant psychologique que serait l'individualisme, émergent-ils à la fin du XVII^e siècle ? Et pourquoi pas avant ? Ni plus tard ? Comment, en fond, expliquer un fait social par un autre fait social, pour reprendre l'adage durkheimien, et où s'arrête une telle tautologie ?

Tout d'abord, il faut remarquer que l'on trouve la mention d'idées analogues au concept d'individualisme bien avant l'instauration du système capitaliste, ce dès le Moyen-Âge^{3 4 5 6}. Phénomène non plus spécifiquement moderne, qu'est-ce qui explique, dès lors, que certaines périodes historiques, ou bien encore certaines sociétés, puissent être considérées comme plus individualistes que d'autres ?

L'anthropologie évolutive s'attache à étudier la diversité des comportements humains dans le contexte de leur histoire évolutive. En particulier, la psychologie « évolutionniste », « évolutive » ou « évolutionnaire », s'intéresse à nos mécanismes de pensée, appréhendés à partir de la théorie de l'évolution biologique. Elle fait de notre environnement et de ses implications pour le fitness de chaque individu, à savoir son succès reproductif, un facteur explicatif central de notre psyché. Sur cette base, elle propose ainsi d'expliquer, ultimement, les modifications des préférences psychologiques humaines qui semblent in-

1. Jean M. Twenge, *Generation Me : Why today's young Americans are more confident, assertive, entitled—and more miserable than ever before*. Pages : vii, 292, New York, NY, US, 2006 (Generation Me : Why today's young Americans are more confident, assertive, entitled—and more miserable than ever before.)

2. *The Century of the Self*, Documentaire BBC : <https://www.youtube.com/watch?v=jymMjNc0igI>

3. Jean-Charles Payen, *L'humanisme médiéval et la redécouverte de l'individu en Occident du XIII^e siècle à la fin du XIII^e*, 1985

4. Anthony Giddens, *Modernity and Self-Identity : Self and Society in the Late Modern Age*, Stanford, 1991, URL : <http://www.sup.org/books/title/?id=2660> (visité le 09/09/2023)

5. Aron-J. Gourevitch, *La Naissance de l'individu dans l'Europe médiévale*, 1997

6. {and} C. S. Heppleston. Rosenwein Barbara H., "Y avait-il un « moi » au haut Moyen Âge ?", 2005, URL : <http://www.jstor.org/stable/40957516>.

tervenir au fil de notre Histoire par des variations du niveau de ressources offertes par notre milieu.

Cette précedence périodiquement donnée à l'individu sur le groupe, dans certaines sociétés, pourrait donc être réductible à des considérations matérielles, entendues au sens large (en incluant par exemple l'existence d'institutions). Encore faut-il définir ce qu'est l'individualisme du point de vue de la psychologie individuelle : quel(s) mécanisme(s) de pensée, précisément, pourrai(ent) motiver ces comportements que l'on qualifie synthétiquement d' « individualistes » ? Il semble que la notion d' « agentivité » soit la plus à même de remplir ce rôle. Quoique définie de façon disparate, cette dimension psychologique fait globalement référence à la capacité d'un individu, réelle ou telle qu'elle est perçue par ce dernier, à agir sur le monde qui l'entoure, ce dans le but de poursuivre des objectifs propres.

En dehors de toute considération hiérarchique, Vallacher⁷ développe notamment l'idée que l'on pourrait placer chaque individu le long d'un continuum agentif, selon qu'il envisage le plus souvent ses actions à l'aune de leurs détails mécaniques (faiblement agentif), ou plutôt de significations plus abstraites (hautement agentif).

Pour mesurer un tel degré d'agentivité, l'auteur fait appel au langage : dans son étude, les participants doivent choisir, parmi deux formulations, celle qui décrit le mieux, selon eux, l'activité en question. A titre d'exemple, l'action « nettoyer la maison » (*cleaning the house* en anglais) est alternativement décrite comme le fait de « faire état de sa propre hygiène » (*showing one's cleanliness*) ou bien simplement de « passer l'aspirateur » (*vaccuming the floor*). Choisir la première formulation est censée dénoter relativement plus d'agentivité qu'opter pour la seconde.

Là où l'auteur ne s'aventure pas, c'est dans la possibilité de systématiser une telle mesure. L'idée que l'utilisation d'un lexique spécifique puisse être corrélée à la possession de certains traits psychologiques n'est pas nouvelle (voir la section I.2.2 *Individualisme et pronoms : même méthode, nouveau raisonnement*), y compris s'agissant d'agentivité. Néanmoins, elle n'a que très rarement été étendue, outre le choix d'un lexique spécifique, à la mobilisation de structures grammaticales ou syntaxiques particulières, qui pourraient s'avérer préférentiellement associées à un état d'esprit plus agentif.

Ce mémoire s'attache donc à déterminer si l'élaboration d'un indice global d'agentivité, construit par agrégation de composantes linguistiques diverses, aussi bien lexicales que « structurelles », est envisageable. Si tel est le cas, l'objectif ultime de cette recherche consistera à mesurer le « score » agentif obtenu par différentes œuvres publiées dans une société donnée au fil du temps, afin de tester l'hypothèse selon laquelle ce score corréle avec les variations du niveau de ressources de la société en question.

7. Robin R. Vallacher et Daniel M. Wegner, « Levels of personal agency : Individual variation in action identification. » *Journal of Personality and Social Psychology*, 57-4 (1989), Place : US Publisher : American Psychological Association, p. 660-671, DOI : [10.1037/0022-3514.57.4.660](https://doi.org/10.1037/0022-3514.57.4.660)

Pour ce faire, nous avons opté pour l'analyse de textes littéraires. Ceux-ci présentaient l'avantage de constituer un corpus conséquent, couvrant une période très large, tout en étant fournis en formules linguistiques suffisamment élaborées. En outre, l'on suppose ici, toujours selon notre paradigme évolutionnaire, que les auteurs de ces œuvres, quoique fictionnelles, se doivent toujours de répondre *a minima* aux attentes de leur public. Ils parleraient ainsi, plus ou moins consciemment, un langage propre à évoquer la psychologie de toute une époque.

Première partie

Agentivité et langage

Chapitre 1

Une définition de l'agentivité

1.1 *The Big Two* : Agentivité et Communion

Parmi toute la littérature psychologique qui cherche à définir la structure de la personnalité humaine, le modèle des *Big Five* est souvent retenu. Comme son nom l'indique, il s'agit d'un modèle descriptif de la personnalité fondé sur cinq traits centraux (Neuroticisme, Extraversion, Ouverture, Agréabilité et Conscienciosité). Construits par analyse factorielle, ces traits représentent les intercorrélations de comportements spécifiques et répétés, et sont largement supposés constituer les dimensions indépendantes les plus élevées dans la structure hiérarchique de notre personnalité. En d'autres termes, la personnalité d'un individu, entendue ici comme un ensemble de traits comportementaux émis au sein de classes spécifiques définies par leur contenu dans un éventail de situations, et de manière cohérente dans le temps, serait tout au plus réductible à la combinaison de ces cinq traits fondamentaux.

Toutefois, il est apparu récemment que ces facteurs n'étaient pas totalement indépendants, à savoir qu'il existait toujours des corrélations entre certains d'entre eux^{1 2}. De là ont été définies deux dimensions plus larges encore, qui constitueraient les strates les plus génériques au moment de décrire la structure de notre personnalité, et pour l'instant considérées comme véritablement orthogonales : l'Agentivité et la Communion, communément désignées par le modèle des *Big Two*³.

1. Jack Block, « A contrarian view of the five-factor approach to personality description. » *Psychological Bulletin*, 117-2 (1995), Place : US Publisher : American Psychological Association, p. 187-215, DOI : [10.1037/0033-2909.117.2.187](https://doi.org/10.1037/0033-2909.117.2.187)

2. Vincent Egan, Ian Deary et Elizabeth Austin, « The NEO-FFI : Emerging British norms and an item-level analysis suggest N, A and C are more reliable than O and E. » *Personality and Individual Differences*, 29-5 (2000), Place : Netherlands Publisher : Elsevier Science, p. 907-920, DOI : [10.1016/S0191-8869\(99\)00242-1](https://doi.org/10.1016/S0191-8869(99)00242-1)

3. Ces dimensions sont plus ou moins équivalentes aux facteurs Alpha et Beta proposés par John M. Digman, « Higher-order factors of the Big Five. » *Journal of Personality and Social Psychology*, 73-6 (1997), Place : US Publisher : American Psychological Association, p. 1246-1256, DOI : [10.1037/0022-3514.73.6.1246](https://doi.org/10.1037/0022-3514.73.6.1246), et synthétisés chez Wiggins par les métaconcepts d'Agentivité et de Communion (Jerry S. Wiggins, « Agency and communion as conceptual coordinates for the understanding and measurement

Du latin *agens*, qui signifie « agir » ou « faire quelque chose », l'agentivité est en réalité un concept aux multiples facettes, qui peut revêtir plusieurs significations plus ou moins intriquées.

Tout d'abord, l'agentivité renvoie à « l'expérience du contrôle de ses propres actions et, à travers elles, des événements du monde extérieur ». C'est ce qu'on appelle plus précisément le « sentiment d'agentivité »⁴. En ce sens, l'agentivité est très proche du concept de « locus de contrôle », qui correspond à la croyance qu'un individu entretient sur la relation entre les actions qu'il entreprend et leurs résultats. Cette croyance peut être externe : la personne croit que des facteurs indépendants de sa volonté sont plus importants que ses propres actions au moment de façonner leurs résultats, ou bien interne : il se pense principal maître à l'œuvre, ce qui correspondrait à un plus grand sentiment d'agentivité.

En relation avec le modèle des *Big Two* cité plus haut, l'agentivité correspond également à « l'une des deux modalités fondamentales de l'existence humaine qui se manifeste dans le contenu de notre perception de nous-mêmes et des autres, dans notre comportement social et dans la manière dont notre personnalité fonctionne »⁵. Contrairement à la communion, qui vise à se rapprocher des autres, l'agentivité est centrée sur la poursuite et la réalisation d'objectifs personnels. En ce sens, l'agentivité et la communion peuvent être considérées comme des aspects complémentaires de l'adaptation humaine, qui peuvent interagir pour façonner le comportement social^{6 7}. Ils répondent en effet à deux besoins fondamentaux en société : le besoin de s'assurer une position sociale avantageuse (ou l'affirmation de soi, étroitement liée à l'agentivité), mais qui doit être modulé avec le besoin d'être intégré dans le groupe (la coopération, soutenue par cette autre dimension psychologique qu'est la communion).

Enfin, l'agentivité peut être comprise comme une relation spécifique entre l'individu et son environnement social, impliquant à la fois un sentiment d'agentivité et une activité réelle, qui se manifeste par la réalisation d'un objectif.

S'appuyant sur cette nature protéiforme qu'il cherche à théoriser, Seligman⁸ considère que l'agentivité est synthétiquement composée de trois éléments : l'efficacité (l'état d'esprit selon lequel on peut atteindre un objectif spécifique maintenant), l'optimisme

of interpersonal behavior. » Dans *Thinking clearly about psychology : Essays in honor of Paul E. Meehl*, Vol. 1 : *Matters of public interest* ; Vol. 2 : *Personality and psychopathology*. Minneapolis, MN, US, 1991, p. 89-113)

4. Chambon V. Haggard P, « Sense of agency », *Curr Biol.* 2012 May 22;22(10) :R390-2. (), DOI : [doi:10.1016/j.cub.2012.02.040](https://doi.org/10.1016/j.cub.2012.02.040). PMID:22625851.

5. David Bakan, *The duality of human existence : An essay on psychology and religion*. Pages : 242, Oxford, England, 1966 (The duality of human existence : An essay on psychology and religion.)

6. Kenneth Locke, « Agency and Communion in Social Comparison », dans 2014, p. 11-38, DOI : [10.1017/cbo9781139035583.003](https://doi.org/10.1017/cbo9781139035583.003)

7. J. S. Wiggins, « Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. »...

8. [seligman_m_maymin_p_agency_2022](#)

(l'état d'esprit selon lequel on peut atteindre cet objectif dans un avenir lointain) et l'imagination (l'état d'esprit selon lequel on peut atteindre de nombreux objectifs).

Hitlin et Edler⁹ ont également proposé une typologie intéressante de l'agentivité, plus directement fondée sur les horizons temporels des acteurs dans les situations d'action, et ses implications pour le soi, comme le détaille le tableau ci-après. Ce modèle nous servira d'heuristique connexe au moment de joindre les approches théoriques de l'agentivité et la psychologie sociale :

Table 1. Types of Agency

Agency "Type"	Analytical Scope	Temporal Scope	Characteristics	Representative Theorists
Existential	All circumstances	All temporal horizons	Pre-reflective capacity to defy social dictates Fundamental element of "free will"	Mead
Pragmatic	Novel situations	"Knife's edge" present moment	Ability to innovate when routines break down	Dewey/Joas Heise/ Smith-Lovin
Identity	Routine situations	Situationally-oriented goal attainment	Capacity to act within socially prescribed role expectations	Stryker/Burke
Life course	Life pathways	Long-range future life plans	Umbrella term for retrospective analysis of decisions made at turning points and transitions	Elder/Clausen

FIGURE 1.1 – Typologie de l'agentivité s'appuyant sur l'horizon temporel de l'action - Hitlin et Edler (2007)

Ainsi, sous quelque angle que ce soit, l'agentivité implique une activité volontaire et une concentration sur les objectifs de l'individu, qui va de pair avec l'élargissement de son horizon temporel ; et peut être considérée comme une différence individuelle. Plus finement encore, Vallacher¹⁰ intègre la dimension d'agentivité dans sa « théorie de l'identification de l'action », pour en donner une définition qui servira d'inspiration à l'ensemble de ce projet de recherche. Selon lui, étant donné qu'un acte peut être identifié de nombreuses façons par l'individu qui l'exécute, les dispositions comportementales seront inévitablement insuffisantes pour représenter les régularités du comportement, ce qui remet en question

9. Steven Hitlin et Jr. Glen H. Elder, « Time, Self, and the Curiously Abstract Concept of Agency », *Sociological Theory*, 25-2 (2007), _eprint : <https://doi.org/10.1111/j.1467-9558.2007.00303.x>, p. 170-191, DOI : [10.1111/j.1467-9558.2007.00303.x](https://doi.org/10.1111/j.1467-9558.2007.00303.x)

10. R. R. Vallacher et D. M. Wegner, « Levels of personal agency : Individual variation in action identification. »...

la validité même de toute tentative globale de classification des comportements dans des catégories de traits de personnalité *a priori*. Un schéma d'organisation plus raisonnable serait donné par la structure cognitive individuelle des identifications de l'action.

1.2 La théorie de l'identification de l'action, ou l'existence d'un continuum agentif

Pour Vallacher¹¹, le niveau d'agentivité personnelle est défini comme « une dimension indépendante qui permet de distinguer à quel point un individu a organisé ses actions en catégories abstraites et significatives qui peuvent servir à canaliser le comportement en tendances dispositionnelles ». En bref, chaque individu pourrait être placé sur un continuum agentif, mesurant son niveau d'identification des actions qu'il entreprend quotidiennement.

À l'extrémité supérieure du spectre, chez les « agents de haut niveau », la primauté est donnée à la signification de l'action au sens large. Ces agents manipulent des interprétations plus abstraites, distales et centrales de leurs actions, relatives à leurs effets causaux, à la signification véhiculée par la société et à leurs implications à long-terme pour le soi. Les actions sont ainsi dites « orientées vers un but » (*purposed-oriented*). Un tel niveau d'agentivité est associé à une meilleure planification, à une moindre impulsivité et à une plus grande maîtrise de soi¹², ce quelles que soient les variations contextuelles ; ainsi qu'à une représentation plus consciente et plus stable du soi.

À l'extrémité inférieure du spectre, les agents dits « de bas niveau » se concentrent à l'inverse sur les détails mécaniques de l'action. Ils génèrent des interprétations plus concrètes, locales et périphériques de leurs actions, qui concernent principalement leurs motivations proximales. Ces dernières sont ainsi qualifiées de *means-* ou *process-oriented*, à savoir orientées vers leurs moyens, ou leur processus. Un tel niveau d'agentivité est associé à une préférence pour le court-terme, à un moindre contrôle de soi (plus d'impulsivité, moins d'auto-motivation et moins de cohérence dans le comportement au fil du temps), ainsi qu'à une représentation moins consciente du soi - comme le suggèrent des descriptions de soi moins directement liées aux traits de caractère - et plus volatile, ces individus étant globalement plus sensibles au retour d'information externe¹³.

Une remarque s'impose ici : Vallacher parle « d'identification d'une action » au sens de son interprétation par l'individu, et non d'identification « à » l'action, ou du degré d'investissement identitaire que l'individu engage dans cette même action. Cela étant,

11. *Ibid.*

12. Kentaro Fujita, « Seeing the forest beyond the trees : A construal-level approach to self-control. » *Social and Personality Psychology Compass*, 2-3 (2008), Place : United Kingdom Publisher : Wiley-Blackwell Publishing Ltd., p. 1475-1496, DOI : [10.1111/j.1751-9004.2008.00118.x](https://doi.org/10.1111/j.1751-9004.2008.00118.x)

13. *Ibid.*

les identifications des actions qualifiées de « haut niveau » par Vallacher correspondent de fait à un engagement plus direct de sa propre identité au moment d'interpréter ces dernières, l'action étant notamment appréhendée au regard de ses implications pour la conception de son moi. Nous rapprocherons ainsi directement, dans la suite de cette étude, l'agentivité « haut niveau » à une identification à l'action, entre autres facteurs bien sûr.

Par ailleurs, selon Vallacher, « l'identité d'acte particulière qui assume la prépotence pour une personne à un moment donné reflète un compromis entre les préoccupations de compréhension globale [agentivité de haut niveau] et l'exécution efficace de l'action [agentivité de bas niveau] », cette dernière restant la réponse la plus optimale face aux actions nouvelles et/ou complexes, nécessitant une approche plus procédurale pour être exécutées correctement.

A titre d'illustration, un individu peut interpréter son échec à un examen d'abord comme la conséquence d'un manque de chance, du fait d'avoir été confronté à un professeur trop sévère, ou encore d'avoir dû aider ses parents dans une tâche et ainsi de ne pas avoir eu assez de temps pour étudier... ce qui correspondrait à un locus de contrôle externe ou, en nos termes, à une moindre agentivité. Au contraire, il peut s'en rendre directement responsable, en reconnaissant n'avoir pas suffisamment révisé, ce qui témoignerait d'un locus de contrôle interne, ou d'une plus grande agentivité.

Bien entendu, chacune de ces justifications peut être exprimée et détaillée de différentes manières, plus ou moins agentives, et explique qu'on ne réduise pas ici l'agentivité au locus de contrôle, dont la dichotomie constitue la limite. En effet, afin d'expliquer pourquoi il a personnellement failli à un examen, un individu peut alléguer qu'il a simplement fait une impasse sur un chapitre, procrastiné, ou bien qu'il a souvent peur de l'échec et tend ainsi à le prévenir en sapant toutes ses chances de réussite. On sent ici la gradation entre ces différentes réponses en termes d'implication pour le soi, ce bien qu'elles relèvent toutes d'un locus de contrôle interne. De même, en adoptant un locus de contrôle externe, un individu peut simplement affirmer, comme suggéré plus haut, « qu'il n'a pas eu de chance », en notant l'emploi du verbe statif *avoir* qui sera commenté plus tard. Mais il peut également déclarer « qu'il a dû aider ses parents », en utilisant cette fois le verbe modal *devoir*. Là encore, la seconde formulation, bien que toujours déresponsabilisante, semble placer le locuteur dans une posture relativement plus active que le simple fait d'être frappé par la malchance : l'obligation implique en effet, quoique par la négative, un processus de choix, complètement absent de la première interprétation, et dont on pourrait supposer qu'il engage davantage le moi, à savoir la manière dont l'individu se conçoit à partir de cette action (par exemple comme quelqu'un d'excessivement obéissant).

Ce projet consiste donc à tenter de déterminer s'il est possible d'identifier des marqueurs linguistiques fiables permettant d'évaluer, de manière systématique, la teneur agentive d'un texte, et par extension le niveau d'agentivité de son locuteur.

Chapitre 2

Le langage comme reflet de la psychologie humaine

2.1 Transitivité et *foregrounding* : des indices corrélés de l'efficacité de l'action

Hopper et Thompson¹ définissent la transitivité comme l'efficacité (*effectiveness*) avec laquelle une action se déroule. Loin de se limiter à la présence d'un Agent et d'un patient, au sens grammatical (un Agent grammatical étant défini comme « une cause unique à partir de laquelle une chaîne ininterrompue de contrôle conduit à l'effet », ou plus simplement comme la responsabilité première d'une entité dans l'effet rapporté sur le monde), elle inclurait les composantes suivantes :

	HIGH	LOW
(1) A. PARTICIPANTS	2 or more participants, A and O. ¹	1 participant
B. KINESIS	action	non-action
C. ASPECT	telic	atelic
D. PUNCTUALITY	punctual	non-punctual
E. VOLITIONALITY	volitional	non-volitional
F. AFFIRMATION	affirmative	negative
G. MODE	realis	irrealis
H. AGENCY	A high in potency	A low in potency
I. AFFECTEDNESS OF O	O totally affected	O not affected
J. INDIVIDUATION OF O	O highly individuated	O non-individuated

FIGURE 2.1 – Elements constitutifs de la transitivité selon Hopper et Thompson (1980)

Selon eux, plus une clause a de caractéristiques dans la colonne *haut* (*high*), plus elle est transitive. Ainsi, une phrase avec deux participants comme « Jerry aime la bière

1. Paul Hopper et Sandra Thompson, « Transitivity in Grammar and Discourse », *Language*, 56 (juin 1980), DOI : [10.2307/413757](https://doi.org/10.2307/413757)

», où le patient *bière* ne reçoit pas vraiment d'action, peut en réalité être moins bien notée qu'une phrase avec un seul participant, telle que « Susan est partie ».

Aussi, les auteurs notent que dans toutes les langues examinées au fil de leur étude (qui dépasse largement le cadre de l'anglais, l'approche étant précisément trans-culturelle), les éléments constitutifs de la transitivité covarient largement et systématiquement. Cela signifie qu'à chaque fois qu'une paire obligatoire de deux éléments de transitivité apparaît dans la morphosyntaxe ou la sémantique d'une clause, les éléments appariés se trouvent toujours du même côté de l'échelle de la transitivité haute/basse.

De là, ils considèrent la transitivité comme une propriété centrale de l'utilisation du langage. Probablement en raison d'une limitation psychologique dans le traitement du discours, elle servirait de signal morphosyntaxique des parties du discours qui doivent être stockées pour un traitement séquentiel immédiat (les parties dites « au premier plan » ou *foregrounded*, qui communiquent la structure temporelle du récit), par opposition aux parties qui doivent être stockées pour une référence future ou un accès concomitant (les parties « en arrière-plan », ou *backgrounded*, et qui ne contribuent pas immédiatement ni de manière cruciale à l'objectif du locuteur mais l'assistent, l'amplifient ou le commentent simplement). En bref, la probabilité qu'une clause reçoive une interprétation « de premier plan » est conçue comme proportionnelle à la hauteur de cette clause sur l'échelle de la transitivité.

Sur la base de cet article, et en prenant le degré d'agentivité du participant comme caractéristique référentielle (ce que les auteurs appellent *potency* ou *agency*), on peut donc postuler que si un participant s'exprime, ou est exprimé, comme ayant plus d'agentivité, toute composante supplémentaire de transitivité présente dans la clause considérée devrait tendre corrélativement vers l'extrémité « haute » de l'échelle. Par exemple, si un événement est évoqué dans une phrase contenant un Agent au sens grammatical, composante linguistique qui semble ainsi la plus directement liée à la notion d'agentivité, cet événement aura tendance à être une action plutôt qu'un état. Et dans le cas particulier d'une action, cette dernière aura tendance à être (interprétée comme) plus volitive et/ou téléique, c'est-à-dire représentée comme ayant une fin inhérente, naturelle ou volontaire. Enfin, si un objet est présent avec l'Agent dans une clause donnée, il aura tendance à être plus individué.

Une telle décomposition nous conduit à adopter une approche par faisceau d'indices, en tentant de construire un indice global d'agentivité intégrant toutes ces composantes, et qui devrait être globalement corrélé à l'agentivité du locuteur. L'idée ici est que même si certaines des composantes linguistiques incluses dans l'indice s'avèrent en réalité non corrélées, voire inversement corrélées avec l'agentivité par rapport à nos prédictions, l'indice devrait globalement évoluer dans la même direction que l'agentivité, tandis que des analyses successives pourront permettre de déterminer la qualité et la pertinence de chaque marqueur qui le compose.

Cela étant, la mesure du niveau d'agentivité d'un texte nécessite une modulation des indices proposés par Hopper et Thompson, l'agentivité n'étant plus simplement une composante parmi d'autres de la transitivité mais une variable dépendante à part entière. En ce sens, si l'on considère comme Hopper et Thompson que l'imposition de combinaisons linguistiques particulières dépend fondamentalement de la fonction dominante attribuée au langage, nous pourrions supposer qu'à la place du compromis entre *foregrounding* et *backgrounding* que les auteurs mettent en avant, ou bien au-delà de ce dernier, un compromis s'opère entre un usage purement « communicatif » de la langue et un usage « performatif » (au sens d'un langage conçu d'abord comme vecteur de construction identitaire par le locuteur-même), qui déterminerait l'utilisation préférentielle de constructions morphosyntaxiques potentiellement non chevauchantes, au-delà des seuls marqueurs lexicaux. Une telle conception fait directement écho à l'arbitrage théorisé par Vallacher² entre la bonne exécution d'une action (ici, raconter une histoire à quelqu'un) et la compréhension globale de ce qu'elle signifie pour soi (ici, faire du récit lui-même un enjeu identitaire, sujet récurrent en littérature et qui en ferait un champ d'étude privilégié, comme suggéré plus loin dans ce rapport).

Alternativement, si l'on considère que le langage sert toujours d'abord à communiquer, en retenant donc le fondement psychologique proposé par Hopper et Thompson, il est possible d'envisager qu'un langage plus agentif corresponde à un autre type de *foregrounding* que celui mentionné par les auteurs, motivé par cette dimension psychologique spécifique qu'est l'agentivité. Plutôt que sa structure temporelle, le locuteur s'attacherait alors à mettre en avant la structure identitaire de son récit, ce qu'il dit sur son *soi*. Là encore, ce *foregrounding* pourrait impliquer des composantes linguistiques spécifiques, ne correspondant pas totalement à celles qui participent de la transitivité, ou efficacité temporelle, d'un discours. Néanmoins, l'on s'attendrait à trouver un certain nombre de convergences linguistiques entre ces deux motifs, l'agentivité étant crucialement dépendante de la capacité à adopter une perspective plus long-termiste, tel que le suggérerait le modèle de Hitlin et Edler précédemment cité et les différentes manières de considérer le soi en fonction du cadre temporel adopté. Le soi peut en effet être considéré comme un instrument conceptuel nécessaire, un liant sémantique dont la fonction originelle serait précisément de permettre la pensée exploratoire³, c'est-à-dire d'être orienté vers le futur, et ainsi de s'autoréguler à plus long terme par la rétrospection et l'anticipation, plutôt que de moduler sa performance dans le présent, focalisation caractéristique d'individus peu agentifs.

2. R. R. Vallacher et D. M. Wegner, « Levels of personal agency : Individual variation in action identification. »...

3. Thomas T. Hills et Stephen Butterfill, « From foraging to autonoeitic consciousness : The primal self as a consequence of embodied prospective foraging », *Current Zoology*, 61–2 (avr. 2015), _eprint : <https://academic.oup.com/cz/article-pdf/61/2/368/32970825/czoolo61-0368.pdf>, p. 368-381, DOI : [10.1093/czoolo/61.2.368](https://doi.org/10.1093/czoolo/61.2.368)

Si l'agentivité au sens grammatical (*potency* ou *agency*) du participant, sa volitionnalité et l'aspect téléique de la clause en question, participant de sa transitivité selon Hopper et Thompson, peuvent également être considérés comme dénotant de manière concordante un degré supérieur d'agentivité, au sens de Vallacher, cela est cependant moins évident s'agissant de l'individuation de l'objet grammatical, définie par Hopper et Thompson de la manière suivante :

INDIVIDUATED	NON-INDIVIDUATED
proper	common
human, animate	inanimate
concrete	abstract
singular	plural
count	mass
referential, definite	non-referential

FIGURE 2.2 – Éléments participant de l'individuation de l'objet grammatical selon Hopper et Thompson (1980)

De fait, bien qu'elle soit considérée comme une composante transitive « faible » dans le tableau ci-dessus, nous nous attendons plutôt à ce que l'abstraction soit corrélée avec un niveau supérieur d'agentivité, étant donnée l'interprétation plus conceptuelle de leurs actions qui caractérise les agents « haut niveau ». Il en va de même pour les modes de l'irréel, classés par les auteurs dans la partie « basse » de l'échelle transitive.

La direction de la corrélation, si elle existe, pour la composante d'affirmation n'est quant à elle pas évidente s'agissant d'agentivité. Cela peut s'expliquer par le fait qu'Hopper et Thompson définissent la transitivité comme l'*efficacité* d'une action, à savoir son caractère plus ou moins réalisé, là où l'agentivité inclut également, pour reprendre la tripartition de Seligman^{4 5}, l'optimisme et l'imagination, que nous subsumerons sous la notion de « capacité de distanciation psychologique », ainsi que « l'auto-efficacité » ou « sentiment d'agentivité », qui n'est que partiellement corrélé à l'efficacité des actions entreprises par l'individu.

2.2 Individualisme et pronoms : même méthode, nouveau raisonnement

Le langage a souvent été utilisé comme un indicateur de dimensions psychologiques ou traits de personnalité spécifiques, ce à différents niveaux d'analyse. L'utilisation des

4. Martin Seligman, « Psychological history and predicting the future », *Possibility Studies & Society*, 1–1 (2023), _eprint : <https://doi.org/10.1177/27538699221128224>, p. 206-210, DOI : [10.1177/27538699221128224](https://doi.org/10.1177/27538699221128224)

5. *Ibid.*

pronoms en particulier, et leur lien avec la notion d'individualisme, a fait l'objet de nombreuses études. Nous précisons que toutes ces études traitent de l'anglais, et non du français. Néanmoins, et malgré l'existence de différences certes importantes entre ces deux langues, il semble n'y avoir aucune raison de penser qu'une approche psycholinguistique ne puisse s'appliquer au français également, et qu'elle révélerait qui plus est les mêmes associations que celles décrites ci-après, étant donnés les marqueurs linguistiques relativement simples considérés ici, ainsi que les similitudes culturelles entre ces deux pays.

Dans une étude de 1998, Kashima et Kashima⁶ ont constaté que l'augmentation de la richesse et le changement climatique prédisaient l'individualisme, mais que le résultat était modulé par l'utilisation de la langue : les cultures individualistes auraient tendance à favoriser l'utilisation explicite des pronoms de la première personne, tandis que les cultures collectivistes permettraient aux locuteurs d'abandonner les pronoms de la première personne dans la structure superficielle des phrases⁷. De manière cohérente, l'utilisation de langues spécifiques, comme l'anglais, a été associée à des concepts de soi plus indépendants chez les personnes bilingues⁸. Plus généralement, des recherches antérieures ont établi un lien entre les pronoms de la première personne du singulier (tels que « je », « moi ») et l'individualisme d'une part, et entre les pronoms de la première personne du pluriel (tels que « nous ») et le collectivisme d'autre part^{9 10 11}.

Au sein d'une même culture, Twenge¹² a pu observer que, dans les livres américains publiés entre 1960 et 2008, l'utilisation des pronoms de la première personne du pluriel avait diminué de 10%, tandis que les pronoms de la première personne du singulier avaient augmenté de 42%, et les pronoms de la deuxième personne (« tu », « toi ») quadruplé. Cela suggérerait d'après l'auteure la montée d'un individualisme aux Etats-Unis, avec une tendance parallèle vers moins de collectivisme.

6. Emiko S. Kashima et Yoshihisa Kashima, « Culture and language : The case of cultural dimensions and personal pronoun use. » *Journal of Cross-Cultural Psychology*, 29-3 (1998), Place : US Publisher : Sage Publications, p. 461-486, DOI : [10.1177/0022022198293005](https://doi.org/10.1177/0022022198293005)

7. Voir également Feng Yu, Theodore Peng, Kaiping Peng, Shi Tang, Chuan Shi Chen, Xiaojun Qian, Pei Sun, Tingting Han et Fangyuan Chai, « Cultural value shifting in pronoun use. » *Journal of Cross-Cultural Psychology*, 47-2 (2016), Place : US Publisher : Sage Publications, p. 310-316, DOI : [10.1177/0022022115619230](https://doi.org/10.1177/0022022115619230)

8. Viorica Marian et Margarita Kaushanskaya, « Self-construal and emotion in bicultural bilinguals. » *Journal of Memory and Language*, 51-2 (2004), Place : Netherlands Publisher : Elsevier Science, p. 190-201, DOI : [10.1016/j.jml.2004.04.003](https://doi.org/10.1016/j.jml.2004.04.003)

9. Wendi L. Gardner, Shira Gabriel et Angela Y. Lee, « "I" value freedom, but "we" value relationships : Self-construal priming mirrors cultural differences in judgment. » *Psychological Science*, 10-4 (1999), Place : United Kingdom Publisher : Blackwell Publishing, p. 321-326, DOI : [10.1111/1467-9280.00162](https://doi.org/10.1111/1467-9280.00162)

10. E. S. Kashima et Y. Kashima, « Culture and language : The case of cultural dimensions and personal pronoun use. »...

11. Jinkyung Na et Incheol Choi, « Culture and first-person pronouns. » *Personality and Social Psychology Bulletin*, 35-11 (2009), Place : US Publisher : Sage Publications, p. 1492-1499, DOI : [10.1177/0146167209343810](https://doi.org/10.1177/0146167209343810)

12. J. M. Twenge, W. Keith Campbell et Brittany Gentile, « Changes in pronoun use in American books and the rise of individualism, 1960-2008. » *Journal of Cross-Cultural Psychology*, 44-3 (2013), Place : US Publisher : Sage Publications, p. 406-415, DOI : [10.1177/0022022112455100](https://doi.org/10.1177/0022022112455100)

Ces résultats peuvent être examinés à la lumière des recherches sur l'utilisation individuelle des pronoms : dans la correspondance personnelle et le discours, les pronoms singuliers de la première personne ont également été associés à une plus grande focalisation sur l'individu, ainsi qu'à un statut inférieur, à l'honnêteté, à la dépression et à un style plus personnel et plus expressif. À l'inverse, les personnes de statut supérieur utiliseraient davantage de pronoms de la première personne du pluriel et de la deuxième personne¹³.

Si l'on tente de faire dialoguer ces deux niveaux d'analyse, culturel et individuel, les résultats de Twenge suggèrent donc que le langage (américain) aurait « évolué pour devenir moins hiérarchique, plus centré sur soi et plus expressif (peut-être dépressif) », bien que cela ne soit en accord qu'avec certains éléments de l'individualisme seulement (augmentation de l'expression émotionnelle, moins de hiérarchie dans l'ensemble, plus d'égoïsme), mais pas nécessairement avec d'autres (plus d'excès de confiance et de revendication de statut, ce qui contraste avec l'association, au niveau individuel, entre un statut inférieur, l'honnêteté et la dépression d'une part, et l'utilisation de pronoms à la première personne du singulier d'autre part). L'utilisation de la deuxième personne en particulier, est associée à un statut supérieur et augmenterait avec le temps, de sorte que les résultats en matière de statut sont au mieux décrits comme mixtes.

Dans le même ordre d'idées, une analyse linguistique des 10 chansons américaines les plus populaires entre 1980 et 2007¹⁴ a mis en évidence des changements dans l'utilisation de mots censés refléter des changements psychologiques : au fil du temps, l'utilisation de mots liés à la focalisation sur soi et au comportement antisocial aurait augmenté, tandis que les mots liés à la focalisation sur l'autre, aux interactions sociales et à l'émotion positive semblent avoir diminué.

Enfin, Twenge¹⁵ a constaté que, par rapport aux générations précédentes, un plus grand nombre d'étudiants américains se considèrent aujourd'hui comme « supérieurs à la moyenne » s'agissant d'attributs tels que les capacités académiques, la volonté de réussir, les capacités de leadership, la capacité à s'exprimer en public, la confiance en soi et la capacité à écrire (ce en tenant compte de la race, du sexe, des performances et du temps passé à étudier, ce dernier étant en fait en baisse ; et en dépit du fait que la population universitaire soit devenue moins sélective). L'auteure conclut que de vastes changements culturels mettant l'accent sur une vision positive de soi (aux États-Unis du moins) ont apparemment entraîné une amélioration des auto-évaluations dans les domaines agitifs,

13. John Nerbonne, « The Secret Life of Pronouns. What Our Words Say About Us », *Literary and Linguistic Computing*, 29 (avr. 2014), p. 139-142, DOI : [10.1093/llc/fqt006](https://doi.org/10.1093/llc/fqt006)

14. C. Nathan DeWall, Richard S. Pond Jr., W. K. Campbell et J. M. Twenge, « Tuning in to psychological change : Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. » *Psychology of Aesthetics, Creativity, and the Arts*, 5-3 (2011), Place : US Publisher : Educational Publishing Foundation, p. 200-207, DOI : [10.1037/a0023195](https://doi.org/10.1037/a0023195)

15. J. M. Twenge, W. K. Campbell et B. Gentile, « Generational increases in agentic self-evaluations among American college students, 1966–2009. » *Self and Identity*, 11-4 (2012), Place : United Kingdom Publisher : Taylor & Francis, p. 409-427, DOI : [10.1080/15298868.2011.576820](https://doi.org/10.1080/15298868.2011.576820)

tandis que les auto-évaluations des attributs communautaires, tels que la compréhension des autres, la coopération et la spiritualité, auraient soit diminué, soit stagné. Aussi, les produits culturels, notamment dans leur forme linguistique, fonctionneraient à leur tour pour entretenir et accélérer de tels changements générationnels, vers la démonstration de traits psychologiques de plus en plus individualistes (en incluant dans cette notion aussi bien l'estime de soi, l'agentivité et le narcissisme).

Cependant, l'ambivalence du lien entre l'individualisme au niveau culturel et individuel mentionnée plus haut pose question, et pourrait laisser penser que l'agentivité est ici définie de manière trop restrictive. En effet, Twenge semble assimiler l'agentivité à un trait de personnalité parmi d'autres, qui résulterait de l'intériorisation d'un individualisme croissant au niveau culturel, les deux se renforçant alors mutuellement. En outre, cet individualisme est compris comme une valeur culturelle donnant la priorité à l'individu sur la collectivité, et est donc étroitement lié à l'affirmation de son propre statut dans le monde, vécu positivement. Non seulement ce raisonnement ne s'inscrit pas dans une approche évolutionnaire que nous expliciterons ci-après, mais nous adoptons une définition plus globale de l'agentivité personnelle, qui ne se résume pas à un jugement positif sur ses propres capacités individuelles, ou estime de soi. Twenge elle-même semble y faire allusion lorsqu'elle considère que ces auto-évaluations plus mélioratives des étudiants américains pourraient résulter, plutôt que d'un égocentrisme plus prononcé, du fait que «les étudiants récents interprètent les attributs de manière plus large que les générations précédentes, trouvant plus de façons de se considérer comme supérieurs à la moyenne »¹⁶. Cette « surinterprétation » correspondrait, en nos termes, à un niveau d'agentivité plus élevé, dans la mesure où un individu qui tend à identifier ses actions de manière plus abstraite au quotidien pourrait être corrélativement plus désireux et/ou capable d'extrapoler n'importe quelle notion, au-delà des actions seules, et d'en trouver ainsi une certaine définition selon laquelle on pourra dire qu'il est particulièrement performant - ce d'autant plus s'il s'agit de compétences métacognitives alors *de facto* plus développées.

Si l'on peut s'attendre à ce qu'un niveau d'agentivité plus élevé favorise l'apparition de certains traits individualistes (narcissisme, extraversion, etc.), il n'y a donc pas de raison de postuler une correspondance biunivoque entre les deux concepts. La tendance au neuroticisme, par exemple, pourrait précisément être associée à un « débordement » de l'agentivité « haut niveau », ce contrairement à ce que l'individualisme appréhendé au niveau culturel semble prédire. Dans ce cas, l'individu serait incapable de passer à un niveau de conceptualisation engageant relativement moins son moi dans des situations nouvelles et/ou complexes, et serait donc amené à s'investir outre-mesure dans ce qui risque d'être mal exécuté, menaçant ainsi l'image de soi (comme le reflète le langage

16. David Dunning, Judith A. Meyerowitz et Amy D. Holzberg, « Ambiguity and self-evaluation : The role of idiosyncratic trait definitions in self-serving assessments of ability. » *Journal of Personality and Social Psychology*, 57-6 (1989), Place : US Publisher : American Psychological Association, p. 1082-1090, DOI : [10.1037/0022-3514.57.6.1082](https://doi.org/10.1037/0022-3514.57.6.1082)

utilisé dans les communications personnelles).

Selon cette logique, il est alors possible de dépasser la conception péjorative de l'individualisme, définie comme un simple égocentrisme, pour le comprendre plutôt comme le corollaire naturel d'une plus grande agentivité, dont on rappelle l'orthogonalité supposée avec la communion.

Deuxième partie

Questions de recherche et méthodologie

Chapitre 3

L'évolution de nos préférences psychologiques, un paradigme évolutionnaire

3.1 Les fondements de la psychologie évolutive

Comme pressenti plus haut, le pouvoir explicatif des grands mouvements culturels appliqués aux changements dans la psychologie et le comportement des individus peut paraître assez limité. En l'occurrence, l'idée qu'avec la montée de l'individualisme, appréhendé comme une évolution culturelle *sui generis*, chaque individu aurait tendance à devenir de plus en plus égocentrique, contribuant ainsi à enraciner l'avènement de cette nouvelle ère du « je » au niveau collectif (qui n'existe en réalité que par la somme de ses membres) frustre par son apparente circularité.

Le paradigme évolutionnaire constitue ainsi une explication alternative intéressante, où l'Histoire est bien celle des individus avant d'être celle du groupe. Selon lui, notre plasticité adaptative, définie comme la capacité à moduler notre phénotype (à savoir l'ensemble de nos caractéristiques observables) en fonction de l'environnement, impliquerait notamment la modulation de nos préférences psychologiques. Ces changements psychologiques seraient les mécanismes « proximaux » permettant *in fine* à chaque individu d'optimiser sa stratégie d'allocation des ressources en fonction des transformations de son milieu (cette stratégie constituant le mécanisme « ultime » dans l'optimisation de son succès reproductif individuel).

Car le phénotype optimal d'un individu, qui inclut, outre ses caractères physiques, ses stratégies comportementales, traits de personnalité et préférences psychologiques, varie de fait en fonction des ressources disponibles dans son environnement, entendues au sens large (matérielles, institutionnelles, etc). En particulier, les modèles d'allocation optimale des ressources montrent que les individus devraient investir dans les besoins dont

le rendement marginal est le plus élevé. Et étant donné que les besoins ont des rendements marginaux décroissants, lorsque les individus ont plus de ressources à investir, ils commenceraient à allouer des ressources à de nouveaux besoins. A mesure que les individus progressent dans la pyramide des besoins¹, leur psychologie se modifierait, et ils développeraient des préférences leur permettant, ultimement, de tendre vers la satisfaction des besoins devenus les plus adaptatifs dans l'environnement considéré (à savoir ceux qui maximisent leur *fitness*, qui constitue la représentation quantitative du succès reproductif de l'individu).

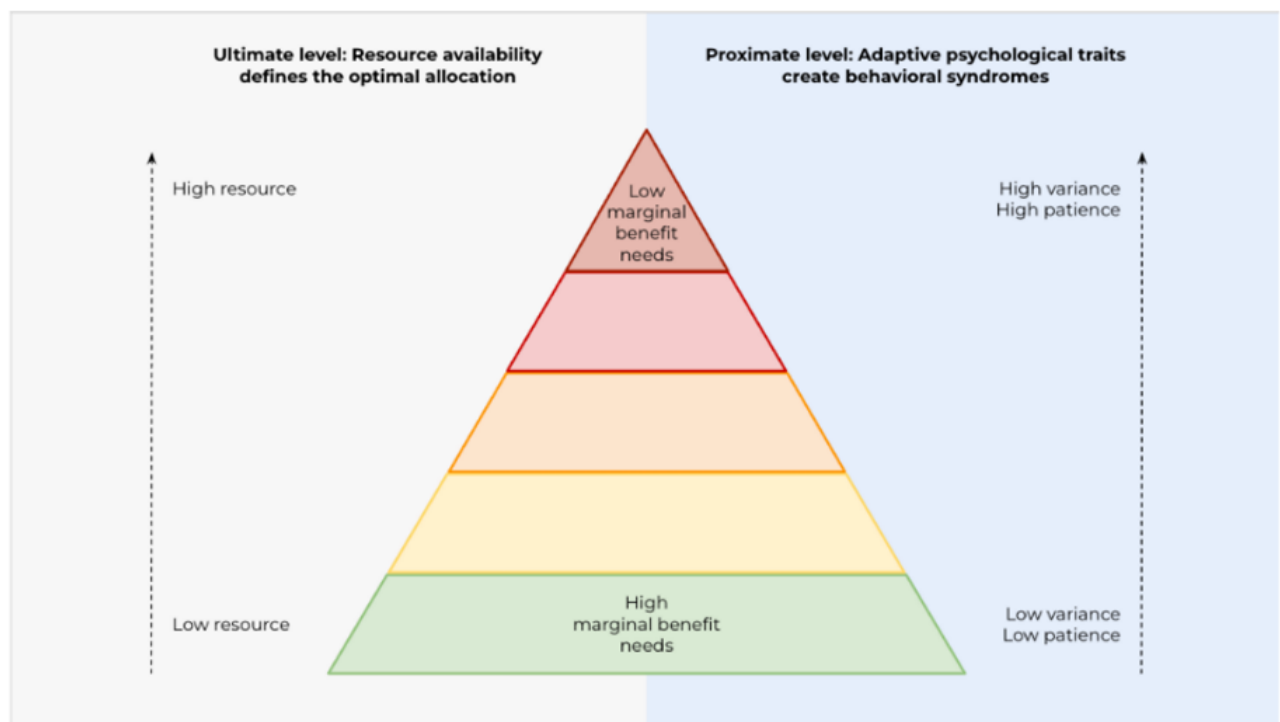


FIGURE 3.1 – La Pyramide des besoins de Maslow (1943) intégrée à un paradigme évolutionnaire

À cet égard, et en ce qui concerne l'agentivité, au fur et à mesure que l'environnement devient plus riche et plus prévisible, il deviendrait de plus en plus adaptatif pour un individu de se penser comme capable d'avoir une prise sur le monde extérieur, précisément parce que l'autonomie devient alors possible, et que les efforts investis sont susceptibles de générer des bénéfices individuels additionnels. A l'inverse, dans un environnement très incertain et/ou hostile, il serait plus adaptatif de se sentir relativement moins agentif et de préférer s'en remettre à l'apprentissage de lois extérieures alors perçues comme « naturelles ». Cela permettrait en effet de s'en accommoder au mieux, puisqu'il serait *de facto* plus difficile, voire impossible, d'infléchir ces forces extérieures en faveur de projets autonomes,

1. A. H. Maslow, « A theory of human motivation. » *Psychological Review*, 50-4 (1943), Place : US Publisher : American Psychological Association, p. 370-396, DOI : [10.1037/h0054346](https://doi.org/10.1037/h0054346)

ces derniers représentant alors un effort relativement vain, voire désavantageux en termes de fitness.

Autant de considérations qui font écho au compromis de Vallacher² mentionné précédemment, mais formulé cette fois en des termes évolutionnaires : lorsque la quantité de ressources dans l'environnement est suffisamment élevée pour que la plupart des actions réalisées soient relativement faciles à prendre en charge (y compris la possibilité de s'appuyer sur un système institutionnel pour déléguer une partie de la satisfaction de ses besoins « les plus bas »), l'individu n'aurait plus besoin de s'attarder sur leurs détails mécaniques, et pourrait se permettre d'allouer ses ressources cognitives à l'interprétation plus abstraite de ses actions, le développement de capacités de planification associées étant susceptible d'optimiser son succès reproductif à plus long terme.

De tout ce qui précède, deux questions de recherche se posent pour nous, dont seule la première sera abordée dans le cadre de ce mémoire :

1. Est-il possible de construire un indice linguistique d'agentivité, c'est-à-dire un faisceau de marqueurs linguistiques qui permette de mesurer le degré d'agentivité d'un texte, et par extension, de son locuteur ?

2. Si tel est le cas, comment l'agentivité, mesurée par un tel indice, varie-t-elle avec le temps et les ressources dans une société donnée ? Sur la base du paradigme adopté, nous supposons que le niveau « moyen » d'agentivité dans une société devrait être positivement corrélé avec son niveau de ressources.

3.2 Les fictions narratives, des « produits dérivés » de l'esprit humain

Si notre ambition ultime est celle d'exploiter notre indice linguistique pour retracer l'évolution de l'agentivité d'individus réels, pourquoi donc l'élaborer à partir de textes fictionnels ? Il se trouve que notre paradigme évolutionnaire fait de la fiction narrative un support particulièrement pertinent pour retracer l'histoire de la psychologie humaine.

De fait, il est possible de considérer les fictions narratives comme des « produits dérivés » de l'esprit humain, à savoir qu'elles n'auraient pas évolué par sélection naturelle mais auraient coopté certaines préférences et mécanismes cognitifs préexistants, sélectionnés avant même l'émergence de la culture symbolique et pour d'autres raisons liées à notre Environnement d'Adaptation Évolutive (l'environnement ancestral auquel notre espèce est adaptée)³. Selon cette hypothèse, la fiction survit car nous tirons certains bénéfices de la production et de la consommation de ces objets culturels attrayants, ce qui

2. R. R. Vallacher et D. M. Wegner, « Levels of personal agency : Individual variation in action identification. »...

3. Edgar Dubourg et Nicolas Baumard, « Why and How Did Narrative Fictions Evolve ? Fictions as Entertainment Technologies », *Frontiers in Psychology*, 13 (2022), DOI : [10.3389/fpsyg.2022.786770](https://doi.org/10.3389/fpsyg.2022.786770)

la rend indirectement adaptative.

Les fictions narratives peuvent ainsi être considérées comme des « technologies de divertissement », à savoir comme des objets conçus par certaines personnes dans le but immédiat d'attirer l'attention des autres, et dans le but ultime de remplir d'autres fonctions évolutives pertinentes qui deviennent plus faciles à atteindre une fois que l'attention des autres a été captée. Cela permettrait d'expliquer pourquoi la fiction est souvent remplie de stimuli exagérés et divertissants, pourquoi elle s'adapte si bien aux préférences changeantes du public qu'elle cible, et pourquoi les producteurs rendent leur fiction de plus en plus attrayante au fil du temps, ce de manière cumulative.

Les fictions peuvent donc servir de miroir pour retracer l'évolution dans le temps de ces préférences psychologiques que chaque lecteur dicte et que chaque auteur coopte, et qui changent en fonction des transformations du milieu.

Chapitre 4

Construction d'un indice d'agentivité global : une approche par faisceau d'indices

4.1 Validation interne : le choix d'un corpus théâtral

Se pose alors la question du corpus à sélectionner pour éprouver notre indice agentif, dont nous détaillerons la composition dans la partie suivante. Les indicateurs linguistiques ayant été choisis sur la base de la langue française, l'étude porte ici sur les productions littéraires françaises.

En particulier, la forme stylistique relativement standardisée du théâtre permet de contrôler certaines variations qui pourraient être liées à des phénomènes plus aléatoires de transmission culturelle plutôt qu'à la sélection et à la plasticité proprement dites, qui font du passage d'une préférence psychologique à une autre la réponse spécifique et optimale à certains changements environnementaux, préférence qui se refléterait en partie dans le choix des mots.

Le genre théâtral présente également l'avantage de disposer de corpus substantiels déjà constitués et librement accessibles en ligne. En l'occurrence, nous avons choisi de retenir le corpus de théâtre classique librement téléchargeable sur le site www.theatre-classique.fr, comprenant les principales œuvres théâtrales publiées en France entre 1550 et 1800.

Surtout, la dualité théâtrale classique entre personnages maîtres et serviteurs fournit un critère de validation interne particulièrement solide pour notre indice agentif. Par définition, il semble en effet raisonnable de considérer que les maîtres sont relativement plus agentifs que leurs serviteurs. Il s'agira alors de comparer le score global de l'indice obtenu pour les répliques prononcées par tous les personnages maîtres et serviteurs respectivement, une première étape pour tester la capacité de l'indice à tracer *quelque chose comme l'agentivité* dans un texte.

Concrètement, nous avons attribué à chaque personnage des pièces étudiées un statut sur la base des mots-clés apparaissant ou non pour le décrire dans la didascalie de chaque pièce. Pour

considérer un personnage comme « serviteur », ou plutôt comme doté d'un « bas statut » (le terme de « serviteur » est entendu ici dans un sens générique), statut pour lequel les mentions dans la didascalie étaient les moins ambiguës, nous avons retenu les mots-clés suivants : « domestique », « esclave », « chambre » (pour « femme de chambre »), « camariste », « servant », « servante », « suivante », « valet », « suite », « confident », « confidente ». Notons que nous avons accès à une liste de tous les « rôles » pris par les personnages dans l'ensemble du corpus, référencée par le site mentionné ci-dessus. C'est à partir de cette liste que nous avons construit notre liste de mots-clés, qui se trouve donc être exhaustive pour ce corpus.

Pour les personnages dits « haut statut », dont les mentions étaient parfois ambiguës ou simplement trop variées pour procéder par liste de mots-clés malgré l'existence d'un tel référencement, nous avons retenu la règle suivante : tout personnage dont le nom X apparaît en complément d'un mot-clé « bas statut » (par exemple « valet de X ») est automatiquement considéré comme appartenant au groupe « haut statut », dans le sens où il a quelqu'un à son service. Cela semblait être la manière la plus neutre et la plus pertinente de procéder, sans être soumis aux variations socio-historiques dont pourrait souffrir le sens d'étiquettes telles que « bourgeois », et en relation avec notre concept d'agentivité - l'idée étant qu'un personnage capable d'en commander un autre est *a fortiori* plus agentif que celui qu'il ou elle commande. Pour chaque pièce considérée, les répliques des personnages qui ne répondaient pas à ces critères dans la didascalie, et n'appartenaient donc à aucun des deux groupes, ont été exclues de l'analyse linguistique.

A partir de là, nous avons codé une manière d'attribuer à chaque personnage inclus dans l'analyse les répliques qu'il ou elle formulait dans l'ensemble de la pièce. Nous avons ensuite concaténé l'ensemble des répliques respectivement prononcées par des personnages « bas » et « haut » statut, afin de pouvoir comparer leur score agentif au niveau de chaque pièce.

4.2 Validation externe : plusieurs propositions

Nous avons également besoin d'une validation externe de notre indice, afin de s'assurer qu'il trace l'agentivité spécifiquement, sans être corrélé à d'autres dimensions psychologiques telles que la communion, auquel cas il ne saurait être informatif.

À cette fin, il existe tout d'abord une vaste littérature établissant un lien entre agentivité et masculinité¹. Examiner en parallèle si l'indice est préférentiellement corrélé avec les répliques de personnages masculins plutôt que féminins était ainsi une première façon de s'assurer de la spécificité de sa mesure.

En outre, il nous a semblé utile de comparer notre indice avec des marqueurs linguistiques spécifiques d'autres dimensions psychologiques, pour lesquelles nous ne nous attendions pas à des corrélations avec le statut du personnage, et par conséquent avec l'indice agentif. Malheureusement, aucune étude à ce jour ne s'est intéressée aux corrélations potentielles entre la communion, *a priori* orthogonale à l'agentivité, et certains marqueurs linguistiques, ce qui

1. Yuval Neria, « Encyclopedia of Personality and Individual Differences », dans 2016, p. 1-3, DOI : [10.1007/978-3-319-28099-8_2026-1](https://doi.org/10.1007/978-3-319-28099-8_2026-1)

suffirait à faire l'objet d'un projet de recherche à part entière.

Il a donc fallu se tourner à nouveau vers le modèle des *Big 5*, en dépit de ses limites. Il se trouve en effet que certains des cinq traits qui le composent entretiennent une relation privilégiée avec les deux dimensions supérieures que sont l'agentivité et la communion.

Pour l'observer, il faut en réalité mobiliser, outre les modèles dimensionnels comme celui des *Big Five*, la notion de circomplexes. Egalement très utilisés en psychologie pour décrire les relations entre un ensemble de variables, les modèles circomplexes sont généralement définis en termes de corrélations systématiques croissantes et décroissantes entre les variables, et peuvent être représentés visuellement sous la forme d'un cercle dans lequel les variables adjacentes sont fortement corrélées et les variables opposées sont inversement corrélées. Ils ont notamment été utilisés pour décrire les dispositions interpersonnelles².

A partir d'un échantillon de 315 hommes et femmes adultes, et en factorisant conjointement les auto-évaluations passées sur les Échelles d'Adjectivité Interpersonnelle révisées de Wiggins³ avec des auto-évaluations, des évaluations par les pairs et des évaluations par le conjoint fondées sur l'Inventaire de Personnalité NEO⁴, ce afin d'examiner les relations entre les deux modèles, McCrae et CostaMCCRAE (Robert R.) et COSTA (Paul T.), « The structure of interpersonal traits : Wiggins's circumplex and the five-factor model. » *Journal of Personality and Social Psychology*, 56-4 (1989), Place : US Publisher : American Psychological Association, p. 586-595, DOI : [10.1037/0022-3514.56.4.586](https://doi.org/10.1037/0022-3514.56.4.586) ont pu constater que le circomplexe interpersonnel était défini par les deux dimensions que sont l'extraversion et l'agréabilité.

2. wiggins 1979

3. Le comportement interpersonnel ayant souvent été caractérisé par des modèles circumplexes, Wiggins a entrepris de créer un ensemble d'échelles adjectives présentant des propriétés circumplexes, et auxquelles on se réfère ici. Voir J. S. Wiggins, « A psychological taxonomy of trait-descriptive terms : The interpersonal domain. » *Journal of Personality and Social Psychology*, 37-3 (1979), Place : US Publisher : American Psychological Association, p. 395-412, DOI : [10.1037/0022-3514.37.3.395](https://doi.org/10.1037/0022-3514.37.3.395)

4. Le NEO-PI est un questionnaire de 181 items développé pour opérationnaliser le modèle à cinq facteurs. Les réponses aux items du NEO-PI se font sur une échelle de Likert en 5 points. Une forme à la troisième personne du NEO-PI a également été mise au point et validée pour être utilisée par des évaluateurs. **costa 1985**

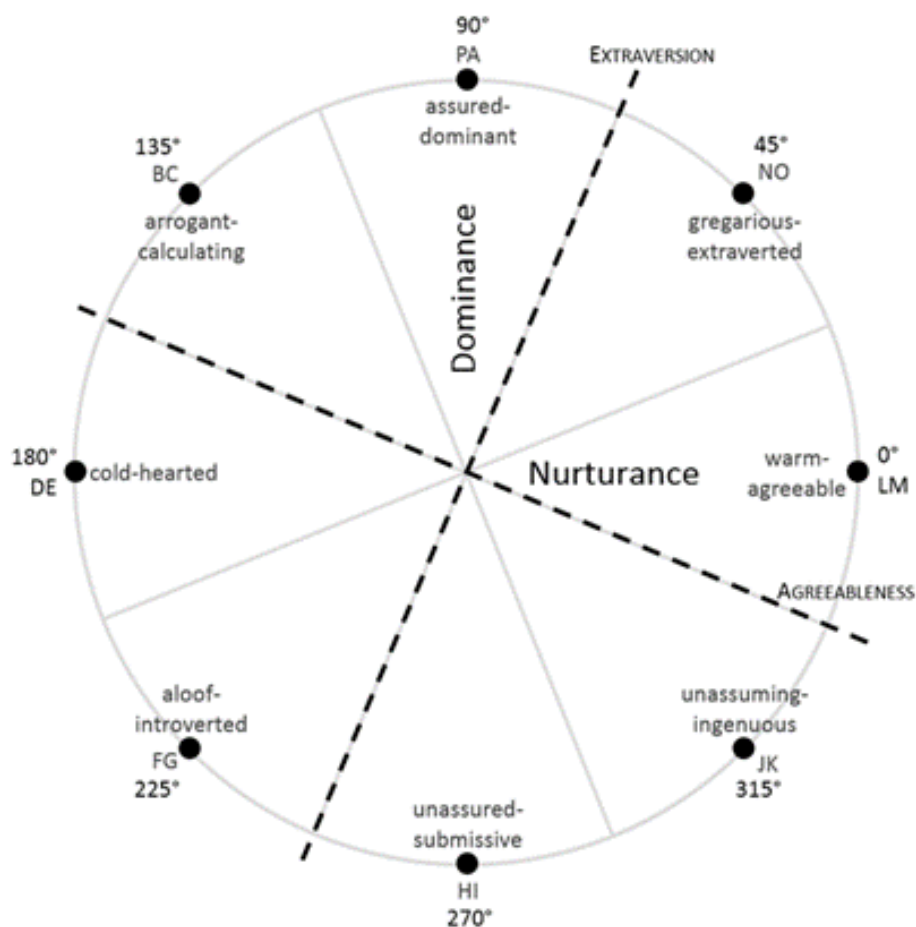


FIGURE 4.1 – Le Circomplexe interpersonnel de Wiggins (1995) et le Modèle à 5 facteurs ou *Big Five*

Comme apparent sur le cercle, les axes de l'agréabilité et de l'extraversion du *Big Five* sont en effet situés à environ 45 degrés des axes familiaux de l'amour et du statut de ce circomplexe, et semblent donc traiter d'aspects différents des mêmes dimensions latentes. A titre de remarque, l'extraversion est définie dans le modèle des *Big Five* comme une activité tournée vers l'extérieur, mais non comme un « comportement prosocial » (elle ne décrit pas, par exemple, à quel point une personne se soucie sincèrement des autres). C'est l'agréabilité qui fait précisément office de liant social, caractérisant quant à elle des personnes plus tolérantes, douces, indulgentes, agréables, altruistes et obligeantes.

L'agréabilité et l'extraversion se rapporteraient ainsi plus étroitement au comportement interpersonnel que les mesures des domaines restants du NEO-PI (névrosisme, conscienciosité et ouverture à l'expérience). Pour autant, nous avons également inclus dans les traits de personnalité préférentiellement associés à l'agentivité qui nous intéressent ici la conscienciosité, sur la base d'un certain nombre de travaux plus récents⁵. De la même façon, outre l'agréabilité, l'ouverture y est parfois positivement corrélée avec la communion, et est donc également prise en compte dans notre étude.

5. John R. Hurley, « Agency and Communion as Related to "Big Five" Self-Representations and Subsequent Behavior in Small Groups », *The Journal of Psychology*, 132-3 (1998), Publisher : Routledge
_eprint : <https://doi.org/10.1080/00223989809599172>, p. 337-351, DOI : [10.1080/00223989809599172](https://doi.org/10.1080/00223989809599172)⁶

4.3 Outils utilisés

Bien que nous détaillions de manière exhaustive les composantes de l'indice global d'agentivité dans le chapitre suivant, elles sont récapitulées et réparties ici dans des catégories définies sur la base des différentes facettes du concept d'agentivité mentionnées plus tôt, et qui ont pour partie inspiré la sélection de ces différents éléments linguistiques supposés avoir trait à l'agentivité :

TABLE 4.1 – Récapitulatif des marqueurs linguistiques retenus

Sentiment d'agentivité Locus de Contrôle	Efficacité de l'action* Disposition objective à l'action	Volition Degré de connotation agentive des événements décrits	Identification à l'action Degré de référence à soi	Compréhension globale Capacité de distanciation psychologique
<i>Verbes</i>	Agentivité grammaticale (orateur-sujet)	Transitivité	Pronoms (individuels vs. collectifs, à la 1ère personne)	Modes irréels
Voix active vs. passive	<i>Verbes</i>	Télicité	Déictiques (proximaux vs. distaux)	Alternance de temps
Modalité (interne vs. externe)	Événements vs. états	Adverbes intentionnels		Degré d'abstraction (adverbes, mots-outils, adjectifs)

** L'agentivité telle que nous la comprenons, et contrairement à Hopper et Thompson⁷, ne semble pas impliquer l'accomplissement le plus total de l'action, et/ou l'affectation la plus complète de l'objet s'il y en a un, mais plutôt la possibilité (le « pouvoir » de l'agent) et l'actualité d'une telle action (le fait qu'il y ait une action évoquée, plutôt qu'un état), ces deux concepts définissant l'efficacité de l'action, à laquelle l'agentivité ne saurait par ailleurs se réduire.*

Dans le tableau ci-dessus, la direction prédite est toujours positive, à savoir que chaque indicateur est supposé être positivement corrélé avec le niveau d'agentivité du personnage locuteur. En italique sont mentionnés les indicateurs pour lesquels la direction est ambiguë. La mention « vs. » signifie que la métrique consiste en un ratio entre les deux notions mises en rapport (calculé dans l'ordre d'apparition de ces dernières). Par exemple, la métrique « Voix active vs. passive » est égale au ratio (nombre d'occurrences de la voix active)/(nombre d'occurrences de la voix passive). Enfin, lorsque cette mention « vs. » apparaît entre parenthèses, cela signifie qu'une métrique existe également pour compter le nombre total d'occurrences du marqueur en question

7. P. Hopper et S. Thompson, « Transitivity in Grammar and Discourse »...

(en plus du ratio mentionné). A titre d'exemple, notre indice agentif inclut à la fois une métrique calculant la fréquence totale des déictiques dans les répliques étudiées, et une métrique calculant le ratio (nombre de déictiques proximaux)/(nombre de déictiques distaux) dans ces dernières.

Concernant l'analyse textuelle, nous avons utilisé le modèle *fr-dep-news-trf* de spaCy⁸, ainsi que le modèle *fr* de pie-extended⁹. Chacun des marqueurs linguistiques a été normalisé en fonction du nombre total de mots prononcés respectivement par les personnages haut et bas statut, dans chaque pièce, à l'exception :

- des signes de ponctuation individuels (rapportés à la ponctuation totale) ;
- des verbes conjugués (cette métrique, ainsi que celle mesurant le gérondif spécifiquement), modaux, transitifs, « statifs » et copules, des verbes dont le sujet est à la première personne et des marqueurs de négation (rapportés au nombre total de verbes) ;
- des métriques de temps et modes (subjonctif etc) (rapportées au nombre total de verbes conjugués)
- du marqueur d'agentivité grammaticale (rapportée au nombre total de phrases) ;
- des duplicata d'indicateurs pour lesquels la contrainte d'avoir un sujet à la première personne nous semblait pertinente à ajouter (rapportés au nombre total de verbes ayant un sujet à la première personne)
- et des ratios, a fortiori.

8. <https://spacy.io/>

9. <https://github.com/hipster-philology/nlp-pie-taggers>

Troisième partie

Détail des composants linguistiques retenus

Chapitre 5

Actionnabilité, volition et locus de contrôle : une escapade sémantique

Nous adoptons ici une approche en priorité « structurelle » du langage, plutôt que sémantique, c'est-à-dire que nous nous limitons aux niveaux grammatical et syntaxique pour la majeure partie de notre analyse. L'idée est de se concentrer de préférence sur ce qui, dans la langue, manifeste l'état d'esprit qui la sous-tend, et qui guiderait supposément l'utilisation de ses composants les plus structurels indépendamment du contenu sémantique particulier exprimé par le locuteur.

Néanmoins, il semble qu'il faille mentionner quelques travaux antérieurs de nature sémantique, en ce qu'ils sont étroitement liés à notre question de recherche, et bien qu'ils présentent des limites qui justifieront, selon nous, une focalisation ultérieure sur les éléments les plus structurels du discours.

5.1 Les limites d'une analyse par mot-clé

Sur la base des travaux de Pietraszkiewicz¹, une analyse lexico-sémantique exploratoire aurait pu consister à simplement compter, pour chaque groupe de répliques de personnages « haut » et « bas » statut, la fréquence relative de tous les lemmes inclus dans son dictionnaire d'agentivité, comme a pu le faire Seligman dans son étude de la Bible².

1. Agnieszka Pietraszkiewicz, Magdalena Formanowicz, Marie Gustafsson Sendén, Ryan L. Boyd, Sverker Sikström et Sabine Szczesny, « The big two dictionaries : Capturing agency and communion in natural language », *European Journal of Social Psychology*, 49–5 (août 2019), p. 871-887, DOI : [10.1002/ejsp.2561](https://doi.org/10.1002/ejsp.2561)

2. Seligman, M., Maymin, P., « Agency in the Bible » (2022)

TABLE 5.1 – Dictionnaire de l’Agentivité de Pietraszkiewicz

Mots
able accomplish* accurac* accurate* achiev* acquir* actualiz* adap- tab* adept* ambition* ambitious* aptitude* aptly aptness aspiration* aspire* aspiring assert* attain* authoritative* autonomous* autonomy capab* careful* choice choices clever* compet* completion confident confidently conquer* conscientious* contemplat* contend* contest* de- cid* decision* decisive* defeat* deliberat* dependable determin* diffi- cult* do doable doing eager* earn earned earning earns easiness easy effective* efficien* effort* empowered enact* endeavor* establish es- tablished establishes establishing exact* expert* fail* fluen* freedom* freely goal goal-oriented goals importan* independ* individualist in- sight* intent* intuition intuitive* keen* know* liberties liberty logic* loner* made make makes making mastered masterful* mastering mas- tery motivat* need needed needing needs objectiv* obtain* opportun* overcame overcome overcomes overcoming persever* persist* persistent pioneer* practic* pragmat* prevail* pride prideful* prior* proactive* productive* productivity proficien* prosper* proud* purpose* pursu* rational* realiz* rebel* recog* reliab* reputation* resilien* resolute* resolution resolv* responsib* reward* risk* savv* score scored scores scoring self self-* should* significant* skill skilled skillful* skills* smart smartly steadfast* strive* striving* struggl* stubborn* succeed* suc- cess* sure take takes taking tenac* think thinking thinks thought took tried tries triumph* trying unaided unyielding* vanquish* victor* will willing* willpower win winner* winning* wins wit wits witting* won won’t you you’* your yours yourself

Catégorie	Mots
Agentivité	tried, victory, knowing, wit, pride, obtain, willing, choice, thinking, purposes, freedom, resolved, needs, opportunity, easy, determined, doing, success, obtained, established, sure, taking, liberty, need, knowledge, making, purpose, able, self, took, known, think, thought, know, take, make, your, do, should, made
Efficacité	able, many, some, which, have, having, but, also, these, those
Optimisme	things, good, joy, yet, hope (mais pas "no hope")
Imagination	desire, new, truth, true, false, ask, asked, know, known

TABLE 5.2 – Les mots agentifs retenus par Seligman

Cette liste aurait d'ailleurs pu être complétée par les travaux de Rouhizadeh³, ce dernier ayant démontré qu'un locus de contrôle interne était associé à des verbes indiquant une « tentative » (*attempt* en anglais), tandis qu'un locus de contrôle externe serait associé à des verbes de cognition, de manque, de sentiments et d'espoir :

(a) Control Relevance			(b) Internal/External Control		
Verb category	r	Verbs	Verb category	r	Verbs
Audio-visual	.15**	see, say, look, talk, hear, speak	Cognition	.08**	know, think, mean, understand
Demand	.10**	need, want, ask	Missing	.07**	miss
Start/end	.08**	start, stop, follow, leave	Feeling	.06**	feel
Auxiliary	.07**	be, have, do	Auxiliary	.06**	be, have, do
Buying	.06**	buy, spend, pay	Hope	.06**	hope, wish
Like	-.06**	like, love	Attempt	-.07**	try, check

FIGURE 5.1 – Rouhizadeh et al. (2018) - Coefficients de régression logistique entre les catégories de verbes et le contrôle (les valeurs positives correspondent à la pertinence du contrôle et au contrôle externe). **: $p < 0,01$ après correction de Benjamini-Hochberg. Les caractéristiques ayant des coefficients positifs sont en bleu et celles ayant des coefficients négatifs sont en rouge

Néanmoins, et outre le fait que ces études ont été menées en anglais et non en français, la principale limite d'une analyse sémantique réside dans sa vulnérabilité aux variations historiques de sens.

Certes, une analyse sémantique latente (LSA) aurait pu servir à calculer la similarité sémantique entre chaque mot unique d'un texte ou d'un énoncé, et un nombre limité de *seed-*

3. Masoud Rouhizadeh, Kokil Jaidka, Laura Smith, H. Andrew Schwartz, Anneke Buffone et Lyle Ungar, « Identifying Locus of Control in Social Media Language », dans *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, 2018, p. 1146-1152, DOI : [10.18653/v1/D18-1145](https://doi.org/10.18653/v1/D18-1145)

*words*⁴, ce pour chaque groupe de répliques dans une pièce considérée. Il aurait également été possible de s'inspirer de la méthode « historiquement verrouillée » de Sneffjella⁵, afin de s'assurer de la stabilité sémantique, dans le temps, des *seed-words* retenus.

Pour autant, ce travail dépassait largement le cadre de notre projet de recherche actuel, qu'il aurait fait reposer sur des méthodes d'analyse elles-mêmes exploratoires dans le cas de Sneffjella. Aussi, notre paradigme et l'approche inspirée d'Hopper et Thompson⁶ justifiaient de se concentrer principalement sur les marqueurs linguistiques infra-sémantiques.

5.2 Les adverbes intentionnels, statut hybride ?

La volition étant un aspect important de l'agentivité qu'il serait difficile d'analyser autrement que par une approche sémantique, nous avons cependant décidé de définir et de mesurer l'occurrence relative d'une liste d'adverbes univoquement intentionnels. Ceux-ci présentaient l'avantage d'être en nombre relativement limité dans la langue française, et d'avoir un sens *a priori* historiquement stable.

TABLE 5.3 – Expressions adverbiales à valeur intentionnelle retenues

« volontairement », « délibérément », « expressément », « intentionnellement », « sciemment », « à dessein », « exprès »

5.3 La modalité : entre sémantique et pragmatique

En guise de clôture de cette escapade sémantique, nous avons également examiné les marqueurs modaux utilisés dans les répliques de chaque groupe de personnages.

La modalité est un concept linguistique qui s'intéresse au statut de la proposition décrivant l'événement⁷, ou plutôt à l'attitude qu'un locuteur peut avoir vis-à-vis d'un événement ou d'un état. Une phrase modalisée situe une proposition sous-jacente, ou préjacente, dans l'espace des possibilités, décrivant des événements avec des informations supplémentaires telles que leur degré de désirabilité, de plausibilité ou de faisabilité.

Dans son sens le plus restrictif, la modalité est ainsi une « catégorie de signification linguistique ayant à voir [spécifiquement] avec l'expression de la possibilité et de la nécessité »⁸.

4. Carlos Diuk, Diego Fernandez Slezak, Iván Raskovsky, Mariano Sigman et Guillermo Cecchi, « A quantitative philology of introspection », *Frontiers in Integrative Neuroscience*, 6 (2012), DOI : [10.3389/fnint.2012.00080](https://doi.org/10.3389/fnint.2012.00080)

5. Bryor Sneffjella, Michel Génereux et Victor Kuperman, « Historical evolution of concrete and abstract language revisited », *Behavior Research Methods*, 51–4 (1^{er} août 2019), p. 1693–1705, DOI : [10.3758/s13428-018-1071-2](https://doi.org/10.3758/s13428-018-1071-2)

6. P. Hopper et S. Thompson, « Transitivity in Grammar and Discourse »...

7. F. Palmer, *Mood and Modality*, 1986, URL : [doi:10.1017/CB09781139167178](https://doi.org/10.1017/CB09781139167178)

8. Kai Von Fintel, « Modality and language », dans *Encyclopedia of Philosophy*, dir. D. Borchert, 2006, t. 2, p. 20–27

On distingue traditionnellement deux grands types de modalité, qui nous intéressent particulièrement ici : la modalité épistémique (également appelée subjective ou hypothétique) et la modalité non épistémique, parfois appelée « orientée vers l'agent » (*agent-oriented*), ou encore « racine déontique » (*deontic root*), « objective » ou « pragmatique ».

La modalité épistémique concerne ce qui est possible, plausible ou nécessaire compte tenu de ce qui est connu et des preuves disponibles. Grâce à cette modalité, les locuteurs expriment un jugement sur le statut factuel de la proposition.

La modalité non épistémique est pour sa part composée de deux sous-modalités principales : l'une se concentrant sur les capacités objectives et le déroulement dynamique des événements (Palmer, 1986), parfois appelée « modalité circonstancielle » ; et l'autre ayant avoir avec les raisons subjectives de donner la priorité à un événement plutôt qu'à un autre⁹, pour laquelle il existe d'autres subdivisions selon que l'événement est priorisé en termes de normes (« déontique »), de désirs/préférences (« boulétique »), ou de buts (« téléologique »).

A noter que Bybee et al.¹⁰ ont élaboré leur propre classification s'agissant de la modalité « orientée vers l'agent », identifiant quatre sous-types : (i) l'obligation, (ii) la nécessité, (iii) la capacité et (iv) le désir. Ces auteurs discutent également d'un certain nombre de ressources pour représenter les connaissances des locuteurs ainsi que leur position à l'égard des événements et des états de fait. En particulier, l'utilisation de verbes modaux tels que les verbes anglais *must*, *should*, ou *may* permettrait aux auditeurs de comprendre comment les locuteurs représentent leurs propres obligations ainsi que celles des autres dans un monde moral principalement construit par la langue (ou plutôt le discours). L'utilisation de prédicats de volition tels que *want*, *would*, *would like*, ou *wish*, pour leur part, mettrait certains états psychologiques internes à la disposition des autres pour qu'ils les comprennent et les évaluent.

Ainsi, quelle que soit la taxonomie, la « modalité orientée vers l'agent » rend compte, par l'utilisation de verbes modaux en particulier, de l'existence de conditions internes et externes à l'accomplissement de l'action exprimée dans le prédicat principal. Avec cette définition, la modalité semble donc particulièrement liée à notre notion d'agentivité.

Cela étant, la modalité ultimement projetée par le locuteur, et/ou comprise par l'auditeur, dépend en réalité du contexte de l'occurrence, les verbes modaux seuls se révélant intrinsèquement polysémiques, ou du moins ambigus. A titre d'exemple, *devoir* dans « Il devrait pleuvoir demain » a une connotation épistémique, tandis que « les visiteurs doivent partir à 6 heures » est plutôt déontique, « Tu te dois d'y aller » boulétique, « Pardon, je dois éternuer » circonstancielle, et « Pour rentrer à temps, tu devrais prendre un taxi » téléologique.

Pour déterminer la modalité d'une clause dans une perspective compositionnelle, Warsnby¹¹ détaille ainsi les caractéristiques, au-delà des verbes modaux et dans cet ordre de désambiguïsation, qui constitueraient ce qu'elle appelle la « contrôlabilité » du participant, définie comme « la capacité d'un agent à choisir de provoquer la situation référentielle représentée par les propo-

9. Paul Portner, *Modality*, 2010, URL : [doi:10.1017/S0022226709990351](https://doi.org/10.1017/S0022226709990351)

10. Bybee, « The evolution of grammar : tense, aspect, and modality in the languages of the world », *Choice Reviews Online*, 32-8 (1^{er} avr. 1995), p. 32-4321-32-4321, DOI : [10.5860/CHOICE.32-4321](https://doi.org/10.5860/CHOICE.32-4321)

11. Anna Warsnby, (*De*)coding modality : the case of *must*, *may*, *måste* and *kan*, Lund, 2006 (Lund studies in English, 113)

sitions des phrases »¹², et qui correspond étroitement à notre définition de l’agentivité, entendue dans un sens linguistique plus étroit. On précisera à nouveau que ces composantes ont été établies pour l’anglais seul. Cela dit, et compte tenu de l’inexistence de telles analyses appliquées au français, il semblait raisonnable d’y puiser une source d’inspiration, à supposer que leurs équivalents français soient interprétés de la même façon par des lecteurs francophones, ce à quoi ce mémoire tente précisément de répondre, *in fine*.

Premièrement, la présence d’un adverbe de but (souvent une clause, comme dans « afin de [faire quelque chose] ») indiquerait une lecture non épistémique. Inversement, les adverbes modaux épistémiques tels que *possiblement*, *peut-être*, *sûrement*, etc., peuvent quant à eux neutraliser, désambiguïser ou renforcer l’interprétation épistémique du modal dans un énoncé. Comme mentionné dans la section précédente, nous avons donc créé une mesure distincte pour les adverbes intentionnels, ainsi que pour les adverbes de doute et d’affirmation (chaque fois identifiés par l’usage d’expressions régulières).

« apparemment », « peut-être », « peut être », « probablement », « vraisemblablement », « éventuellement », « supposément », « présumément », « possiblement », « potentiellement », « hypothétiquement », « a priori », « sans doute », « si jamais », « toutefois », « certainement »

TABLE 5.5 – Expressions adverbiales à valeur d’affirmation retenues

« à vrai dire », « assurément », « à coup sûr », « absolument », « bien sûr », « pour sûr », « sans conteste », « sans contredit », « bien entendu », « effectivement », « en effet », « exactement », « parfaitement », « clairement », « certainement », « évidemment », « fatalement », « forcément », « immanquablement », « indubitablement », « inéluctablement », « inévitablement », « infailliblement », « sûrement », « certes », « en vérité », « oui », « précisément », « sans doute », « nul doute », « sans aucun doute », « sans le moindre doute », « si fait », « tout à fait », « si. », « Si », « Si : », « Si ! », « soit. », « soit », « soit : », « soit ! », « volontiers », « vraiment », « naturellement », « vraisemblablement », « d’accord »

En outre, la contrôlabilité sélectionnerait un agent non générique plutôt que générique, à savoir un agent animé, qui peut être approximé par les pronoms personnels (notamment au théâtre), ou un agent défini, approximé par le nombre de déterminants définis (bien que ce proxy soit très imparfait). Elle sélectionnerait également les événements (y compris les changements

12. Alex Klinge, « The Impact of Context on Modal Meaning in English and Danish », *Nordic Journal of Linguistics*, 19 (1996), p. 35-54, URL : <https://api.semanticscholar.org/CorpusID:145424710>

d'état), principalement interprétés comme déontiques, plutôt que les états (le plus souvent épistémiques), ce qui est encore une fois en alignement avec certaines de nos considérations théoriques au moment de sélectionner les marqueurs linguistiques de l'agentivité, notamment l'inclusion d'un ratio des verbes statifs par rapport aux autres verbes (voir la partie *III. 2.3. Analyse des dépendances syntaxiques*). A cela, nous ajoutons ainsi pour le calcul de l'indice agentif une métrique comptant le nombre relatif de pronoms personnels, ainsi qu'une métrique rapportant le nombre de déterminants définis à ceux indéfinis, sur la base des propositions de Warsnby.

Selon l'auteur, il s'agirait là des conditions explicites rendant inefficace l'influence d'autres caractéristiques.

Lorsqu'elles sont absentes, il est encore possible d'examiner les verbes modaux utilisés spécifiquement comme verbe matrice (c.a.d. principal) d'une clause de complément finie, qui reçoivent généralement une lecture épistémique.

En outre, l'on peut supposer que les énoncés déontiques ont généralement une transitivité plus élevée que les énoncés épistémiques. Ceci est, une fois de plus, cohérent avec nos propres mesures (voir la partie *III. 2.3. Analyse des dépendances syntaxiques*).

« Les situations situées dans le temps passé et les situations situées au moment de l'énonciation (progressives) seraient immuables, et donc hors du contrôle de l'agent », et interprétées comme épistémiques¹³. Inversement, lorsque le temps de référence est postérieur au temps de la modalité exprimée, l'interprétation déontique serait possible (mais pas nécessaire). Il semblait toutefois trop complexe d'implémenter ce critère de façon efficiente, même en l'approximant.

En ce qui concerne la modification aspectuelle, la télicité sélectionnerait de préférence une interprétation déontique, et l'atélicité une interprétation épistémique. En particulier, le passé simple décrirait les événements en fonction du contexte (déontique), tandis que le passé composé ne donnerait qu'une information stative (épistémique). Comme cela est suggéré dans une autre partie de notre analyse (voir la section *III. 2.2. Pronoms et déictiques : la référence au soi* notamment), nous avons inclus des métriques s'intéressant aux temps verbaux utilisés.

Enfin, la présence de la voix passive favoriserait une interprétation épistémique en indiquant un manque de contrôle de la part du sujet, mais aussi la stativité de la situation décrite, ce qui est encore une fois cohérent avec le lien entre voix active et agentivité supposé plus bas (voir la partie *III. 2.3. Analyse des dépendances syntaxiques*).

Warsnby note finalement que les énoncés contenant *must*, (*maste*), *may* ou *can* en anglais, dans lesquels les propositions sont codées de manière à indiquer un manque de contrôle de la part de l'agent, sont bien interprétés de manière épistémique. Dans les énoncés où les caractéristiques pertinentes indiquent que l'agent visé contrôle la situation dénotée par la proposition, en revanche, l'interprétation préférée est déontique.

Cela étant, nous n'avons en réalité pas pu analyser la modalité de manière compositionnelle comme nous le faisons pour la télicité (*III. 3.1. L'analyse aspectuelle de la télicité*), en la définissant pour chaque clause spécifique, ce en raison d'un manque de temps et d'une trop grande complexification technique. Nous avons donc choisi d'estimer plutôt la modalité apparaissant comme la plus récurrente, au niveau global, dans chaque groupe de répliques, en se focalisant sur

13. *Ibid.*

les verbes modaux, potentiellement modulés par les quelques proxys retenus plus haut. Ceci rend notre approche de la modalité particulièrement limitée, l'interprétation des verbes modaux étant dans une large mesure déterminée par l'ensemble de la construction dans laquelle ils s'inscrivent, à l'échelle de chaque clause.

Il reste que l'on peut supposer de façon raisonnable que ces verbes modaux ont une signification préférentielle. En effet, une gradation intuitive semble exister entre les modaux signifiant une obligation (*devoir, avoir besoin de, être supposé faire* etc. - tous ces éléments atténuant l'agentivité et se référant à un locus de contrôle externe), une possibilité (*pouvoir*, notamment au conditionnel, *être susceptible de* etc.), une capacité (*être capable de, pouvoir* à l'indicatif), un désir (*vouloir* [que quelqu'un fasse quelque chose], *désirer, aimer, souhaiter* au conditionnel notamment etc.) ou une action à venir (*être sur le point de, être en train de*, l'emploi du futur etc. - les marqueurs de futurité se chevauchent avec la modalité car ils impliquent une certaine forme d'incertitude). En outre, les modaux utilisés en référence à l'agent et aux règles ou normes (déontiques), qui nous intéressent particulièrement, semblent précisément revêtir les sens les moins ambigus¹⁴.

Pour autant, et compte tenu de notre seconde question de recherche diachronique, il est nécessaire de s'assurer que ces corrélations sémantiques préférentielles sont à peu près stables dans le temps.

Dans plusieurs de ses études, Hilpert¹⁵ utilise la modélisation de l'espace vectoriel sémantique fondée sur les tokens pour déterminer si différents auxiliaires modaux peuvent être distingués en termes de profils dits « collocationnels » (ce que l'on appelle la « sémantique distributionnelle »), et si différents sens du même auxiliaire présentent des préférences collocationnelles divergentes.

Il constate, pour l'anglais, que les paires quasi-synonymes d'expressions modales, telles que *may* et *might*, ou *must* et *have to*, diffèrent dans leurs caractéristiques distributionnelles. En outre, *may* en particulier aurait subi un changement sémantique continu, s'éloignant des significations modales déontiques et des utilisations textuelles « impliquées » (*involved* en anglais) pour se rapprocher des significations épistémiques et d'un degré plus élevé d'informativité. En bref, les collocats verbaux les plus fortement attirés par *may* au XIX^e siècle se prêtent à l'expression d'un sens permissif, tandis que les collocats les plus fréquents dans les données plus récentes se combinent avec *may* pour exprimer des possibilités.

Plus généralement, les aires sémantiques qui contiennent des verbes aux significations abstraites et épistémiques seraient en augmentation. Cela étant, *must*, par exemple, est encore utilisé relativement plus souvent avec un sens déontique. Aussi, Hilpert conclut qu'en tant qu'éléments hautement grammaticalisés, les auxiliaires modaux anglais présentent une assez grande variabilité collocationnelle, mais que sous la surface de cette variabilité, il existe une multitude d'associa-

14. Nadja Nesselhauf, « Mechanisms of language change in a functional system : The recent semantic evolution of English future time expressions », *Journal of Historical Linguistics*, 2-1 (25 juill. 2012), p. 83-132, DOI : [10.1075/jhl.2.1.06nes](https://doi.org/10.1075/jhl.2.1.06nes)

15. Martin Hilpert, « Change in modal meanings : Another look at the shifting collocates of *may* », *Constructions and Frames*, 8-1 (29 sept. 2016), p. 66-85, DOI : [10.1075/cf.8.1.05hil](https://doi.org/10.1075/cf.8.1.05hil) ; M. Hilpert et Susanne Flach, « Disentangling modal meanings with distributional semantics », *Digital Scholarship in the Humanities*, 36-2 (29 sept. 2021), p. 307-321, DOI : [10.1093/llc/fqaa014](https://doi.org/10.1093/llc/fqaa014)

tions qui lient conventionnellement une expression modale donnée à un ensemble particulier d'éléments lexicaux, ou un sens modal spécifique à un contexte linguistique particulier. Ainsi, bien qu'il existe des changements sémantiques, Hilpert peut affirmer que *might*, *must* et *could* ont globalement plus de potentiel épistémique, tandis que *can* et *will* sont stables en termes de sens et déontiques. Pour sa part, *would* demeure assez ambivalent. Alors que *will* n'a pas d'équivalent direct en français, l'analyse du modal *pouvoir* (équivalent de *can*) semblait donc particulièrement pertinente pour notre question, en considérant qu'il possède, en français également, un sens déontique relativement stable.

Mais pour complexifier encore l'analyse, la corrélation de ces verbes modaux avec le degré d'agentivité personnelle, sur la base de la gradation mentionnée plus haut, n'est en réalité pas évidente. La tragédie grecque illustre bien cette inéquation. Bien qu'il s'agisse d'un cas limite du point de vue du locus de contrôle des personnages, apparemment manifestement externe puisque ces derniers sont soumis à la contingence divine (bien que l'on trouve déjà, dans certaines tragédies, des expressions ambiguës telles que le « désir de sacrifice » d'Agamemnon¹⁶), d'aucuns pourrait considérer que ces mêmes personnages identifient en réalité les vexations divines à leur propre destin (faisant tout le tragique de la pièce, précisément). De fait, ils construisent souvent des récits *a posteriori* pour donner un sens à leurs souffrances, et expriment en ce sens une capacité à interpréter cette contingence dans des termes personnels, caractéristique d'un haut niveau d'agentivité.

De la même façon, comme le suppose Rymes¹⁷ dans une étude socio-linguistique s'attachant à analyser le discours de jeunes adolescents ayant abandonné l'école secondaire dans les années 90, pour continuer à se considérer comme de bonnes personnes malgré les actes criminels qu'ils ont commis, ces mêmes jeunes atténuent paradoxalement l'agentivité réelle induite par leur langage, principalement par l'utilisation de verbes modaux associés à un locus de contrôle externe, dans un acte qui pourrait pour sa part être considéré comme particulièrement agentif (ce que Rymes appelle « l'agentivité morale »).

C'est en ce sens que nous avons décliné notre métrique s'attachant aux verbes modaux de différentes manières. En trouvant les équivalents français les plus directs par recoupement de grammaires françaises bien connues et accessibles en ligne¹⁸, nous avons tout d'abord inclus une métrique comptant le nombre relatif de verbes modaux globalement utilisés, comme approximation de l'agentivité discursive ou morale des personnages considérés, ce quel que soit le contrôle sur l'action sémantiquement impliqué par le verbe modal en question. Nous avons en réalité calculé deux versions différentes de cet indicateur, l'une restrictive se concentrant sur les modaux d'obligation/permission et de désir (« pouvoir », « devoir », « vouloir », « falloir + infinitif ») et l'autre plus large (contenant également « espérer », « savoir », « penser », « aller » + infinitif).

Nous avons également, au moyen d'une nouvelle métrique, opposé l'utilisation personnelle

16. vernant_j-p__vidal-naquet_p_mythe_2005

17. Betsy R. Rymes, « The Construction of Moral Agency in the Narratives of High-School Drop-Outs », *Discourse & Society*, 6 (1995), p. 495-516, URL : <https://api.semanticscholar.org/CorpusID:145568427>

18. Martin Beaudoin Jacqueline Ollivier, *Grammaire Française*, 2007 ; Maurice Grevisse Cédric Fairon, *Le Petit bon usage de la langue française : Grammaire*, 2018 ; Yvonne Delatour, *Nouvelle grammaire du français : Cours de civilisation française de la Sorbonne*, 1991

des modaux (à savoir ceux dont le sujet est à la première personne) à leur utilisation globale. En effet, cette utilisation « structurelle » des modaux, qui ne tient pas compte de leur contenu sémantique, pourrait n'être pertinente que lorsqu'on parle de soi (l'agentivité discursive ou morale se restreignant potentiellement aux seuls récits identitaires).

A cela, nous avons enfin ajouté une métrique calculant le rapport entre les modaux traduisant un locus de contrôle interne (« pouvoir », « vouloir » + infinitif), et externe respectivement (« devoir », « falloir » + infinitif), étant donné leur connotation agentive plutôt univoque et stable dans le temps, en tenant compte cette fois du contrôle sémantiquement impliqué par le modal employé.

L'idée d'une utilisation des verbes modaux d'abord « performative », et en cela partiellement indépendante de leur contenu sémantique, est d'ailleurs corroborée par l'observation de tendances globales dans l'utilisation des verbes modaux, ayant affecté chacun d'entre eux de façon indistincte. De nombreux articles^{19 20 21 22} ont en effet constaté un déclin général de la fréquence d'utilisation des verbes modaux au cours du XX^e siècle au moins, bien que ce phénomène pourrait être compensé par une augmentation globale parallèle de l'utilisation des verbes semi-modaux en anglais (à l'exception de *have got to*). En particulier, ces travaux observent que *will* et *must* sont les verbes qui diminuent le plus drastiquement, tandis que l'utilisation de *need* et *be going to* (semi-modaux) augmente. Il s'agirait donc plutôt d'un glissement que d'une disparition pure et simple de la modalité. Pour autant, ce dernier souligne la limite de notre métrique comptant simplement le nombre total de verbes modaux utilisés, qui ne saurait se suffire à elle-même à l'heure où l'expression de la modalité se ferait plus discrète.

Plus problématique encore pour notre analyse, et comme le suggèrait Hilpert, il semblerait que l'on assiste au déclin du sens « intentionnel » (en anglais *'speaker's intention'-sense*) des modaux, notamment pour *will*, ce en faveur d'un sens de « pure prédiction » (*'pure-prediction' sense*)²³, sur la base de ses seules interprétations cependant, et d'un échantillon restreint d'expressions analysées), ce qui est cohérent avec le Bybee-path, aussi appelé processus de « grammaticalisation », une tendance inhérente à la langue dans son ensemble. Selon Bybee²⁴, en cherchant à exprimer un sens plus spécifique, une langue finirait paradoxalement par « blanchir » (*bleaching* en anglais), autorisant alors une plus grande acceptabilité contextuelle avec le temps, de sorte que de nombreuses significations se retrouvent à co-occuper dans un seul mot. Ainsi, plus une expression temporelle future est ancienne par exemple, plus elle exprimerait fréquemment une prédiction non colorée par la caractéristique de l'intention. Le déclin de l'expression de l'intention

19. John Myhill, « Change and continuity in the functions of the American English modals », *ling*, 33-2 (1995), p. 157-212, DOI : [10.1515/ling.1995.33.2.157](https://doi.org/10.1515/ling.1995.33.2.157)

20. Nicholas Smith, « Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English », dans *Modality in Contemporary English*, dir. Roberta Facchinetti, Frank Palmer et Manfred Krug, Berlin, Boston, 2003, p. 241-266, DOI : [doi:10.1515/9783110895339.241](https://doi.org/10.1515/9783110895339.241)

21. Geoffrey Leech, Marianne Hundt, Christian Mair et N. Smith, « Change in Contemporary English : A Grammatical Study » ()

22. Stig Johansson, « Modals and semi-modals of obligation in American English : some aspects of developments from 1990 until the present day », dans *The Verb Phrase in English*, dir. Bas Aarts, Joanne Close, Geoffrey Leech et Sean Wallis, 1^{re} éd., 2013, p. 372-380, DOI : [10.1017/CB09781139060998.016](https://doi.org/10.1017/CB09781139060998.016)

23. N. Nesselhauf, « Mechanisms of language change in a functional system... »

24. Bybee, « The evolution of grammar... »

du sujet et du locuteur avec les marqueurs de futur résulterait donc de changements successifs entre les expressions de temps futur établies, et les nouvelles constructions venant prendre le relais pour exprimer les significations en question à mesure que les expressions de temps futur plus anciennes, du moins celles qui proviennent du matériel lexical, évoluent sur le chemin de Bybee pour exprimer la « prédiction pure ».

Aussi, comme le note Nesselhauf²⁵, les deux types de changements sont probablement motivés, au moins en partie, par des transformations socioculturelles et stylistiques. En particulier, une explication souvent avancée pour rendre compte de ce moindre usage des verbes modaux, ou du moins de leur moindre explicitation des intérêts du locuteur, est celle de la « démocratisation de l'anglais » (les études ne s'intéressent là encore qu'à l'anglais seul). Cette dernière fait référence à une « montée d'alternatives plus conviviales et moins menaçantes, dans une société apparemment plus égalitaire, démocratique et antiautoritaire » qui conduirait à une « tendance des locuteurs à éviter les modes d'interaction inégaux », notamment les marqueurs manifestes d'autorité dans les énoncés véhiculant une directive (Farrelly et al., 2012). Cela expliquerait la préférence pour les modaux exprimant une obligation médiane, comme *need to*, afin d'éviter les marqueurs de pouvoir excessifs tels que *must*. Une construction de la forme *X needs you to* serait ainsi interprétée comme un appel plutôt que comme un ordre direct, psychologisant par-là les directives (certains parlent de « pseudo-démocratisation »).

D'autres phénomènes pourraient encore soutenir, selon Nesselhauf, le déclin du sens intentionnel de *will* en particulier. Tout d'abord, une expression qui n'est pas traditionnellement une expression de futurité mais avec un sens intentionnel similaire pourrait avoir remplacé l'utilisation d'expressions du futur avec ce sens dans certains contextes. En effet, il semble qu'en anglais, *be going to* et le *present progressive* aient remplacé *will* et *shall*, dans une certaine mesure, pour exprimer une prédiction fondée sur l'intention du sujet. En ce sens, nous avons inclus une métrique mesurant l'utilisation du progressif, correspondant en français à l'expression « être en train de [faire quelque chose] ». En outre, la construction *want to* est identifiée par Nesselhauf comme un marqueur de futurité émergent, qui a commencé à être utilisé pour des prédictions fondées sur l'intention du sujet.

Autre possibilité avancée par l'auteur, ce déclin pourrait en partie s'expliquer par un changement stylistique global : alors que dans les exemples du XVIII^e siècle, le sujet des clauses intentionnelles (avec *shall*) diffère du locuteur, de sorte que seule l'intention du locuteur et non celle du sujet est exprimée, en anglais actuel, il semble beaucoup plus probable que l'on choisisse une construction où le locuteur et le sujet sont totalement ou partiellement identiques (c'est-à-dire avec *I* ou *we*, en anglais, comme sujet – l'auteur compare par exemple *My friendship to you shall be steady* avec *I'm always going to be your friend*). Et cela refléterait une tendance plus générale à passer de modes d'expression impersonnels à des modes d'expression plus personnels dans de nombreux types de textes. La diminution globale de la construction passive va également dans ce sens²⁶ ; et il est intéressant de constater que c'est principalement au théâtre et dans la

25. N. Nesselhauf, « Mechanisms of language change in a functional system... »

26. G. Leech, M. Hundt, C. Mair, et al., « Change in Contemporary English : A Grammatical Study »...

conversation fictionnelle que le déclin a eu lieu²⁷.

Enfin, comme l'indique Nesselhauf, « les personnages de fiction pourraient être de plus en plus représentés comme se sentant davantage déterminés par les circonstances extérieures et comme ayant le sentiment que leurs intentions personnelles et celles des autres ont moins d'influence sur ce que l'avenir pourrait leur apporter. Et cela pourrait refléter une évolution plus générale, qui ne se limite pas à la représentation littéraire : à mesure que le réseau d'interconnexions dans la société devient de plus en plus complexe et dense (cf., par exemple, Elias 2000), l'intention propre d'un sujet pourrait avoir moins d'influence sur les événements futurs ou du moins pourrait amener les sujets à supposer que c'est le cas ». Cette affirmation est particulièrement contradictoire avec nos propres hypothèses. Néanmoins, il faut d'abord noter que dans l'analyse de Nesselhauf, ce sentiment potentiel de moindre contrôle sur les événements futurs ne s'étend en réalité pas aux prédictions fondées sur des arrangements pour lesquels le sujet a été personnellement impliqué, ce qui semble difficilement explicable selon la lecture précédente.

A partir de là, et en maintenant la validité de notre analyse modale, nous proposons une interprétation alternative : plutôt qu'une moindre futurité dans l'intention moderne, qui dénoterait une perte réelle ou perçue de contrôle sur les événements futurs, on peut au contraire considérer que le futur est de plus en plus infusé par l'intention de l'agent, ce dernier étant de plus en plus capable de planifier. Le passage de modaux (y compris *will*) précédemment dévolus à l'expression de la seule intention à des marqueurs de futurité, laissant à *will* une connotation épistémique plus neutre, apparaît sous ce nouvel éclairage particulièrement évocateur. Suivant un principe d'économie lexicale, et l'intention étant supposément devenue une dimension psychologique de plus en plus pervasive, il semblerait naturel que la langue l'incorpore dans sa structure même, libérant ainsi de l'espace sémantique pour d'autres nuances moins directement évidentes pour les interlocuteurs, dans la logique du *foregrounding* mentionné plus tôt.

Pour autant, cela limite à nouveau la portée de notre analyse modale. Si nous n'avons plus autant besoin de la modalité pour souligner la dimension intentionnelle d'une proposition, il se pourrait que l'utilisation de verbes modaux, aussi bien internes qu'externes, n'augmente pas avec le temps. Notre analyse modale vise donc dans un premier temps à déterminer s'il existe une différence significative entre les personnages de haut et de bas statut dans l'utilisation des verbes modaux, tandis que le sens de la corrélation, si corrélation il y a, demeure théoriquement incertain. C'est également pour cette raison que nous avons inclus dans nos métriques, et calculé séparément, l'emploi du temps futur.

Un autre marqueur modal plus secondaire, et pouvant traduire un manque d'assertivité, est l'ellipse, elle aussi ajoutée au calcul de notre indice.

27. toutes ces observations s'appuient sur le Corpus ARCHER : <https://www.projects.alc.manchester.ac.uk/archer/>

Chapitre 6

Les proto-marqueurs agentifs

De tout ce qui précède, l'on comprend qu'il semblait pertinent de donner la prévalence aux marqueurs d'agentivité les plus « typiques » pour construire notre indice, et dont la corrélation avec la notion d'agentivité était *a priori* indépendante de la composition des clauses dans lesquels ils s'insèrent.

6.1 Le verbe, un indicateur ambigüe

Le verbe peut être considéré comme le marqueur linguistique prototypique de l'agentivité¹, dénotant une disposition objective à l'action ou, plus implicitement, un locus de contrôle interne. Une composante de notre indice agentif a donc consisté à mesurer la proportion relative de verbes dans chaque groupe de répliques.

Néanmoins, et étant donnée notre acception relativement large de la notion d'agentivité, la direction de la corrélation pour ce proto-marqueur est en réalité incertaine. En effet, il s'agit également d'une partie du discours à connotation plus « concrète »², à laquelle des agents « haut niveau » pourraient préférer l'utilisation de parties du discours plus abstraites (voir la partie *III. 3.3. Les marqueurs corrélés d'abstraction*).

6.2 Pronoms et déictiques : la référence au soi

Comme mentionné au chapitre I., le choix d'un pronom à la première personne rapproche notre attention du soi³, par rapport à un pronom à la troisième personne (bien qu'il faille probablement faire une distinction entre les récits indirects à la troisième personne et ceux utilisant le style indirect libre, qui semble largement compenser la distance induite par le choix du pronom au niveau syntaxique et expressif).

1. M. Formanowicz, Janin Roessel, Caterina Suitner et Anne Maass, « Verbs as linguistic markers of agency : The social side of grammar. » *European Journal of Social Psychology*, 47–5 (2017), Place : US Publisher : John Wiley & Sons, p. 566-579, DOI : [10.1002/ejsp.2231](https://doi.org/10.1002/ejsp.2231)

2. Francesca Strik Lievers, Marianna Bolognesi et Bodo Winter, « The linguistic dimensions of concrete and abstract concepts : lexical category, morphological structure, countability, and etymology », *Cognitive Linguistics*, 32–4 (25 nov. 2021), p. 641-670, DOI : [10.1515/cog-2021-0007](https://doi.org/10.1515/cog-2021-0007)

3. Cindy Chung et James Pennebaker, « The Psychological Functions of Function Words » ()

Néanmoins, la dimension orale du théâtre ne donnait pas accès au critère de choix d'une personne par l'auteur, pour faire parler son personnage.

Cela étant, il restait possible de comparer l'utilisation relative des pronoms personnels individuels (première et deuxième personne du singulier) et collectifs (première et deuxième personne du pluriel) par les personnages-mêmes, en supposant qu'un ratio en faveur des premiers dénote une plus grande focalisation sur soi, et ainsi une agentivité plus « haut niveau ».

Dans la même logique, le choix du temps principal du récit peut être interprété à l'aune du degré d'identification au personnage souhaité par le lecteur, et ultimement par l'auteur (dans la mesure où il répond à une demande du public, mais aussi parce qu'il peut lui-même chercher à vivre son personnage de manière plus intime), traduisant ainsi une plus ou moins grande focalisation sur soi. Là encore, la spécificité du genre théâtral obligeait à se limiter à l'analyse des temps utilisés par les personnages-mêmes, qui ont été comptés séparément (futur, présent, passé simple et imparfait - à savoir tous ceux dont nous disposons compte tenu des outils linguistiques utilisés).

Les déictiques constituent un autre point d'intérêt linguistique. Ceux-ci désignent des termes qui ne prennent leur sens qu'en fonction de la situation d'énonciation dans laquelle ils sont utilisés (c.a.d. le locuteur, le lieu et le moment d'énonciation), et dont la distance induite entre ce qui est mentionné et le « centre déictique » (le locuteur) peut être plus ou moins proche. Intuitivement, l'utilisation de déictiques proximaux (plus proches de soi) plutôt que distaux semble devoir être corrélée à une plus grande agentivité, toujours au sens d'une plus grande identification « à » l'action. Nous avons ainsi comparé l'utilisation relative, par chaque groupe de personnages, des marqueurs déictiques français proximaux « ceci », « voici », « -ci », « ici », « (-)deçà », « maintenant » ; par rapport aux marqueurs distaux « cela », « voilà », « (-)là », « (-)delà » et « alors ».

6.3 Analyse des dépendances syntaxiques : agentivité et voix grammaticales, verbes statifs, transitivité

6.3.1 Agentivité grammaticale

D'un point de vue linguistique, la notion d'Agent, entendue comme rôle sémantique (voir I.2.1. *Transitivité et foregrounding : des indices corrélés de l'efficacité de l'action*) est essentiellement portée par le verbe et ne peut être réduite à la seule présence d'un sujet humain. Cependant, il semble qu'il y ait une corrélation constante entre le type de sujet utilisé et l'animation (*animacy*) du participant, cette dernière étant une condition préalable à toute forme d'agentivité. Une façon d'approcher l'agentivité des locuteurs, en évitant toute analyse sémantique, consistait donc à examiner leur degré d'animation, lui-même véhiculé par le type de sujet utilisé. De fait, Ahearn⁴ constate l'existence d'un continuum récurrent d'« animation » allant des pronoms de première et deuxième personne (animation élevée) aux référents inanimés exprimés par des

4. Laura M. Ahearn, « Language and Agency », *Annual Review of Anthropology*, 30-1 (1^{er} oct. 2001), Publisher : Annual Reviews, p. 109-137, DOI : [10.1146/annurev.anthro.30.1.109](https://doi.org/10.1146/annurev.anthro.30.1.109)

expressions nominales indéfinies (animation faible), comme synthétisé dans le graphique suivant :

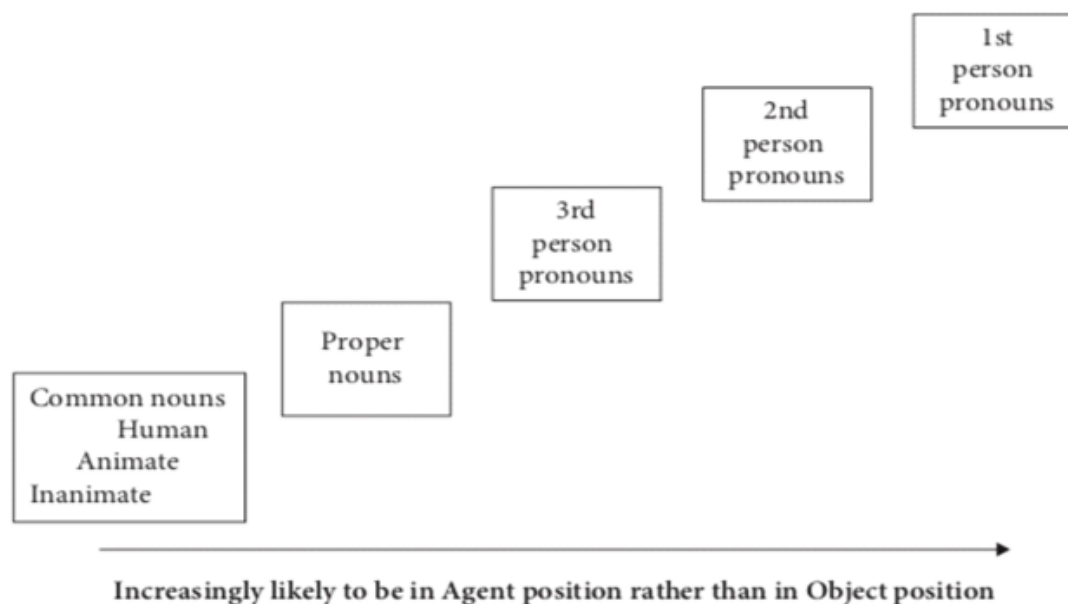


FIGURE 6.1 – Représentation des dénominations personnelles en fonction du degré d'*animation* qu'elles évoquent - Ahearn (2001)

Là encore, la forme orale du théâtre rendait impossible l'utilisation de ce marqueur tel quel pour estimer l'agentivité du locuteur, puisque tous les personnages parlent à la première personne. Nous pouvions toutefois mesurer la proportion relative de pronoms à la première personne utilisés en position sujet, ou explicitement comme agents à la voix passive (en utilisant le *parser* syntaxique de spaCy) dans chaque groupe de répliques, afin d'approximer très grossièrement leur agentivité grammaticale respective.

Pour la suite de l'analyse, et compte tenu de cette approximation très imparfaite, nous ne nous sommes pas limités à l'analyse des seules clauses dites « agentives » (celles comprenant un Agent au sens grammatical ou sémantique), mais nous avons plutôt dupliqué la mesure des indicateurs pour lesquels il était particulièrement difficile de savoir si leur relation avec l'agentivité dépendait du fait que le locuteur soit ou non le sujet de la clause, de la même façon que pour les verbes modaux. Nous avons ainsi dupliqué les mesures suivantes : verbes, copules, ratio de l'utilisation de la voix active vs. passive.

A titre de remarque, la gradation mentionnée ci-dessus a encore été utilisée dans l'analyse de la télélicité : un personnage plus agentif pourrait en effet également être amené à prêter plus d'agentivité aux autres personnages qu'il ou elle mentionne, et ainsi à utiliser relativement plus de pronoms dénotant une animation particulière y compris lorsqu'il ou elle se réfère à d'autres individus. Il semble de fait qu'il existe un biais cognitif généralisé favorisant l'interprétation télélique d'une phrase lorsqu'elle contient un sujet animé, et inversement une interprétation atélique dans le cas d'un sujet inanimé⁵. De manière cohérente, de nombreuses recherches indiquent que

5. Tim Graf, Markus Philipp, Xiaonan Xu, Franziska Kretzschmar et Beatrice Primus, « The in-

les entités inanimées sont incompatibles (ou du moins relativement moins compatibles que les entités animées) avec les verbes qui connotent l'intentionnalité, la planification ou impliquent des processus cognitifs, car ces entités inanimées n'ont pas les moyens de mener à bien des actions intentionnelles (voir ci-après).

6.3.2 Voix active vs. passive

Dans le même ordre d'idée, bien que cette fois-ci non soutenue par des travaux antérieurs, l'utilisation d'une voix active plutôt que passive (et malgré l'agentivité grammaticale qu'implique cette dernière) nous a semblé devoir être corrélée avec un degré d'agentivité plus élevé ⁶.

6.3.3 Verbes statifs

Un type de verbe pourrait constituer une exception univoque au lien verbe-agentivité : les verbes statifs ou d'état, qui se réfèrent principalement aux états émotionnels (et potentiellement durables) des agents, plutôt qu'à des actions. Il est vrai que l'on peut considérer les verbes exprimant des attitudes et des sentiments propositionnels comme évoquant des formes d'action mentale. Cependant, les *verba sentiendi* (verbes de perception) sont généralement considérés comme moins « actifs ». Une autre composante linguistique de notre indice agentif a donc consisté à comparer la proportion relative de verbes purement statifs (ici approximés par l'utilisation des verbes *être* et *avoir* en excluant les emplois temporels, ainsi que par l'étiquette « copule » du *parser* spaCy, les deux métriques étant mesurées séparément) par rapport aux verbes dits « dynamiques » (tous les autres verbes).

En guise de commentaire, et malgré leur supposée « stativité », l'utilisation du progressif avec des verbes d'émotions comme *aimer* (par opposition à *savoir*) semble en réalité être de plus en plus autorisée dans le langage courant, au moins à partir du XIX^e siècle pour l'anglais ⁷. Par ailleurs, Granath et al. notent qu'une telle combinaison entraînerait un changement aspectuel, suggérant alors une « agentivité accrue ». Cela remet en cause l'idée même d'une séparation binaire entre verbes statifs et non statifs, ces derniers se plaçant plutôt sur un continuum. De la même façon en effet, certains verbes appartenant traditionnellement à la catégorie des verbes dynamiques, notamment en référence à des activités et des événements physiques, se produisent rarement avec l'aspect progressif (par exemple *shrug*, *smash*, *suck* et *throw* en anglais). Notre analyse binaire est donc limitée, mais une approche tripartite, voire par degré de stativité, néces-

teraction between telicity and agentivity : Experimental evidence from intransitive verbs in German and Chinese », *Lingua* (, oct. 2017), DOI : [10.1016/j.lingua.2017.08.006](https://doi.org/10.1016/j.lingua.2017.08.006) - il s'agit là d'une études psycholinguistique, et non simplement théorique

6. P. Hopper et S. Thompson, « Transitivity in Grammar and Discourse »...

7. Solveig Granath et Michael Wherrity, « "I'm loving you - and knowing it too" : Aspect and so-called stative verbs », dans 2014, URL : <https://api.semanticscholar.org/CorpusID:142062193>

sitait de recourir à de la modélisation^{8 9 10}, ce qui dépassait le cadre technique de cette analyse. Aussi, nous avons tout de même mesuré l’usage relatif du progressif, considéré comme une indication potentiellement plus fine que l’opposition verbes statifs/autres verbes du dynamisme connoté par les événements, et indirectement de l’agentivité du locuteur (comparez simplement « je mange » avec « je suis en train de manger »).

6.3.4 Transitivité

Comme mentionné plus haut (I.2.1. *Transitivité et foregrounding : des indices corrélés de l’efficacité de l’action*), la transitivité semble assez intuitivement liée à une perspective plus agentive, dans la mesure où le locuteur conçoit la possibilité d’une action directe sur un objet (grammatical). Une mesure de la proportion relative des verbes utilisés transitivement, plus proches de notre prototype de causalité¹¹, a donc logiquement été ajoutée à notre indice.

8. Judith L. Klavans et Martin Chodorow, « Degrees of Stativity : The Lexical Representation of Verb Aspect », dans *COLING 1992 Volume 4 : The 14th International Conference on Computational Linguistics*, 1992, URL : <https://aclanthology.org/C92-4177>

9. Annemarie Friedrich et Alexis Palmer, « Automatic prediction of aspectual class of verbs in context », dans *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)*, Baltimore, Maryland, 2014, p. 517-523, DOI : [10.3115/v1/P14-2085](https://doi.org/10.3115/v1/P14-2085)

10. Matthew Sims, Jong Ho Park et David Bamman, « Literary Event Detection », dans *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy, 2019, p. 3623-3634, DOI : [10.18653/v1/P19-1353](https://doi.org/10.18653/v1/P19-1353)

11. Abigail W. Konopasky et Kimberly M. Sheridan, « Towards a Diagnostic Toolkit for the Language of Agency », *Mind, Culture, and Activity*, 23-2 (2016), Publisher : Routledge _eprint : <https://doi.org/10.1080/10749039.2015.1128952>, p. 108-123, DOI : [10.1080/10749039.2015.1128952](https://doi.org/10.1080/10749039.2015.1128952)

Chapitre 7

Traverser les distances psychologiques : l'agentivité et la planification

Les gens sont capables de penser à l'avenir, au passé, à des endroits éloignés, à la perspective d'une autre personne et à des alternatives contrefactuelles. Sans nier le caractère unique de chaque processus, Trope et al.¹ ont proposé qu'ils constituent des formes différentes de franchissement d'une certaine distance psychologique. Selon eux, la distance psychologique est égocentrique : son point de référence est le soi, ici et maintenant (voir sur ce point la partie *III. 2.2. Pronoms et déictiques : la référence au soi*), et les différentes façons dont un objet peut être éloigné de ce point (dans le temps, dans l'espace, dans la distance sociale et dans l'hypothéticité) constituent différentes dimensions de la distance.

Aussi, plus les événements sont éloignés de l'expérience directe, plus ils sont susceptibles d'être représentés en des termes abstraits traduisant l'essence perçue des événements (constructions de haut niveau), plutôt qu'à travers les détails plus concrets et accessoires des événements (constructions de bas niveau), ce qui correspond étroitement à notre définition du continuum agentif. Cela est d'ailleurs corroboré par l'association, trouvée chez Woike², entre agentivité et différenciation (qui s'oppose au processus cognitif d'intégration³) :

« Parce que la différenciation implique de percevoir des différences et l'indépendance entre les attributs des stimuli sociaux, elle peut être utilisée plus fréquemment par les personnes qui ont une orientation vers l'individualisme et l'indépendance, car elle

1. Yaacov Trope et Nira Liberman, « Construal-level theory of psychological distance. » *Psychological Review*, 117–2 (2010), Place : US Publisher : American Psychological Association, p. 440-463, DOI : [10.1037/a0018963](https://doi.org/10.1037/a0018963)

2. Barbara A. Woike, « The use of differentiation and integration processes : Empirical studies of "separate" and "connected" ways of thinking. » *Journal of Personality and Social Psychology*, 67–1 (1994), Place : US Publisher : American Psychological Association, p. 142-150, DOI : [10.1037/0022-3514.67.1.142](https://doi.org/10.1037/0022-3514.67.1.142)

3. Les termes « séparé » et « connecté » ont souvent été utilisés pour décrire deux conceptions différentes de l'individu par rapport au monde social : l'une comme autonome, indépendante et séparée des autres, et l'autre comme relationnelle, interdépendante et indépendante des autres. Pour Woike, cela suggère qu'il existe des processus cognitifs spécifiques impliqués dans la promotion et le maintien de ces deux orientations : la différenciation, qui consiste à percevoir un certain nombre d'aspects différents dans un groupe de stimuli donné, et l'intégration, qui consiste à établir des connexions ou des liens entre des stimuli différenciés

remplit une fonction particulière pour elles. Par exemple, Belenky et al. (1986) ont suggéré que la distance psychologique peut être maintenue en séparant le soi des objets sociaux afin de percevoir les choses de manière objective ou comme n'ayant aucun rapport avec soi. En outre, la pensée critique et la prise de décision sont, par définition, des procédures de comparaison qui impliquent de voir les différences relatives entre les choses. L'acte de séparer une question en la décomposant en plusieurs parties et en percevant ses contrastes et ses distinctions peut aider les percepteurs à porter des jugements et à prendre des décisions en limitant leur implication personnelle dans le contexte interpersonnel, ce qui leur permet de conserver une orientation individualiste. »

Le degré de planification ou de capacité rétrospective de l'agent évoqués par une clause est ainsi une autre caractéristique linguistique importante.

7.1 L'analyse aspectuelle de la télélicité

La télélicité, ou « délimitation aspectuelle », est la propriété aspectuelle d'une phrase verbale (ou de la phrase dans son ensemble) qui indique qu'une action ou un événement a une fin spécifique. Du grec *telos*, elle désigne plus précisément la « présence d'une finalité, intrinsèque ou envisagée, inhérente à l'événement, qui se manifeste mutuellement dès le moment où la situation a commencé et est évoquée », et au-delà de laquelle l'événement ne peut plus se poursuivre⁴. A titre d'exemple, « courir vers quelqu'un » est considéré comme ayant un aspect plus télélique que « jouer dans l'herbe » ou encore « marcher près d'un lac ».

Par définition, la télélicité est donc liée à l'intentionnalité de l'agent. La manière dont un événement est évoqué (et réciproquement interprété) comme étant doté d'une valeur aspectuelle plus ou moins télélique peut donc nous informer en partie quant à la mesure dans laquelle le locuteur se représente lui-même, ou ses *alter ego*, comme capable d'agir volontairement sur le monde, par opposition à une idéologie plus fataliste par exemple. La télélicité a donc également avoir, et ce d'une manière encore plus structurelle et compositionnelle, avec la notion d'agentivité.

En nous inspirant de modèles d'apprentissage automatique de la télélicité d'une clause (cités ci-après), qui donnaient accès aux principales caractéristiques utilisées par l'algorithme de classification (et dont la pertinence pour un lecteur humain était systématiquement vérifiée par la théorie linguistique ou par des études psycholinguistiques ultérieures), nous avons défini un certain nombre d'éléments linguistiques qui semblaient être décisifs dans la détermination de la télélicité d'une phrase. Cette fois, et contrairement à la modalité, l'analyse se fait au niveau de la phrase-même (voire de la clause), puisque la télélicité est définie de manière intrinsèquement compositionnelle (de telle sorte qu'une approximation au niveau des répliques concaténées des personnages n'avait pour le coup pas grand sens). En effet, si elle va souvent de pair avec les prédicats de réalisation et d'accomplissement, pour reprendre la classification de Vendler⁵, et s'il

4. Ilse Depraetere, « (A)telicity and intentionality », *Linguistics*, 45-2 (20 janv. 2007), DOI : [10.1515/LING.2007.008](https://doi.org/10.1515/LING.2007.008)

5. Zeno Vendler, « "Verbs and Times." » *The Philosophical Review*, vol. 66, no. 2, 1957, pp. 143-160

semble assez difficile d'exprimer un événement télique avec un prédicat de type activité ou état, les premiers ne peuvent garantir la télicité d'un événement.

En anglais comme en français, la télicité (un aspect interne) n'est pas non plus représentée par un marquage morphologique, contrairement à l'aspect externe (par exemple le progressif), bien qu'une corrélation ait été établie entre télicité et aspect parfait d'une part, et atélicité et progressif d'autre part ⁶.

Il existe cependant certains critères qui détermineraient automatiquement la télicité d'une clause, et qui constituent donc des tests de télicité souvent mentionnés dans les articles théoriques. Notamment, le test de l'adverbe de temps demande si une phrase verbale est compatible avec le syntagme prépositionnel « dans un temps X » (télique) ou « pendant un temps X » (atélique). Un autre test fréquemment cité est celui de la compatibilité de la clause avec un adverbe d'intention tel que *volontairement*.

Concrètement, nous avons donc défini un score « télique » pour chaque phrase, variant entre 0 (score initial) et 1 (phrase univoquement télique). Sur cette base, notre code teste tout d'abord la présence des marqueurs « automatiques » suivants au niveau de la phrase (nous n'avons en réalité pas été en mesure d'effectuer une décomposition syntaxique au niveau de la clause *per se*, mais cette dernière est approximée par la présence d'un verbe, comme expliqué ci-après).

Notons qu'une fois encore, nous avons tenté de ne retenir que les déterminants les plus structurels (non sémantiques) pour conclure à la télicité d'une phrase, afin de contrôler autant que possible pour un biais d'interprétation moderne, qui pourrait voir un aspect télique dans des formes lexicales qui n'en évoquaient en réalité aucun chez les lecteurs contemporains de l'œuvre (le degré d'agentivité étant précisément censé s'accroître avec le temps).

Les critères automatiques sont donc les suivants : a) « en X temps », qui rend la phrase systématiquement télique et fixe son score à 1 ; et « pendant X temps », qui la rend atélique et fixe son score final à 0.

b) la présence d'un adverbe intentionnel (critère d'ailleurs limité dans son application, puisque le nombre d'adverbes univoquement volitifs est restreint), auquel cas le score phrastique est égal à 1 (voir la section I.1.2. *Les adverbes intentionnels, statut hybride ?*).

c) un ajout personnel, mais étroitement lié au test d'intentionnalité présenté ci-dessus : les expressions de but (dont nous avons dressé une liste aussi exhaustive que possible, de la même manière que pour les adverbes intentionnels), pour lesquelles la présence d'au moins une

(), DOI : <https://doi.org/10.2307/2182371> : Les activités et les accomplissements se distinguent des réalisations et des états dans la mesure où les deux premières permettent l'utilisation d'aspects continus et progressifs. Les activités et les accomplissements se distinguent les unes des autres par leur délimitation. Les activités n'ont pas de point final (un point avant lequel l'activité a eu lieu et après lequel elle ne peut plus continuer, comme pour « Jean a dessiné un cercle »), alors que les accomplissements en ont un. En ce qui concerne les réalisations et les états, les réalisations sont instantanées, mais les états sont durables. Les réalisations et les accomplissements se distinguent l'un de l'autre par le fait que les réalisations ont lieu immédiatement (comme dans « reconnaître » ou « trouver »), alors que les accomplissements s'approchent progressivement d'un point final (comme dans « peindre un tableau » ou « construire une maison »).

6. William Frawley, *Linguistic semantics*. Pages : xvii, 533, Hillsdale, NJ, US, 1992 (Linguistic semantics.) : cette étude est purement théorique

expression fixe également le score à 1.

TABLE 7.1 – Expressions à valeur de but retenues

« pour », « finalité », « afin de/que », « à dessein de », « dans l’objectif/le but/l’intention/le dessein de », « de façon à/de cette façon/de telle façon que », « histoire de/que », « à la (seule) fin de »

Si la phrase contient au moins l’un de ces marqueurs, en considérant *a priori* comme marginaux les cas de combinaisons contradictoires entre ces derniers (quant à leur évocation téléique ou atéléique), l’analyse s’arrête ici pour la phrase en question, et le code nous fait passer à la suivante.

Si la phrase ne contient pas de tels critères « automatiques », le score de la phrase reste à 0, et l’analyse passe à un niveau plus fin : le code s’exécute pour chaque verbe de la phrase (et l’arborescence syntaxique de spaCy est telle qu’il n’était pas possible de limiter l’analyse aux verbes étiquetés comme « ROOT », sans perdre une partie de l’information qui nous intéressait pour la téléicité). Tous les verbes sont donc analysés, y compris les verbes à l’infinitif qui dépendent d’un verbe conjugué (dans ce cas, certaines étiquettes de dépendance intéressantes sont associées au verbe à l’infinitif en question, et non au « ROOT »). Le code ajoute ensuite +1, ou respectivement -1, au score de téléicité du semblant de clause (à savoir du verbe en question), également initialisé à 0, chaque fois qu’il rencontre l’un des marqueurs linguistiques suivants, sélectionnant une interprétation téléique ou atéléique :

1. Comme mentionné ci-dessus, un temps parfait sélectionnerait une interprétation téléique. Compte tenu des étiquettes disponibles, nous avons inclus le passé simple (présente dans le modèle fr de pie-extended) et le passé composé (identifié par la présence d’un auxiliaire de temps qui fait partie des étiquettes spaCy), en considérant qu’il s’agissait dans tous les cas d’un événement passé, à savoir terminé d’une manière ou d’une autre. Ces derniers ajoutent donc 1 au score clausal. A l’inverse, l’imparfait (étiquette du modèle fr) et le progressif (identifiée par l’expression régulière « en train de ») sélectionneraient une interprétation atéléique, et réduisent ainsi le score de 1.

2. Alors que les phrases atéléiques peuvent être intransitives ou avoir un objet direct optionnel qui modifie le prédicat, les phrases téléique requièrent un participant à l’événement projeté en position d’objet (qui peut être réalisé comme un objet direct quantifié ou comme le sujet dans une construction non-accusative). La présence d’un objet grammatical (étiquette spaCy) ajoute donc 1 au score clausal.

3. En particulier, un objet quantifié (article défini, numéral) ou un objet dénombrable (*count nouns* en anglais) nous obligent souvent à dériver une interprétation téléique et, inversement, un objet indénombrable une interprétation atéléique (*mass nouns*). La présence d’un article défini ou numéral devant l’objet ajoute donc 1 au score clausal.

4. Selon Krifka⁷, la pluralité de l’objet direct signifie que l’événement doit être associé

7. Susan Rothstein, « What are Incremental Themes ? » ()

à une pluralité d'événements « en cours de réalisation », dont le nombre n'est pas précisé. De même, le point final de l'accomplissement est déterminé par le moment où le point final de tous ces événements est atteint, mais rien n'est dit sur leur nombre ni sur le fait qu'ils aient lieu simultanément ou séquentiellement. La localisation du point final dans le cas des « pluriels nus » (*bare plurals* en anglais) n'est donc pas identifiable, et ils sélectionnent pour l'atélicité. La présence du pluriel indéfini *des* devant l'objet, équivalent français le plus proche à nos yeux, soustrait donc 1 au score clausal.

Ajout personnel, mais qui nous a semblé intuitivement participer à l'individuation de l'objet grammatical en question : la présence d'un complément d'objet de tout type (adjectif, locution possessive [objet] de X etc.) associé à l'objet ajoute 1 au score clausal (concrètement, une étiquette spaCy « amod », « nod », « nmod » ou « appos » dans les dépendances syntaxiques de l'objet considéré).

5. Une phrase prépositionnelle (PP) directionnelle peut également affecter la télicité : alors qu'un prédicat d'activité ne permet normalement pas l'interprétation télique, l'ajout d'une PP directionnelle peut contribuer à l'interprétation télique. En particulier (mais valant uniquement pour l'anglais ici, bien que nous réfléchissions plus tard à la possibilité d'un équivalent français), Walková⁸ distingue scalarité et télicité. Elle considère que la scalarité est une condition nécessaire mais non suffisante pour que l'effet de marquage de la télicité se produise. Ce qui compte vraiment, c'est en réalité que l'échelle en question soit bornée (verbes ou syntagmes téliques) ou non (atéliques). Atteindre la limite de l'échelle (notons qu'elle inclut dans son analyse les échelles à 2 points sous-jacentes aux événements ponctuels, qui eux sont toujours limités) équivaut à atteindre le point final de l'événement (c'est pourquoi la télicité est définie comme un point final temporel), bien que l'événement puisse en principe se poursuivre avec ce dernier type de limite.

Ainsi, les « particules scalaires » (celles qui peuvent affecter la structure argumentale de la racine du verbe et renforcer la télicité - *down, off, out, over, through* et *up* en anglais) marquent la télicité lorsqu'elles se réfèrent à une échelle lexicalisée dans un objet direct qui ne peut pas devenir un sujet non-accusatif.

Plus précisément, les particules scalaires ne peuvent apparaître qu'avec un objet direct quantifié. Lorsque la quantité est univoque (déterminant cardinal), elles renforceraient une sorte de télicité redondante ; et pour les syntagmes définis et possessifs (quantité ambiguë), elles imposeraient la télicité. Inversement, s'il n'y a pas d'autre constituant syntaxique « +scalaire », les « particules non-scalaires » (*about, along, around, on, away*) n'affecteraient pas la structure argumentale du verbe, de sorte qu'elles ne pourraient renforcer la télicité. De manière cohérente, elles semblent même ajouter, dans ce cas, un sens qui pourrait être interprété comme dénotant un niveau inférieur d'agentivité (l'action étant vécue comme non intentionnelle et/ou intéressante en elle-même, dans son déroulement, plutôt que pour son résultat - pour reprendre la distinction souvent évoquée en linguistique entre *path* et *manner*).

Ainsi, et en considérant pour notre propos la possibilité d'une équivalence en français, les « verbes à particule » (à savoir ici les verbes suivis d'une préposition introduisant une construction

8. Milada Walková, « Particle verbs in English : Telicity or scalarity? », *Linguistics*, 55-3 (2017), p. 589-616, DOI : [doi:10.1515/ling-2017-0005](https://doi.org/10.1515/ling-2017-0005)

verbale subordonnée) ajoutent 1 au score clausal.

Autre initiative personnelle : si la préposition en question véhicule sémantiquement l’idée de direction (« jusqu’à » et « vers » - à l’exclusion de leur utilisation dans des locutions adverbiales temporelles pour éviter la redondance avec ce qui suit), elle ajoute 1 au score clausal. Nous avons au passage dupliqué cette métrique, en comptant séparément le nombre relatif de prépositions directionnelles (*jusqu’à*, *vers*) et de manière (*par*, *avec*, *sans*) dans l’ensemble des répliques concaténées par statut (en les standardisant par le nombre total de mots et le nombre total de prépositions).

6. Depraetere⁹ suggère également (et c’est sa définition de la télélicité que nous avons citée en introduction, car elle était la plus pertinente pour notre question de recherche) que l’intentionnalité est nécessaire à la télélicité. Selon elle, certains types de constituants quantifiés (phrases nominales, numériques, et adverbiaux) ne seront associés à la fonction de fin inhérente que s’ils sont « dans l’horizon » d’une intention mutuellement manifeste. L’intentionnalité est donc une propriété *a priori* exclusive des sujets animés. La présence d’un pronom sujet ou d’un nom propre ajoute ainsi 1 au score clausal.

7. Une corrélation entre la télélicité et la voix passive a également été établie par Frawley¹⁰. Ce dernier soutient que la logique du passif étant de « promouvoir le destinataire de l’action à la position de sujet », il mettrait l’accent sur le résultat du processus encodé par le verbe dans la mesure où le destinataire est le résultat de l’action, produisant ainsi une interprétation télélique. Pour autant, nous considérons plutôt que la voix active est la plus à même de renforcer la télélicité d’une clause (en renforçant la transitivité du verbe, par souci de cohérence avec ce dernier critère). La présence d’une construction passive (donnée par spaCy) soustrait donc 1 au score clausal.

8. Les modèles d’apprentissage automatique récemment développés pour déterminer la télélicité des clauses^{11 12 13 14} confirment les marqueurs mentionnés ci-dessus, à savoir que les transformateurs y prêtent également attention lorsqu’ils concluent sur la télélicité des clauses. Ces modèles nous ont encore permis d’ajouter un critère adverbial, les études en question ayant montré qu’il était également valable pour l’interprétation aspectuelle des participants humains. Nous avons donc construit en amont des listes d’adverbes temporels, locatifs et de manière, à partir du croisement de plusieurs grammaires françaises. Nous avons également dressé une liste,

9. I. Depraetere, « (A)telicity and intentionality »...

10. W. Frawley, *Linguistic semantics*...

11. Eric V. Siegel et Kathleen R. McKeown, « Learning Methods to Combine Linguistic Indicators : Improving Aspectual Classification and Revealing Linguistic Insights », *Computational Linguistics*, 26-4 (déc. 2000), p. 595-628, DOI : [10.1162/089120100750105957](https://doi.org/10.1162/089120100750105957)

12. Alessandra Zarcone et Alessandro Lenci, « Computational Models for Event Type Classification in Context », dans *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08)*, Marrakech, Morocco, 2008, URL : http://www.lrec-conf.org/proceedings/lrec2008/pdf/315_paper.pdf

13. A. Friedrich et Danyana Gateva, « Classification of telicity using cross-linguistic annotation projection », dans *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, 2017, p. 2559-2565, DOI : [10.18653/v1/D17-1271](https://doi.org/10.18653/v1/D17-1271)

14. Yiyun Zhao, Jian Gang Ngui, Lucy Hall Hartley et Steven Bethard, « Do pretrained transformers infer telicity like humans ? », dans *Proceedings of the 25th Conference on Computational Natural Language Learning*, Online, 2021, p. 72-81, DOI : [10.18653/v1/2021.conll-1.6](https://doi.org/10.18653/v1/2021.conll-1.6)

aussi exhaustive que possible, des expressions temporelles pouvant être utilisées pour construire des compléments à valeur adverbiale temporelle (*heures, minutes, etc.*), en localisant leur occurrence à des positions syntaxiques spécifiques qui devraient le plus souvent correspondre à une telle valeur.

TABLE 7.2 – Expressions adverbiales de manière retenues

« à bras-le-corps », « à califourchon », « à la légère », « à la va-comme-je-te-pousse », « à la va-vite », « à l'aveuglette », « à loisir », « à nouveau », « de nouveau », « à tire-d'aile », « à tire-larigot », « à tort », « à tue-tête », « bel et bien », « bon marché », « d'arrache-pied », « de guingois », « par hasard », « pour de bon », « tant bien que mal », « comme », « comment », « ainsi », « aussi », « bien », « debout », « également », « ensemble », « franco », « gratis », « incognito », « mal », « mieux », « pis », « plutôt », « presque », « quasi », « recta », « vite », « volontiers »

TABLE 7.3 – Expressions adverbiales de lieu retenues

« à proximité », « aí », « aux abords de », « aux abords du », « aux abords d », « ailleurs », « alentour », « alentours », « arrière », « attendant », « au diable Vauvert », « autour », « en bas », « vers », « çà », « céans », « chez », « ci », « à côté », « sur le côté », « au côté », « deçà », « delà », « dedans », « dehors », « derrière », « dessous », « dessus », « devant », « à droite », « à gauche », « exa », « à l'extérieur », « face à », « face au », « hái », « en haut », « ici », « là », « à l'intérieur », « léans », « loin », « nulle part », « où », « par monts et par vaux », « partout », « près de », « près du », « près d », « proche de », « proche du », « proche d », « quelque part », « sus », « y »

TABLE 7.4 – Expressions adverbiales temporelles retenues

« à présent », « a présent », « à l'approche d », « a l'approche d », « à l'approche de », « a l'approche de », « à l'approche du », « a l'approche du », « à l'approche des », « a l'approche des », « sur le moment », « sur l'heure », « un temps », « un moment », « un instant », « un jour », « un matin », « une nuit », « un soir », « une soirée », « un beau jour », « un beau matin », « un beau soir », « un hiver », « un été », « un printemps », « un automne », « une année », « d'un moment à l'autre », « d'un instant à l'autre », « un de ces jours », « au début », « au commencement », « à la fin », « a la fin », « en fin », « au milieu », « aux environs de », « aux environs d », « actuellement », « anciennement », « simultanément », « de jour », « de nuit », « en simultané », « fréquemment », « régulièrement », « antan », « après », « alors », « auparavant », « aussitôt », « autrefois », « avant », « bientôt », « dans peu », « de suite », « de ce pas », « de temps en temps », « déjà », « demain », « d'main », « depuis », « derechef », « dernièrement », « désormais », « de tout temps », « de tous temps », « d'abord », « dorénavant », « encore », « enfin », « ensuite », « entre-temps », « entretemps », « quelque temps », « hier », « hui », « illico », « immédiatement », « instantanément », « sous peu », « jadis », « jamais », « jusqu'à », « jusqu'au », « jusqu'aux », « jusque », « longtemps », « lors », « maintenant », « maishui », « meshui », « méshui », « momentanément », « naguère », « naguères », « parfois », « présentement », « prochainement », « puis », « quelquefois », « rarement », « récemment », « sans délai », « sans perdre un instant », « sans attendre », « sans tarder », « sans trop tarder », « sitôt », « soudain », « soudainement », « souvent », « subito », « sur-le-champ », « sur le champ », « sur le coup », « sur le coup de », « sur les coups de », « tantôt », « tard », « tardivement », « tôt », « toujours », « tout à coup », « tout d'un coup », « tout le long de », « tout le long du », « tout le long d », « tu suite », « tandis », « quand », « lorsque », « lorsqu », « une fois que », « une fois qu », « pendant que », « pendant qu », « cependant que », « cependant qu », « chaque fois que », « chaque fois qu », « toutes les fois que », « toutes les fois qu », « tant que », « tant qu », « à mesure », « a mesure », « le temps que », « le temps qu », « dès », « en attendant », « le temps de », « le temps d », « le temps du », « durant », « pendant », « au cours de », « au cours d », « l'espace d'un instant », « dans l'instant », « au bout d », « au bout de », « à partir de », « à partir d », « lundi », « mardi », « mercredi », « jeudi », « vendredi », « samedi », « dimanche », « pièça »

TABLE 7.5 – Lexique univoquement temporel retenu en complément des expressions adverbiales précédentes pour l'analyse regex

« coucher », « lever », « an », « ans », « année », « années », « soir », « soirs », « soirée », « soirées », « matin », « matins », « matinée », « matinées », « jour », « jours », « nuit », « nuits », « journée », « journées », « midi », « midis », « minuit », « heure », « heures », « minute », « minutes », « seconde », « secondes », « mois », « janvier », « février », « mars », « avril », « mai », « juin », « juillet », « août », « septembre », « octobre », « novembre », « décembre », « siècle », « siècles », « saison », « saisons », « hiver », « printemps », « été », « automne », « étés », « hivers », « printemps », « automnes », « semaine », « semaines », « temps », « moment », « moments », « instant », « instants », « fois »

Finalement, le score de la clause (c.a.d. de chaque verbe étudié) a été normalisé, suivi du score de la phrase (la normalisation prenant comme paramètre le nombre de clauses présentes dans la phrase en question, qui détermine les scores maximum et minimum utilisés pour la normalisation) et du score total du texte, calculé en faisant la moyenne des scores de toutes les phrases analysées.

7.1.1 Les modes de l'irréel et la pensée hypothétique

Les modes de l'irréel font référence à l'ensemble des modes grammaticaux indiquant qu'une certaine situation ou action n'est pas connue pour s'être produite au moment où le locuteur parle. Sur la base de la classification de Wikipedia (valable pour l'anglais et le français)¹⁵, nous avons ainsi mesuré la proportion relative des formes indicatives par rapport aux formes non affirmatives (subjonctif, conditionnel, interrogatif et impératif).

Comme mentionné précédemment, nous avons encore mesuré l'utilisation du futur, en considérant qu'il dénotait également une modalité épistémique irréaliste.

Enfin, nous avons quantifié la fréquence d'alternance entre les temps pour les personnages de haut et de bas statut, en supposant qu'une plus grande capacité de distanciation psychologique devrait se traduire, en moyenne, par des allers-retours plus fréquents entre des considérations rétrospectives, présentes et anticipatrices.

7.1.2 Les marqueurs corrélés d'abstraction

Dans son article de 2021, Lievers¹⁶ énumère une panoplie de marqueurs linguistiques qui peuvent être utilisés comme proxys pour estimer le niveau d'abstraction d'un texte, sans avoir à

15. [https://fr.wikipedia.org/wiki/Mode_\(grammaire\)](https://fr.wikipedia.org/wiki/Mode_(grammaire))

16. F. Strik Lievers, M. Bolognesi et B. Winter, « The linguistic dimensions of concrete and abstract concepts... »

mobiliser un modèle d'apprentissage particulièrement complexe pour cette tâche. Parmi tous ceux qu'il propose, le plus facile à mettre en œuvre était la fréquence relative d'utilisation des parties du discours principalement associées à un sens plus abstrait (à savoir les adverbes ; les *mots-outils*, c.a.d. les déterminants et adjectifs non qualificatifs, conjonctions, prépositions excluant les locatifs, pronoms, auxiliaires, modaux, outils interrogatifs, qualificatifs ; et les adjectifs), par rapport à ceux associés à un sens plus concret (noms et verbes).

Comme on l'a expliqué, cet indicateur implique une incertitude quant à la direction de la corrélation attendue pour le nombre de verbes utilisés par chaque groupe de personnages, les verbes étant intuitivement liés à l'agentivité dans le sens d'une plus grande disposition à agir, réalisée ou non, mais inversement corrélés si l'on considère le niveau d'abstraction supplémentaire que les agents « haut niveau » sont censés mobiliser. C'est ce qui motive l'utilisation d'un indice agentif global, qui devrait résister à de telles ambiguïtés individuelles.

Chapitre 8

Le langage de la personnalité : indices d'extraversion et d'agréabilité

Une précision s'impose concernant l'analyse des marqueurs linguistiques associés respectivement à l'extraversion et à l'agréabilité, censés garantir la spécificité de notre indice agentif.

En nous appuyant notamment sur les travaux de Pennebaker¹, qui a principalement utilisé le LIWC² pour ses études, bien qu'il ait également été en mesure d'identifier des associations linguistiques plus « structurelles » que nous avons déjà pris en compte, nous avons pu dresser la liste des indicateurs linguistiques ci-dessous³ :

1. James W Pennebaker et Laura A King, « Linguistic Styles : Language Use as an Individual Difference » (); J. Pennebaker et C. Chung, « Expressive Writing, Emotional Upheavals, and Health » (, janv. 2007)

2. Le LIWC consiste notamment en une collection de dictionnaires répertoriant les mots relatifs à des catégories sémantiques spécifiques, et créés par les utilisateurs eux-mêmes : <https://www.liwc.app/>

3. Pour une liste complète des références parcourues avant de parvenir à cette sélection : James W Pennebaker et L. A. King, « Linguistic Styles : Language Use as an Individual Difference »...; Francis Heylighen et Jean-Marc Dewaele, « Variation in the Contextuality of Language : An Empirical Measure », *Foundations of Science*, 7-3 (1^{er} sept. 2002), p. 293-340, DOI : [10.1023/A:1019661126744](https://doi.org/10.1023/A:1019661126744); J. Pennebaker et C. Chung, « Expressive Writing, Emotional Upheavals, and Health »...; Jon Oberlander et Scott Nowson, « Whose thumb is it anyway? : classifying author personality from weblog text », dans *Proceedings of the COLING/ACL on Main conference poster sessions* -, Sydney, Australia, 2006, p. 627-634, DOI : [10.3115/1273073.1273154](https://doi.org/10.3115/1273073.1273154); F. Mairesse, M. A. Walker, M. R. Mehl et R. K. Moore, « Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text », *Journal of Artificial Intelligence Research*, 30 (28 nov. 2007), p. 457-500, DOI : [10.1613/jair.2349](https://doi.org/10.1613/jair.2349); Alastair Gill, S. Nowson et J. Oberlander, « What Are They Blogging About? Personality, Topic and Motivation in Blogs », *Proceedings of the International AAAI Conference on Web and Social Media*, 3-1 (19 mars 2009), p. 18-25, DOI : [10.1609/icwsm.v3i1.13949](https://doi.org/10.1609/icwsm.v3i1.13949); Tal Yarkoni, « Personality in 100,000 Words : A large-scale analysis of personality and word use among bloggers », *Journal of Research in Personality*, 44-3 (juin 2010), p. 363-373, DOI : [10.1016/j.jrp.2010.04.001](https://doi.org/10.1016/j.jrp.2010.04.001); Gokul Chittaranjan, Jan Blom et Daniel Gatica-Perez, « Who's Who with Big-Five : Analyzing and Classifying Personality Traits with Smartphones », dans *2011 15th Annual International Symposium on Wearable Computers*, San Francisco, CA, USA, 2011, p. 29-36, DOI : [10.1109/ISWC.2011.29](https://doi.org/10.1109/ISWC.2011.29); Margaret L. Kern, Johannes C. Eichstaedt, H. A. Schwartz, Gregory Park, Lyle H. Ungar, David J. Stillwell, Michal Kosinski, Lukasz Dziurzynski et Martin E. P. Seligman, « From “Sooo excited!!!” to “So proud” : Using language to study development. » *Developmental Psychology*, 50-1 (2014), p. 178-188, DOI : [10.1037/a0035048](https://doi.org/10.1037/a0035048); H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, M. Kosinski, David Stillwell, M. E. P. Seligman, *et al.*, « Personality, Gender, and Age

TABLE 8.1 – Récapitulatif des marqueurs linguistiques distinctifs identifiés pour les deux catégories de traits d’intérêt du modèle des *Big Five*

<u>Extraversion</u> (et Conscienciosité)	<u>Agréabilité</u> (et Ouverture)
Points d’interrogation (-)	Points d’interrogation (+)
<u>Articles (-)</u>	Articles (+)
<u>Négations (-)</u>	Passé (-)
<u>Style implicite (verbes, adverbes, pronoms) (+)</u>	
<u>Style formel (par opposition à contextuel) (-)</u>	
<u>Richesse lexicale (Uber index) (-)</u>	
<u>WPS (<i>words per sentence</i>) (+)</u>	

Les mots soulignés sont ceux dont l’utilisation a été directement associée aux deux traits les plus déterminants que sont l’extraversion et l’agréabilité dans la littérature (à savoir les traits les plus directement associés à l’agentivité et à la communion, respectivement). En gras, le marqueur a également été associé de manière cohérente au trait de caractère plus secondaire mentionné (respectivement la conscienciosité et l’ouverture). Le signe + ou - indique la corrélation positive ou négative attendue avec le trait de personnalité en question.

Concernant la ponctuation, nous avons ajouté, en plus des points d’interrogation, des métriques retraçant l’usage de tous les autres signes de ponctuation, certaines des études mentionnées laissant penser qu’il existerait là aussi des associations préférentielles selon la personnalité du locuteur, bien que ces signes n’étaient *a priori* pas distinctifs s’agissant de la corrélation avec l’extraversion/conscienciosité et l’agréabilité/ouverture respectivement (à savoir qu’ils étaient corrélés à ces deux catégories de traits simultanément, et ce dans la même direction – par exemple positivement corrélés à la fois à l’extraversion et à l’agréabilité ; ou bien encore qu’ils étaient inversement corrélés à chacun des deux traits mentionnés au sein desdites catégories – par exemple positivement corrélés à l’agréabilité mais négativement corrélés à l’ouverture, ce qui limitait à

in the Language of Social Media : The Open-Vocabulary Approach », dir. Tobias Preis, *PLoS ONE*, 8-9 (25 sept. 2013), e73791, DOI : [10.1371/journal.pone.0073791](https://doi.org/10.1371/journal.pone.0073791) ; Fabio Celli et Massimo Poesio, « PR2 : A Language Independent Unsupervised Tool for Personality Recognition from Text » (, févr. 2014) Yash Mehta, Samin Fatehi, Amirmohammad Kazameini, Clemens Stachl, Erik Cambria et Sauleh Eetemadi, « Bottom-Up and Top-Down : Predicting Personality with Psycholinguistic and Language Model Features », dans *2020 IEEE International Conference on Data Mining (ICDM)*, Sorrento, Italy, 2020, p. 1184-1189, DOI : [10.1109/ICDM50108.2020.00146](https://doi.org/10.1109/ICDM50108.2020.00146) ; Jennifer E. Caplan, Kiki Adams et R. L. Boyd, « Personality and Language », dans *The Wiley Encyclopedia of Personality and Individual Differences*, dir. Bernardo J. Carducci, Christopher S. Nave et Christopher S. Nave, 1^{re} éd., 2020, p. 311-316, DOI : [10.1002/9781118970843.ch52](https://doi.org/10.1002/9781118970843.ch52) ; Sanja Stajner et Seren Yenikent, « A Survey of Automatic Personality Detection from Texts », dans *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain (Online), 2020, p. 6284-6295, DOI : [10.18653/v1/2020.coling-main.553](https://doi.org/10.18653/v1/2020.coling-main.553)

première vue leur robustesse pour servir de marqueur distinctif).

Pour la même raison, nous avons inclus dans nos métriques le comptage du nombre total de mots, de phrases, ainsi que des mots de plus de 6 lettres, tous étudiés en association avec la personnalité du locuteur mais se révélant, dans les études en question, non distinctifs.

La détermination de la formalité du style correspond au calcul de l'indice suivant : $F = (\text{fréquence du nom} + \text{fréquence de l'adjectif} + \text{fréquence de la préposition} + \text{fréquence de l'article} - \text{fréquence du pronom} - \text{fréquence du verbe} - \text{fréquence de l'adverbe} - \text{fréquence de l'interjection} + 100)/2$. Les personnes plus extraverties ont été identifiées comme utilisant un langage moins formel (et corollairement plus contextuel) que les personnes plus introverties.

Comme mentionné dans le tableau, un style de pensée plus implicite, caractérisant les individus plus extravertis, fait référence à l'utilisation relativement plus importante de verbes, d'adverbes et de pronoms, pour lesquels nous avons construit une métrique séparée, en plus du suivi individuel de ces différentes parties du discours.

Enfin, la richesse lexicale correspond au calcul de l'indice de « Uber » suivant :

$$U = \frac{\log(tokens)^2}{\log(tokens) - \log(types)} \quad (8.1)$$

Les *tokens* correspondent au nombre de mots (tels qu'identifiés par spaCy), et les *types* au nombre de mots uniques ou distincts utilisés (à l'exclusion des *mots-outils* ou de la ponctuation), le rapport des deux fournissant une approximation de la diversité lexicale des répliques d'un personnage. Nous avons également calculé, séparément, le ratio du nombre de *tokens* différents utilisés par rapport au nombre total de mots, bien que ce marqueur ait été corrélé négativement à la fois à l'extraversion et à l'agréabilité.

De façon très exploratoire, nous avons également intégré dans notre code deux métriques s'attachant à approximer deux styles de pensée censés être relativement exclusifs : un style de pensée « catégorique », positivement corrélé avec le succès académique dans une étude menée par Pennebaker⁴, et dont la prévalence serait, à titre purement spéculatif, attendue chez les individus relativement plus agentifs (l'agentivité étant liée à l'assertivité, qui devrait logiquement se traduire par des réussites personnelles relativement plus nombreuses, y compris dans le domaine scolaire); et un style de pensée plus « dynamique », quant à lui corrélé négativement avec le succès académique et qui pourrait ainsi davantage caractériser les individus relativement moins agentifs. D'un point de vue linguistique, un style catégorique se traduit par l'utilisation d'articles, de prépositions, et dans une certaine mesure de conjonctions et de négations (bien que les résultats soient contradictoires pour ces deux derniers marqueurs) relativement plus fréquente. A l'inverse, un style de pensée dynamique ou « flexible » correspond à l'usage préférentiel de verbes auxiliaires, de pronoms et d'adverbes⁵

4. James W. Pennebaker, Cindy K. Chung, Joey Frazee, Gary M. Laverne et David I. Beaver, « When Small Words Foretell Academic Success : The Case of College Admissions Essays », *PLOS ONE*, 9–12 (déc. 2015), Publisher : Public Library of Science, p. 1-10, DOI : [10.1371/journal.pone.0115844](https://doi.org/10.1371/journal.pone.0115844)

Quatrième partie

Résultats et discussion

Chapitre 9

Des résultats prometteurs

Pour construire notre indice global d’agentivité, calculé au niveau de chaque pièce, nous avons tout d’abord normalisé les mesures de chaque métrique linguistique afin qu’elles soient comprises entre 0 et 1. Nous avons ensuite additionné, pour une pièce considérée, le score de toutes les métriques calculées par statut (haut/bas) ou respectivement par genre (masculin/féminin), des personnages ayant prononcé les répliques concaténées, et sur la base desquelles chaque métrique avait été calculée dans la pièce en question. La somme obtenue correspond ainsi, pour chaque pièce, au score d’une « méga-variable » calculée par statut (resp. par genre), et constituant la première ébauche de notre indice agentif. Ce score est compris entre 0 et 1.

Au total (en considérant toutes les pièces analysées dans notre corpus), nous disposons de 888 observations afin de tester l’association entre « mégascore » agentif et statut des personnages, et de 891 observations s’agissant de l’association entre le mégascore agentif et le genre des personnages. A partir de tous ces points, nous tentons donc d’estimer la corrélation entre l’indice d’agentivité ainsi construit et la variable catégorielle *statut* (resp. *genre*), à l’aide de différents modèles de régression linéaire détaillés ci-après :

- Un modèle de régression simple calculant l’impact du Statut (resp. Genre) sur le mégascore agentif.

- Un modèle de régression logit calculant à l’inverse l’impact du mégascore agentif sur le Statut (resp. Genre) des personnages locuteurs.

A noter que pour l’analyse s’intéressant au statut, nous avons en réalité construit non pas deux mais quatre modèles de régression au total, en dupliquant chaque modèle selon que l’on prenait en compte, ou non, la direction de corrélation attendue (positive ou négative) pour chacune des métriques linguistiques avec le score agentif. On notera que les verbes sont comptés positivement ici, y compris les verbes modaux et l’usage de la modalité en général (en ce qu’elle reflète supposément une agentivité « morale »), ce malgré l’ambiguïté mentionnée dans les sections correspondantes pour ces marqueurs. L’on considère en effet dans cette première ébauche que le verbe reste, intuitivement, un proto-marqueur agentif qu’il serait difficile d’ignorer, en sachant que si la direction s’avère contraire à celle prédite, cela ne devrait pas être suffisamment important pour affecter la direction du score global étant donnée la multitude d’autres métriques incluses dans son calcul. Et c’est en ce sens que nous conservons malgré tout un calcul plus «

grossier » du score agentif, qui somme simplement toutes les métriques dont on pense qu'elles ont avoir avec la notion d'agentivité, quelle que soit la direction de la corrélation postulée. Nous avons également supprimé du calcul du score agentif, dans la seconde version des modèles, les métriques pour lesquelles la direction de corrélation demeurerait trop ambiguë, à savoir : le nombre de parenthèses, points, virgules, points-virgules, deux points, guillemets, apostrophes, tirets et l'emploi du gérondif¹.

Ce problème ne s'est pas posé concernant le genre, pour lequel nous n'avions pas de quoi supposer, *a priori*, une direction de corrélation précise pour chaque métrique de l'indice, par rapport au genre. Dans ce cas en effet, nous supposons simplement que le score agentif, même s'il comprend certaines métriques en réalité inversement corrélées à ce dernier et ayant donc un score plus faible quand ce degré d'agentivité augmente, devrait globalement corrélér positivement avec le genre masculin, contrairement à celui féminin.

Concernant le statut des personnages tout d'abord, les résultats sont donc présentés ici respectivement sans et avec prise en compte des directions de corrélation propres à chaque métrique (à savoir que dans la seconde version de chacun des modèles, les métriques suivantes ont été comptées négativement au moment de calculer le mégascore agentif : ponctuation, ellipses, ratio du nombre de *tokens* différents utilisés par rapport au nombre total de mots ou lemmes - deux versions d'une même métrique, indice de richesse lexicale, noms, articles, adverbes de manière, adverbes de doute, négations, présent, imparfait, indicatif, verbes statifs, copules, prépositions de manière, formalité, style « dynamique »).

En moyenne, le mégascore agentif est de 23.5 pour les personnages de haut statut contre 22.5 pour ceux de bas statut (respectivement 7.35 contre 6.24 en prenant en compte les directions).

TABLE 9.1 – Score agentif par statut - modèle 1

Statut	Moyenne	Ecart-type
Haut	23.5	0.826
Bas	22.5	1.22

1. Ces métriques n'ont d'ailleurs pas été mentionnées du tout, ou seulement très brièvement, dans le corps de ce mémoire. En guise de précision donc, les signes de ponctuation supprimés dans la seconde version des modèles étaient régulièrement associés à des traits de personnalité d'intérêt du *Big Five* dans les études correspondantes parcourues, mais ce de façon contradictoire et *a fortiori* non distinctive. Quant au gérondif, l'on avait supposé qu'il pourrait venir s'ajouter aux proxy téléiques individuels - hors analyse compositionnelle, en ce qu'il dénoterait en l'occurrence un aspect atélique par la focalisation sur le déroulement de l'action qu'il sous-tend. Cela dit, cette interprétation demeurerait très personnelle et incertaine, n'étant soutenue par aucune étude spécifique. Il est donc également supprimé du calcul du score dans la seconde version des modèles. On précisera enfin, au passage, que la suppression des métriques comptant le nombre de phrases et de mots prononcés dans le calcul de l'indice agentif ne change pas la significativité des résultats qui suivent, de sorte que cet indice ne se contente pas de révéler le fait qu'un maître parle plus souvent, et/ou plus longtemps, que son serviteur dans la pièce, toutes les autres métriques incluses étant normalisées en fonction du nombre total de mots dans chacun des deux groupes de répliques précisément afin de prévenir ce biais.

TABLE 9.2 – Score agentif par statut - modèle 2

Statut	Moyenne	Ecart-type
Haut	7.35	0.940
Bas	6.24	1.05

Aussi, un t.test nous révèle que la différence de moyenne du mégascore agentif entre ces deux statuts est chaque fois significative ($p < 0.001$).

En appliquant le modèle de régression linéaire simple suivant :

$$Score\ agentif \sim Statut \quad (9.1)$$

On trouve de façon cohérente qu'il existe une association linéaire significative entre le mégascore agentif et la variable catégorique *statut*, ce avec et sans prise en compte des directions de corrélation individuelles ($p < 0.001$).

D'après le premier modèle, le fait de passer de la catégorie « bas » à « haut » statut, toutes choses égales par ailleurs, augmente la valeur moyenne du score agentif de 1.01, contre 1.11 dans le second modèle.

Le R^2 pour chacun des deux modèles est de 0.19 et 0.24 respectivement, avec un sigma (écart-type du terme d'erreur) de 1.04 et 1. Le second modèle explique donc une plus grande proportion de la variance du score agentif, et génère des prédictions plus proches des points de données réels (sigma plus faible). Il peut ainsi être considéré comme ayant une meilleure adéquation aussi bien en termes de R^2 que de minimisation des erreurs de prédiction.

Puisqu'il n'était pas aisé de savoir si l'on s'attendait à ce que le statut permette de prédire le langage utilisé dans la pièce, ou bien plutôt à ce que l'utilisation d'un certain langage prédise le statut des personnages, nous avons également analysé le modèle de régression logistique suivant :

$$Statut \sim Score\ agentif \quad (9.2)$$

Là encore, l'association est chaque fois très significative ($p < 0.001$), la seconde version du modèle (avec prise en compte des directions individuelles) présentant de façon cohérente un meilleur « fit » que la première (déviante résiduelle et AIC plus faible), ainsi qu'un coefficient de score agentif un peu plus élevé.

Exprimé en ratio de vraisemblance (et non plus en logarithme de ce dernier), le premier modèle indique que pour chaque augmentation d'une unité du score agentif des répliques considérées, les chances que leurs locuteurs soient de haut statut sont 2.8 fois plus élevées, toutes choses égales par ailleurs (l'intervalle de confiance à 95% associé étant : [2.29, 3.53]). Le second modèle (avec prise en compte des directions individuelles) présente quant à lui un odds-ratio de 3.03 ([2.49, 3.75]).

S'agissant des modèles logistiques, nous avons encore testé leurs performances prédictives au moyen d'une validation croisée dix fois. La précision de chacun des modèles (sans et avec prise en compte des direction de corrélation respectivement) est de 75% et 79% respectivement, leur sensibilité de 78% et 77%, et ainsi leur F1-score de 77% et 78%. Si le second modèle est

légèrement plus précis que le premier (relativement moins de faux positifs), leurs performances prédictives globales sont donc équivalentes.

TABLE 9.3 – Résultats de l'évaluation des modèles logistiques relatifs au statut en validation croisée

	precision	recall	f1-score
Modèle 1	0.754	0.782	0.768
Modèle 2	0.786	0.774	0.780

Concernant le genre des personnages cette fois, pour lequel nous avons systématiquement conservé toutes les métriques sans duplication des modèles, on trouve qu'en moyenne, le mégascore agentif est de 29 pour les personnages féminins contre 29.1 pour les locuteurs masculins. Un t.test nous confirme que cette différence de moyenne du mégascore agentif entre les deux genres n'est pas du tout significative ($p = 0.9921$).

TABLE 9.4 – Score agentif par genre

Genre	Moyenne	Ecart-type
Masculin	29.1	0.830
Féminin	29.0	1.06

En appliquant le modèle de régression linéaire simple suivant :

$$Score\ agentif \sim Genre \quad (9.3)$$

On trouve qu'il existe une association linéaire faiblement significative ($p = 0.0159$) entre ces deux variables que sont le mégascore agentif et genre du personnage.

Le fait de passer de la catégorie « féminin » à « masculin » augmente plus précisément la valeur moyenne du score agentif de 0.15, toutes choses égales par ailleurs.

Le R^2 du modèle est de 0.007, et son sigma de 0.95. Cela suggère que le genre, en tant que seul prédicteur, n'explique pas une grande partie de la variabilité du mégascore agentif. Sa significativité pratique serait donc limitée. Les prédictions du modèle sont du reste relativement proches des points de données réels (sigma faible).

Pour les mêmes raisons que précédemment, nous avons également analysé le modèle de régression logistique suivant :

$$Genre \sim Score\ agentif \quad (9.4)$$

Là encore, l'association est faiblement significative ($p = 0.0374$).

Aussi, le modèle indique que pour chaque augmentation d'une unité du score agentif des répliques considérées, les chances que leurs locuteurs soient des personnages masculins sont 1.2 fois plus élevées, toutes choses égales par ailleurs (l'intervalle de confiance à 95% associé étant : [1.01, 1.45]).

Nous avons à nouveau testé ses performances prédictives au moyen d’une validation croisée dix fois. La précision du modèle est de 54%, sa sensibilité de 52%, et ainsi son F1-score de 53%. Ce modèle est donc, logiquement, très peu performant lorsqu’il s’agit de prédire le genre du personnage locuteur, sur la base du mégascore agentif seul.

TABLE 9.5 – Résultats de l’évaluation du modèle logistique relatif au genre en validation croisée

precision	recall	f1-score
0.539	0.512	0.529

Ces résultats constituent donc une première confirmation de nos intuitions théoriques, à savoir que le niveau d’agentivité d’un personnage, tel que nous l’avons défini théoriquement et sur la base d’une validation interne à la pièce (c’est-à-dire a priori informés par le statut du personnage), semble être identifiable de manière relativement systématique par un ensemble de marqueurs linguistiques au niveau agrégé.

Il est également rassurant que la seconde version de notre modèle linéaire simple s’attachant à expliquer le mégascore agentif par le statut du personnage se révèle globalement plus performant que la première. À savoir que la prise en compte des directions de corrélation propre à chaque marqueur linguistique rend de fait le niveau d’agentivité des locuteurs (approximé par leur statut) plus « capable » d’expliquer leur mégascore agentif, ainsi calculé.

Néanmoins, la première tentative de validation externe de notre indice agentif global, au moyen du genre des personnages, est pour sa part peu concluante. De fait, l’association entre genre et mégascore agentif n’est que très faiblement significative, et les performances prédictives d’un modèle de régression logistique tentant d’expliquer ledit mégascore par le genre des locuteurs sont à peine supérieures à celles qu’obtiendrait un modèle de classification aléatoire.

Cela étant, la notion de genre ne se limite pas à l’agentivité, et est susceptible de charrier un certain nombre d’autres dimensions qui pourraient précisément interférer avec le degré d’agentivité d’un individu au moment de façonner son langage. Il n’est donc pas étonnant que la corrélation soit moins évidente que pour le statut. Ce d’autant plus que, n’ayant pas eu le temps de nous plonger dans la littérature s’intéressant aux associations préférentielles entre genre et langage, nous n’avons pas entrepris d’affiner le calcul de notre indice par la prise en compte de la direction de corrélation précise attendue pour chaque métrique prise isolément (et il semblait peu justifié, à ce stade, de simplement répliquer les directions de corrélation prédites pour le statut et non pour le genre spécifiquement). Aussi, on notera que la direction globale prédite est quant à elle respectée, à savoir que l’indice agentif est bien préférentiellement associé au genre masculin, bien que cette préférence soit légère, par rapport à celui féminin.

Chapitre 10

Discussion

10.1 Limites de la présente recherche

Notre étude, très exploratoire, connaît bien entendu certaines limites.

Tout d’abord, un certain nombre de travaux sur lesquels nous nous appuyons afin de définir, *a priori*, nos marqueurs linguistiques agentifs ne triangulent pas les données en faisant notamment appel, outre l’analyse de corpus par « lecture distante », à des expériences psycholinguistiques. Certains marqueurs sont donc étiquetés agentifs sur la base de la seule appréciation subjective de l’auteur, et certains résultats sont d’ailleurs en partie mis en cause par des études contradictoires (pour la modalité notamment).

Aussi, toutes ces études valent pour l’anglais et non le français. Pour autant, notre recherche s’intéressant en priorité à des marqueurs « structurels » (c.a.d. aux niveaux grammatical et syntaxique) relativement simples, la transposition des conclusions au français ne nous semblait pas déraisonnable (la présence d’un objet grammatical par exemple, semble devoir produire un effet relativement similaire quelle que soit la langue considérée, d’autant plus entre des locuteurs appartenant à des sociétés culturellement proches comme le sont les Etats-Unis et la France). Nous sommes toutefois conscients qu’il existe des différences notables entre ces deux langues, s’agissant notamment du cadrage dit « satellitaire » pour l’anglais et « verbal » pour le français. En effet, l’on distingue depuis Talmy¹ deux grands types de langues : celles dites « encadrées par le verbe », où l’emplacement et le mouvement sont encodés par le verbe principal d’une clause, une adjonction facultative se chargeant alors d’encoder la manière dont se déroule l’évènement ; et celles à cadre « satellitaire », où la localisation et le mouvement sont encodés par un élément associé au verbe, laissant le verbe libre d’encoder la manière. Aussi, cette typologie binaire traditionnelle en linguistique, entre « chemin » et « manière », est directement liée à la notion de télélicité, une clause télélique s’intéressant par définition davantage à la direction de l’évènement (sa finalité) qu’à la manière dont il s’est déroulé. Notre marqueur télélique, défini de manière compositionnelle et sur la base d’études anglophones, s’avère donc très incertain, ce d’autant plus qu’il se fonde également sur des choix personnels. En particulier, l’analyse des prépositions

1. Leonard Talmy, « Path to Realization : A Typology of Event Conflation », dans 1991, URL : <https://api.semanticscholar.org/CorpusID:60877805>

pour décider de la télélicité d'une clause (selon qu'elles soient « directionnelles » - « jusqu' », « jusqu' », « vers » – ou indiquant une manière de faire – « sans », « avec », « par ») pourrait ne pas être pertinente s'agissant du français (bien que certaines études aient depuis contesté la validité de cette typologie binaire entre deux cadrages linguistiques du « chemin » et de la « manière »²). C'est pour cette raison que nous avons intégré à l'analyse le calcul isolé de proxy téléliques plus « grossiers », à savoir le temps du discours (parfait ou imparfait) et l'individuation de l'objet grammatical (présence d'un article défini, d'un déterminant possessif, démonstratif, numéral etc), calculés séparément de cette métrique compositionnelle.

La confection des listes d'adverbes pose également la question de leur solidité. Si les adverbes français exprimant une intention, un doute, ou une affirmation semblent *a priori* exister en nombre limité, et que ces derniers ont été identifiés en prenant soin de recouper plusieurs grammaires françaises officielles, ces métriques doivent être interprétées avec précaution. Elles reposent en effet sur des choix personnels, et se placent à un niveau d'analyse sémantique qui reste soumis au risque de variation historique quant au sens et à la fréquence d'utilisation des mots considérés, variation pouvant être soutenue par des phénomènes de transmission culturelle non signifiants pour notre étude. Le noème d' « intentionnalité » pourrait en outre être exprimé par différents éléments lexicaux selon les époques, selon un processus de glissement similaire à celui qui semble s'opérer entre verbes modaux et semi-modaux ce dernier siècle. L'augmentation de l'utilisation d'adverbes intentionnels ne saurait donc renseigner que de façon très partielle l'évolution de l'expression de l'intentionnalité, en général, à travers le temps.

Enfin, en l'absence de validation externe, nos modèles d'analyse ne disent pour l'instant rien de la spécificité de notre indice agentif global, qui pourrait tout autant corrélérer avec le niveau de communion des personnages, par exemple.

10.2 Analyses successives

Etant donnée l'échec de la validation externe de notre indice par le genre des personnages, nous ne pouvons en effet pas encore être certains de son caractère distinctif au moment d'estimer le niveau d'agentivité d'un locuteur. Il faudrait encore étudier, en parallèle, les marqueurs d'extraversion et d'agréabilité, qui correspondent pour rappel aux deux traits de personnalité du modèle des *Big 5* les plus directement corrélés à l'agentivité et à la communion respectivement. Pour que notre indice agentif soit spécifique, il faudrait alors que le mégascore d'extraversion ainsi obtenu soit significativement et positivement corrélé au statut des personnages, et que celui d'agréabilité soit à l'inverse non significativement corrélé au statut (ou même négativement corrélé à ce dernier, bien que l'agentivité et la communion soient *a priori* orthogonales).

Nous avons pour ce faire déjà identifié les éléments linguistiques préférentiellement associés à l'extraversion (et dans une moindre mesure à la conscienciosité), et à l'agréabilité (et dans une moindre mesure à l'ouverture). Cette analyse pourrait être complétée par un usage plus global du LIWC, qui sert de support à la plupart des études s'intéressant au lien entre traits de personnalité

2. Thomas Egan, « Manner and Path : evidence from a multilingual corpus », *CogniTextes*, 12 (sept. 2015), DOI : [10.4000/cognitextes.788](https://doi.org/10.4000/cognitextes.788)

et langage, en incluant ses composants sémantiques (qui n'ont pas été listés dans le tableau de la section III. 4. *Le langage de la personnalité : indices d'extraversion et d'agréabilité* étant donnée la focalisation « structurelle » de ce mémoire).

A titre subsidiaire, il serait également intéressant d'estimer l'interaction entre le statut et le genre des personnages au moment d'expliquer leur score agentif. Cette analyse n'a pu être menée ici car elle supposait de faire tourner notre code à nouveau (le temps de calcul étant conséquent), en calculant les métriques par personnage, et non plus sur leurs répliques concaténées. En effet, bien que le genre ne soit pas un prédicteur significatif de l'agentivité en tant que tel, il se pourrait que le fait d'être un homme exacerbe l'effet du statut sur le mégascore agentif.

Si tant est que notre indice global agentif soit distinctif, une analyse ultérieure consisterait à analyser plus finement l'importance relative de chaque métrique qui compose cet indice global, prise individuellement et/ou en les regroupant en sous-catégories selon certains critères tels que le niveau d'analyse auquel se situe la métrique (global, lexical, syntaxique, aspectuel etc.) – autant de regroupements qui permettraient de conserver suffisamment de puissance statistique tout en mettant en lumière les dimensions linguistiques les plus significatives au moment de prédire l'agentivité d'un texte. Nous pourrions alors affiner notre indice en lui ôtant les métriques s'avérant en réalité non significativement corrélées, prises individuellement, au statut des personnages, et de là le rendre potentiellement plus performant encore dans la reconnaissance de l'agentivité d'un locuteur.

Enfin, l'objectif ultime de cette étude est diachronique : il s'agit de tester l'hypothèse d'une corrélation positive entre niveau d'agentivité et niveau de ressources dans une société donnée, sur la base d'un paradigme évolutionnaire. Cela nécessiterait de construire un modèle de régression prédisant le niveau d'agentivité global d'un corpus de pièces, tel qu'estimé par l'indice agentif cette fois calculé sur l'ensemble des répliques, à partir d'un indicateur de développement économique (principalement le PIB par habitant et/ou le taux d'urbanisation), ce pour une société donnée et sur plusieurs fenêtres temporelles, en incluant potentiellement l'auteur de la pièce en tant qu'effet aléatoire³. Nous supposons qu'une association linéaire significative existe entre de tels indicateurs et le mégascore agentif des pièces étudiées, ce sur la base des observations recueillies à différentes périodes. En considérant que le niveau de ressources global augmente linéairement depuis 1600 au moins, nous nous attendons plus précisément à ce que le mégascore agentif des pièces augmente avec le temps, sur cette période du moins. Aussi, une analyse à l'échelle des personnages de la pièce, en distinguant à nouveau entre « haut » et « bas » statut, pourrait nous permettre de comprendre la façon dont ce score augmente : la différence entre les scores agentifs des personnages « haut » et « bas » statut se creuse-t-elle ? Ou bien tous les personnages deviennent-ils plus agentifs ? Conformément à notre paradigme évolutionnaire, l'on s'attendrait à ce que tous les personnages, quel que soit leur statut dans la pièce, s'expriment de manière plus agentive avec le temps. Un tel résultat serait d'ailleurs cohérent avec la théorie littéraire qui note souvent l'émergence d'une certaine porosité entre les rôles de maîtres et de serviteurs au théâtre à partir du XVIII^e siècle, à l'image des pièces de Marivaux.

3. Pour une référence méthodologique qui s'appuie sur le même corpus de théâtre que celui utilisé ici : N. Baumard, Elise Huillery, Alexandre Hyafil et Lou Safra, « The cultural evolution of love in literary history », *Nature Human Behaviour*, 6–4 (7 mars 2022), p. 506-522, DOI : [10.1038/s41562-022-01292-z](https://doi.org/10.1038/s41562-022-01292-z)

Sur ce point, on s'attendrait ainsi à ce que l'ajout d'une variable d'interaction entre l'année de publication de la pièce et le statut des personnages dans nos précédents modèles de régression révèle un affaiblissement de l'association linéaire entre statut et indice agentif avec le temps (à savoir une interaction significativement négative entre le statut « bas » et l'année de publication), ce à mesure que les serviteurs deviennent supposément plus agentifs, jusqu'à emprunter le rôle, et ainsi peut être le langage, de leurs maîtres.

Néanmoins, ne disposant pour ce corpus que de 156 années de publication distinctes, nous ne bénéficions pas d'une puissance statistique suffisante pour inclure une telle interaction *per se*. Alternativement, nous pouvons simplement constater que la différence entre le score agentif des personnages « haut » et « bas » statut semble en effet varier avec la date de publication de l'oeuvre :

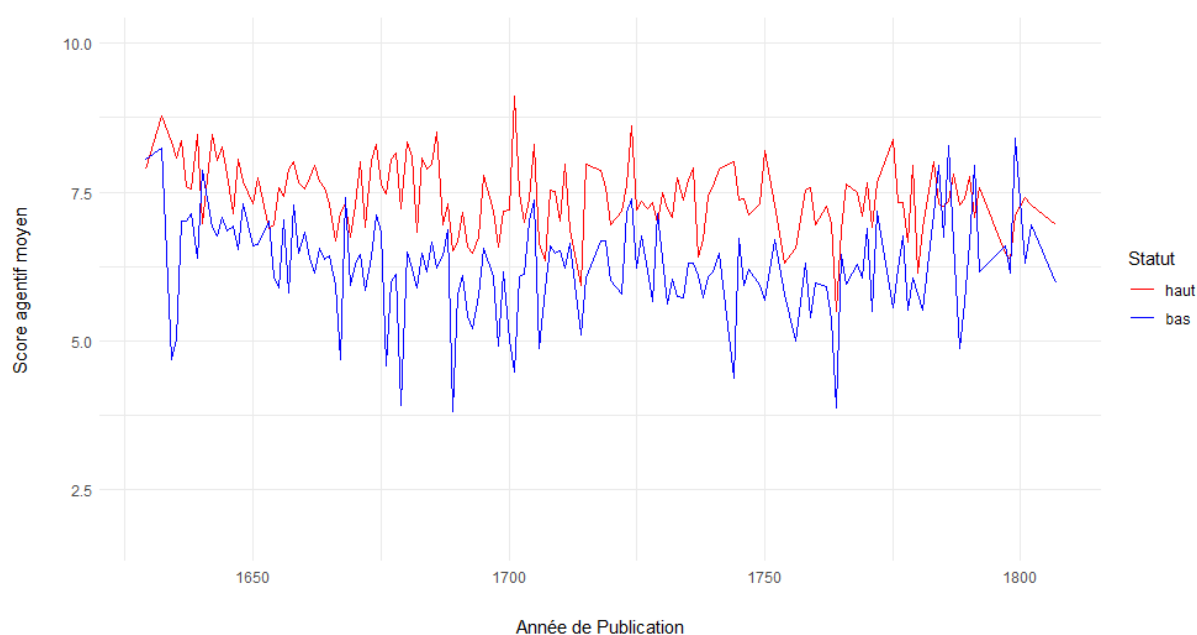


FIGURE 10.1 – Evolution du score agentif moyen des personnages haut et bas statut, selon l'année de publication

Nous avons ici réduit la période d'analyse pour ne s'intéresser qu'à l'intervalle de temps [1629-1807]. Les valeurs étant beaucoup plus espacées et extrêmes à partir de 1808, elles risquaient en effet d'influencer les résultats de façon disproportionnée. En outre, l'on précise que le score agentif analysé ci-après est celui calculé en prenant en compte la direction de corrélation prédite pour chaque métrique prise isolément (version 2 des modèles liés au statut).

De là, un modèle linéaire hiérarchique ou « à effets mixtes » incluant l'année comme intercept aléatoire permettait une première estimation de la façon dont l'effet du statut varie d'une année à l'autre. Plus précisément, en calculant la corrélation intraclasse (ICC), il apparaît qu'environ 10% de la variabilité totale du score agentif est attribuable aux différences entre les années sur la période 1629-1807.

En s'essayant à un premier regroupement des années de publication au sein de deux intervalles distincts ([1629-1699] et [1700-1807]), correspondant approximativement aux deux siècles

d'analyse, et pour lesquels le nombre d'observations était globalement similaire (respectivement 450 et 414), l'on a également pu comparer les coefficients obtenus par notre modèle linéaire simple précédent (expliquant le score agentif par le statut des personnages) entre ces deux sous-périodes. L'on observe que la corrélation négative significative ($p < 0.001$) entre un statut « bas » et le score agentif diminue d'un intervalle à l'autre (passant de -1.163 sur la période 1629-1699 à -1.097 sur 1700-1807), ce qui est cohérent avec nos prédictions.

Malgré les limites de la présente étude, et précisément contre ces limites-mêmes concernant l'élaboration de certaines métriques individuelles, il semble donc que la conduite d'analyses successives soit justifiée (en premier lieu la détermination du caractère distinctif de notre indice). Notre indice agentif global est en effet significativement et très fortement corrélé à l'appartenance des personnages locuteurs à un « haut » statut, relativement aux personnages de « bas » statut, et cela n'est pas simplement lié à une différence dans la longueur ou la fréquence des répliques entre ces deux types de personnages, mais bien à des marqueurs linguistiques plus fins. Il apparaît même que l'on puisse déjà discerner des variations diachroniques dans cette association, bien que leur significativité, ainsi que leur corrélation avec le niveau de ressources environnant, reste encore à prouver.

Conclusion

*« Nous autres esclaves, nous sommes doués contre nos maîtres
d'une pénétration »*

Ce que Marivaux confine à une relation entre une esclave et son maître, qui obligerait la première à un effort d'analyse des attitudes et du langage de celle qui la soumet dans la seule intention de ne pas s'y laisser prendre, et sûrement de prévenir quelques mauvais traitements, pourrait ainsi valoir beaucoup plus globalement. « Nous sommes doués *comme* nos maîtres d'une pénétration », voilà peut-être ce qu'aurait pu dire, à la place, le personnage esclave ou serviteur du XVIII^e, et dévoilant par-là l'état d'esprit de toute une époque. C'est du moins la thèse soutenue dans ce mémoire : celle d'un nivellement par le haut du degré d'agentivité individuel au cours des derniers siècles, déterminé par l'augmentation du niveau des ressources environnantes.

Pour espérer pouvoir retracer et fonder une telle évolution, il nous fallait passer par un support objectif : le langage s'est naturellement imposé comme un terrain d'analyse privilégié.

A travers l'élaboration d'un indice linguistique d'agentivité global, il s'avère finalement possible de prédire, avec une assez grande précision, le statut des locuteurs à l'origine des répliques analysées dans une pièce donnée. Si cet indice manque encore de validité externe, il est un premier pas vers une compréhension plus fine des structures linguistiques profondes, grammaticales et syntaxiques (et outre l'utilisation de pronoms personnels à la première personne), qui pourraient bien soutenir notre discours à l'heure d'un « individualisme » exacerbé, notion intimement liée à celle d'agentivité. Et par là-même, il ouvre la porte à une étude plus approfondie des grands changements psychologiques de l'histoire humaine.

Bien entendu, si notre paradigme fait de ce degré croissant d'agentivité, ou d'individualisme tel que nous l'avons redéfini, une tendance purement mécanique, portée par l'amélioration du confort de vie, cela n'exclut pas qu'il puisse être devenu « mal-adapté » à notre environnement actuel, au sens proprement évolutionnaire du terme. Comme mentionné dans la partie I. 2.2. *Individualisme et pronoms : même méthode, nouveau raisonnement*, l'augmentation apparente des troubles neurotiques sur la période récente⁴ pourrait en ce sens être interprété comme symptomatique d'une stratégie d'identification « haut niveau » dont la quasi-permanence serait devenue sous-optimale dans nos sociétés modernes. En feignant d'ignorer ici des inégalités sociales bien entendu non négligeables, il semble en effet que tout un chacun dispose aujourd'hui, en Occident du moins, d'un niveau de ressources suffisant pour s'engager dans une interprétation particulièrement abstraite et distale de ses actions. Le moindre acte entrepris deviendrait alors le théâtre d'une délibération identitaire qu'il semble difficile de pouvoir supporter au quotidien sans risquer de se perdre, bien loin de l'agir, dans les méandres du moi. Et c'est peut-être ce que la notion d'individualisme, dans son sens péjoratif, pointe précisément du doigt. De là l'importance, également, d'études psycholinguistiques qui devraient venir compléter ces analyses dans un effort de triangulation des données, et qui pourraient envisager y trouver, à terme, des pistes de diagnostic clinique par le biais du langage.

4. J. M. Twenge, « The age of anxiety? The birth cohort change in anxiety and neuroticism, 1952–1993. » *Journal of Personality and Social Psychology*, 79–6 (2000), Place : US Publisher : American Psychological Association, p. 1007-1021, DOI : [10.1037/0022-3514.79.6.1007](https://doi.org/10.1037/0022-3514.79.6.1007)

Annexes

.1 Disponibilité des données du mémoire

Illustrations, code et données du mémoire disponibles sur Github à cette adresse : https://github.com/CamillePerault/Agency_in_language__an_evolutionary_perspective.

Version L^AT_EX disponible sur Github, à cette adresse https://github.com/CamillePerault/redaction_Agency_in_language__an_evolutionary_perspective.

.2 Annexe psycho-linguistique

The Behavior Identification Form

Item	<i>M</i>	Item-total <i>r</i>	Item	<i>M</i>	Item-total <i>r</i>
1. Making a list	0.73	.30	14. Climbing a tree	0.40	.37
a. Getting organized ^a			a. Getting a good view ^a		
b. Writing things down			b. Holding on to branches		
2. Reading	0.87	.29	15. Filling out a personality test	0.69	.31
a. Following lines of print			a. Answering questions		
b. Gaining knowledge ^a			b. Revealing what you're like ^a		
3. Joining the Army	0.47	.39	16. Toothbrushing	0.79	.41
a. Helping the Nation's defense ^a			a. Preventing tooth decay ^a		
b. Signing up			b. Moving a brush around in one's mouth		
4. Washing clothes	0.57	.34	17. Taking a test	0.53	.35
a. Removing odors from clothes ^a			a. Answering questions		
b. Putting clothes into the machine			b. Showing one's knowledge ^a		
5. Picking an apple	0.62	.31	18. Greeting someone	0.74	.35
a. Getting something to eat ^a			a. Saying hello		
b. Pulling an apple off a branch			b. Showing friendliness ^a		
6. Chopping down a tree	0.61	.33	19. Resisting temptation	0.48	.34
a. Wielding an axe			a. Saying "no"		
b. Getting firewood ^a			b. Showing moral courage ^a		
7. Measuring a room for carpeting	0.70	.38	20. Eating	0.59	.43
a. Getting ready to remodel ^a			a. Getting nutrition ^a		
b. Using a yardstick			b. Chewing and swallowing		
8. Cleaning the house	0.60	.37	21. Growing a garden	0.65	.32
a. Showing one's cleanliness ^a			a. Planting seeds		
b. Vacuuming the floor			b. Getting fresh vegetables ^a		
9. Painting a room	0.65	.41	22. Traveling by car	0.78	.30
a. Applying brush strokes			a. Following a map		
b. Making the room look fresh ^a			b. Seeing countryside ^a		
10. Paying the rent	0.65	.48	23. Having a cavity filled	0.47	.41
a. Maintaining a place to live ^a			a. Protecting your teeth ^a		
b. Writing a check			b. Going to the dentist		
11. Caring for houseplants	0.40	.35	24. Talking to a child	0.66	.32
a. Watering plants			a. Teaching a child something ^a		
b. Making the room look nice ^a			b. Using simple words		
12. Locking a door	0.89	.31	25. Pushing a doorbell	0.92	.28
a. Putting a key in the lock			a. Moving a finger		
b. Securing the house ^a			b. Seeing if someone's home ^a		
13. Voting	0.66	.34			
a. Influencing the election ^a					
b. Marking a ballot					

FIGURE 2 – Le « formulaire d'identification des comportements » utilisé par Vallacher (1989) afin d'estimer le degré d'agentivité de ses participants

.3 Résultats détaillés des modèles de régression

```
Call:
lm(formula = sum_normalized_var ~ Status, data = df_catego_status)

Residuals:
    Min       1Q   Median       3Q      Max
-6.6074 -0.4318  0.1028  0.6276  2.6017

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 22.47440    0.04948  454.19  <2e-16 ***
Statushigh   1.01031    0.06982   14.47  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.04 on 886 degrees of freedom
Multiple R-squared:  0.1911,    Adjusted R-squared:  0.1902
F-statistic: 209.4 on 1 and 886 DF,  p-value: < 2.2e-16

>
> glance(lm_status) %>% pull(sigma)
[1] 1.0403
> |
```

FIGURE 3 – Résultats de la régression linéaire prédisant le score agentif par le statut - modèle 1

```
Call:
lm(formula = sum_normalized_var_dir ~ Status, data = df_catego_status_dir)

Residuals:
    Min       1Q   Median       3Q      Max
-8.4880 -0.5369  0.0118  0.6010  3.4106

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.24161    0.04736  131.80  <2e-16 ***
Statushigh   1.11086    0.06682   16.62  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9956 on 886 degrees of freedom
Multiple R-squared:  0.2377,    Adjusted R-squared:  0.2369
F-statistic: 276.3 on 1 and 886 DF,  p-value: < 2.2e-16

>
> glance(lm_status_dir) %>% pull(sigma)
[1] 0.9956491
> |
```

FIGURE 4 – Résultats de la régression linéaire prédisant le score agentif par le statut - modèle 2

```

Call:
glm(formula = Status_binary ~ sum_normalized_var, family = binomial,
    data = train_data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1053  -1.0087   0.4766   0.9743   2.9804

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -23.9369     2.5483  -9.393  <2e-16 ***
sum_normalized_var  1.0383     0.1101   9.429  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 863.65  on 622  degrees of freedom
Residual deviance: 734.71  on 621  degrees of freedom
AIC: 738.71

Number of Fisher Scoring iterations: 4

```

FIGURE 5 – Résultats de la régression logistique prédisant le statut par le score agentif - modèle 1

```

Call:
glm(formula = Status_binary ~ sum_normalized_var_dir, family = binomial,
    data = train_data_dir)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5226  -0.9044   0.3407   0.9240   4.2035

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -7.5741     0.7267 -10.42  <2e-16 ***
sum_normalized_var_dir  1.1099     0.1051  10.56  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 863.65  on 622  degrees of freedom
Residual deviance: 702.52  on 621  degrees of freedom
AIC: 706.52

Number of Fisher Scoring iterations: 4

```

FIGURE 6 – Résultats de la régression logistique prédisant le statut par le score agentif - modèle 2

```

Call:
lm(formula = sum_normalized_var_dir ~ Status, data = interval_data)

Residuals:
    Min       1Q   Median       3Q      Max
-3.9992 -0.5563  0.0268  0.6049  2.8235

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.49303    0.06050  123.84  <2e-16 ***
Statuslow    -1.16268    0.08556  -13.59  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9076 on 448 degrees of freedom
Multiple R-squared:  0.2919,    Adjusted R-squared:  0.2903
F-statistic: 184.6 on 1 and 448 DF,  p-value: < 2.2e-16

```

FIGURE 7 – Résultats du modèle linéaire hiérarchique intégrant l'effet de l'année de publication à la prédiction du score agentif par le statut - modèle 1

```

Call:
lm(formula = sum_normalized_var_dir ~ Status, data = interval_data)

Residuals:
    Min       1Q   Median       3Q      Max
-3.8048 -0.4907  0.0024  0.5325  3.4982

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.25087    0.06848  105.89  <2e-16 ***
Statuslow    -1.09686    0.09731  -11.27  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9899 on 412 degrees of freedom
Multiple R-squared:  0.2357,    Adjusted R-squared:  0.2338
F-statistic: 127.1 on 1 and 412 DF,  p-value: < 2.2e-16

```

FIGURE 8 – Résultats du modèle linéaire hiérarchique intégrant l'effet de l'année de publication à la prédiction du score agentif par le statut - modèle 2

Bibliographie

- ABELE (Andrea E.), HAUKE (Nicole), PETERS (Kim), LOUVET (Eva), SZYMKOW (Aleksandra) et DUAN (Yanping), « Facets of the Fundamental Content Dimensions : Agency with Competence and Assertiveness—Communion with Warmth and Morality », *Frontiers in Psychology*, 7 (2016), DOI : [10.3389/fpsyg.2016.01810](https://doi.org/10.3389/fpsyg.2016.01810).
- AHEARN (Laura M.), « Language and Agency », *Annual Review of Anthropology*, 30–1 (1^{er} oct. 2001), Publisher : Annual Reviews, p. 109-137, DOI : [10.1146/annurev.anthro.30.1.109](https://doi.org/10.1146/annurev.anthro.30.1.109).
- BAKAN (David), *The duality of human existence : An essay on psychology and religion*. Pages : 242, Oxford, England, 1966 (The duality of human existence : An essay on psychology and religion.).
- BAUMARD (Nicolas), HUILLERY (Elise), HYAFIL (Alexandre) et SAFRA (Lou), « The cultural evolution of love in literary history », *Nature Human Behaviour*, 6–4 (7 mars 2022), p. 506-522, DOI : [10.1038/s41562-022-01292-z](https://doi.org/10.1038/s41562-022-01292-z).
- BLOCK (Jack), « A contrarian view of the five-factor approach to personality description. » *Psychological Bulletin*, 117–2 (1995), Place : US Publisher : American Psychological Association, p. 187-215, DOI : [10.1037/0033-2909.117.2.187](https://doi.org/10.1037/0033-2909.117.2.187).
- BYBEE, « The evolution of grammar : tense, aspect, and modality in the languages of the world », *Choice Reviews Online*, 32–8 (1^{er} avr. 1995), p. 32–4321–32-4321, DOI : [10.5860/CHOICE.32-4321](https://doi.org/10.5860/CHOICE.32-4321).
- CAPLAN (Jennifer E.), ADAMS (Kiki) et BOYD (Ryan L.), « Personality and Language », dans *The Wiley Encyclopedia of Personality and Individual Differences*, dir. Bernardo J. Carducci, Christopher S. Nave et Christopher S. Nave, 1^{re} éd., 2020, p. 311-316, DOI : [10.1002/9781118970843.ch52](https://doi.org/10.1002/9781118970843.ch52).
- CÉDRICK FAIRON (Maurice Grevisse), *Le Petit bon usage de la langue française : Grammaire*, 2018.
- CELLI (Fabio) et POESIO (Massimo), « PR2 : A Language Independent Unsupervised Tool for Personality Recognition from Text » (, févr. 2014).
- CHITTARANJAN (Gokul), BLOM (Jan) et GATICA-PEREZ (Daniel), « Who's Who with Big-Five : Analyzing and Classifying Personality Traits with Smartphones », dans *2011 15th Annual International Symposium on Wearable Computers*, San Francisco, CA, USA, 2011, p. 29-36, DOI : [10.1109/ISWC.2011.29](https://doi.org/10.1109/ISWC.2011.29).

- CHUNG (Cindy) et PENNEBAKER (James), « The Psychological Functions of Function Words » ().
- DELATOUR (Yvonne), *Nouvelle grammaire du français : Cours de civilisation française de la Sorbonne*, 1991.
- DEPRAETERE (Ilse), « (A)telicity and intentionality », *Linguistics*, 45–2 (20 janv. 2007), DOI : [10.1515/LING.2007.008](https://doi.org/10.1515/LING.2007.008).
- DEWALL (C. Nathan), POND JR. (Richard S.), CAMPBELL (W. Keith) et TWENGE (Jean M.), « Tuning in to psychological change : Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. » *Psychology of Aesthetics, Creativity, and the Arts*, 5–3 (2011), Place : US Publisher : Educational Publishing Foundation, p. 200-207, DOI : [10.1037/a0023195](https://doi.org/10.1037/a0023195).
- DIGMAN (John M.), « Higher-order factors of the Big Five. » *Journal of Personality and Social Psychology*, 73–6 (1997), Place : US Publisher : American Psychological Association, p. 1246-1256, DOI : [10.1037/0022-3514.73.6.1246](https://doi.org/10.1037/0022-3514.73.6.1246).
- DIUK (Carlos), FERNANDEZ SLEZAK (Diego), RASKOVSKY (Iván), SIGMAN (Mariano) et CECCHI (Guillermo), « A quantitative philology of introspection », *Frontiers in Integrative Neuroscience*, 6 (2012), DOI : [10.3389/fnint.2012.00080](https://doi.org/10.3389/fnint.2012.00080).
- DUBOURG (Edgar) et BAUMARD (Nicolas), « Why and How Did Narrative Fictions Evolve ? Fictions as Entertainment Technologies », *Frontiers in Psychology*, 13 (2022), DOI : [10.3389/fpsyg.2022.786770](https://doi.org/10.3389/fpsyg.2022.786770).
- DUNNING (David), MEYEROWITZ (Judith A.) et HOLZBERG (Amy D.), « Ambiguity and self-evaluation : The role of idiosyncratic trait definitions in self-serving assessments of ability. » *Journal of Personality and Social Psychology*, 57–6 (1989), Place : US Publisher : American Psychological Association, p. 1082-1090, DOI : [10.1037/0022-3514.57.6.1082](https://doi.org/10.1037/0022-3514.57.6.1082).
- EGAN (Thomas), « Manner and Path : evidence from a multilingual corpus », *CogniTextes*, 12 (sept. 2015), DOI : [10.4000/cognitextes.788](https://doi.org/10.4000/cognitextes.788).
- EGAN (Vincent), DEARY (Ian) et AUSTIN (Elizabeth), « The NEO-FFI : Emerging British norms and an item-level analysis suggest N, A and C are more reliable than O and E. » *Personality and Individual Differences*, 29–5 (2000), Place : Netherlands Publisher : Elsevier Science, p. 907-920, DOI : [10.1016/S0191-8869\(99\)00242-1](https://doi.org/10.1016/S0191-8869(99)00242-1).
- FORMANOWICZ (Magdalena), ROESSEL (Janin), SUITNER (Caterina) et MAASS (Anne), « Verbs as linguistic markers of agency : The social side of grammar. » *European Journal of Social Psychology*, 47–5 (2017), Place : US Publisher : John Wiley & Sons, p. 566-579, DOI : [10.1002/ejsp.2231](https://doi.org/10.1002/ejsp.2231).
- FRAWLEY (William), *Linguistic semantics*. Pages : xvii, 533, Hillsdale, NJ, US, 1992 (Linguistic semantics.).
- FRIEDRICH (Annemarie) et GATEVA (Damyana), « Classification of telicity using cross-linguistic annotation projection », dans *Proceedings of the 2017 Conference on Empi-*

- rical Methods in Natural Language Processing*, Copenhagen, Denmark, 2017, p. 2559-2565, DOI : [10.18653/v1/D17-1271](https://doi.org/10.18653/v1/D17-1271).
- FRIEDRICH (Annemarie) et PALMER (Alexis), « Automatic prediction of aspectual class of verbs in context », dans *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)*, Baltimore, Maryland, 2014, p. 517-523, DOI : [10.3115/v1/P14-2085](https://doi.org/10.3115/v1/P14-2085).
- FUJITA (Kentaro), « Seeing the forest beyond the trees : A construal-level approach to self-control. » *Social and Personality Psychology Compass*, 2–3 (2008), Place : United Kingdom Publisher : Wiley-Blackwell Publishing Ltd., p. 1475-1496, DOI : [10.1111/j.1751-9004.2008.00118.x](https://doi.org/10.1111/j.1751-9004.2008.00118.x).
- GARDNER (Wendi L.), GABRIEL (Shira) et LEE (Angela Y.), « "I" value freedom, but "we" value relationships : Self-construal priming mirrors cultural differences in judgment. » *Psychological Science*, 10–4 (1999), Place : United Kingdom Publisher : Blackwell Publishing, p. 321-326, DOI : [10.1111/1467-9280.00162](https://doi.org/10.1111/1467-9280.00162).
- GIDDENS (Anthony), *Modernity and Self-Identity : Self and Society in the Late Modern Age*, Stanford, 1991, URL : <http://www.sup.org/books/title/?id=2660> (visité le 09/09/2023).
- GILL (Alastair), NOWSON (Scott) et OBERLANDER (Jon), « What Are They Blogging About ? Personality, Topic and Motivation in Blogs », *Proceedings of the International AAAI Conference on Web and Social Media*, 3–1 (19 mars 2009), p. 18-25, DOI : [10.1609/icwsm.v3i1.13949](https://doi.org/10.1609/icwsm.v3i1.13949).
- GOUREVITCH (Aron-J.), *La Naissance de l'individu dans l'Europe médiévale*, 1997.
- GRAF (Tim), PHILIPP (Markus), XU (Xiaonan), KRETZSCHMAR (Franziska) et PRIMUS (Beatrice), « The interaction between telicity and agentivity : Experimental evidence from intransitive verbs in German and Chinese », *Lingua* (, oct. 2017), DOI : [10.1016/j.lingua.2017.08.006](https://doi.org/10.1016/j.lingua.2017.08.006).
- GRANATH (Solveig) et WHERRITY (Michael), « "I'm loving you - and knowing it too" : Aspect and so-called stative verbs », dans 2014, URL : <https://api.semanticscholar.org/CorpusID:142062193>.
- HAGGARD P (Chambon V.), « Sense of agency », *Curr Biol*. 2012 May 22 ;22(10) :R390-2. (, DOI : [doi:10.1016/j.cub.2012.02.040](https://doi.org/doi:10.1016/j.cub.2012.02.040). PMID:22625851..
- HEYLIGHEN (Francis) et DEWAELE (Jean-Marc), « Variation in the Contextuality of Language : An Empirical Measure », *Foundations of Science*, 7–3 (1^{er} sept. 2002), p. 293-340, DOI : [10.1023/A:1019661126744](https://doi.org/10.1023/A:1019661126744).
- HILLS (Thomas T.) et BUTTERFILL (Stephen), « From foraging to autonoeitic consciousness : The primal self as a consequence of embodied prospective foraging », *Current Zoology*, 61–2 (avr. 2015), _eprint : <https://academic.oup.com/cz/article-pdf/61/2/368/32970825/czoolo61-0368.pdf>, p. 368-381, DOI : [10.1093/czoolo/61.2.368](https://doi.org/10.1093/czoolo/61.2.368).

- HILPERT (Martin), « Change in modal meanings : Another look at the shifting collocates of *may* », *Constructions and Frames*, 8–1 (29 sept. 2016), p. 66-85, DOI : [10.1075/cf.8.1.05hil](https://doi.org/10.1075/cf.8.1.05hil).
- HILPERT (Martin) et FLACH (Susanne), « Disentangling modal meanings with distributional semantics », *Digital Scholarship in the Humanities*, 36–2 (29 sept. 2021), p. 307-321, DOI : [10.1093/llc/fqaa014](https://doi.org/10.1093/llc/fqaa014).
- HITLIN (Steven) et GLEN H. ELDER (Jr.), « Time, Self, and the Curiously Abstract Concept of Agency », *Sociological Theory*, 25–2 (2007), _eprint : <https://doi.org/10.1111/j.1467-9558.2007.00303.x>, p. 170-191, DOI : [10.1111/j.1467-9558.2007.00303.x](https://doi.org/10.1111/j.1467-9558.2007.00303.x).
- HOPPER (Paul) et THOMPSON (Sandra), « Transitivity in Grammar and Discourse », *Language*, 56 (juin 1980), DOI : [10.2307/413757](https://doi.org/10.2307/413757).
- HURLEY (John R.), « Agency and Communion as Related to “Big Five” Self-Representations and Subsequent Behavior in Small Groups », *The Journal of Psychology*, 132–3 (1998), Publisher : Routledge _eprint : <https://doi.org/10.1080/00223989809599172>, p. 337-351, DOI : [10.1080/00223989809599172](https://doi.org/10.1080/00223989809599172).
- JACQUELINE OLLIVIER (Martin Beaudoin), *Grammaire Française*, 2007.
- JOHANSSON (Stig), « Modals and semi-modals of obligation in American English : some aspects of developments from 1990 until the present day », dans *The Verb Phrase in English*, dir. Bas Aarts, *et al.*, 1^{re} éd., 2013, p. 372-380, DOI : [10.1017/CB09781139060998.016](https://doi.org/10.1017/CB09781139060998.016).
- KASHIMA (Emiko S.) et KASHIMA (Yoshihisa), « Culture and language : The case of cultural dimensions and personal pronoun use. » *Journal of Cross-Cultural Psychology*, 29–3 (1998), Place : US Publisher : Sage Publications, p. 461-486, DOI : [10.1177/0022022198293005](https://doi.org/10.1177/0022022198293005).
- KERN (Margaret L.), EICHSTAEDT (Johannes C.), SCHWARTZ (H. Andrew), PARK (Gregory), UNGAR (Lyle H.), STILLWELL (David J.), KOSINSKI (Michal), DZIURZYNSKI (Lukasz) et SELIGMAN (Martin E. P.), « From “Sooo excited!!!” to “So proud” : Using language to study development. » *Developmental Psychology*, 50–1 (2014), p. 178-188, DOI : [10.1037/a0035048](https://doi.org/10.1037/a0035048).
- KLAVANS (Judith L.) et CHODOROW (Martin), « Degrees of Stativity : The Lexical Representation of Verb Aspect », dans *COLING 1992 Volume 4 : The 14th International Conference on Computational Linguistics*, 1992, URL : <https://aclanthology.org/C92-4177>.
- KLINGE (Alex), « The Impact of Context on Modal Meaning in English and Danish », *Nordic Journal of Linguistics*, 19 (1996), p. 35-54, URL : <https://api.semanticscholar.org/CorpusID:145424710>.
- KONOPASKY (Abigail W.) et SHERIDAN (Kimberly M.), « Towards a Diagnostic Toolkit for the Language of Agency », *Mind, Culture, and Activity*, 23–2 (2016), Publisher :

- Routledge _eprint : <https://doi.org/10.1080/10749039.2015.1128952>, p. 108-123, DOI : [10.1080/10749039.2015.1128952](https://doi.org/10.1080/10749039.2015.1128952).
- LEECH (Geoffrey), HUNDT (Marianne), MAIR (Christian) et SMITH (Nicholas), « Change in Contemporary English : A Grammatical Study » ().
- LOCKE (Kenneth), « Agency and Communion in Social Comparison », dans 2014, p. 11-38, DOI : [10.1017/cbo9781139035583.003](https://doi.org/10.1017/cbo9781139035583.003).
- MAIRESSE (F.), WALKER (M. A.), MEHL (M. R.) et MOORE (R. K.), « Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text », *Journal of Artificial Intelligence Research*, 30 (28 nov. 2007), p. 457-500, DOI : [10.1613/jair.2349](https://doi.org/10.1613/jair.2349).
- MARIAN (Viorica) et KAUSHANSKAYA (Margarita), « Self-construal and emotion in bi-cultural bilinguals. » *Journal of Memory and Language*, 51-2 (2004), Place : Netherlands Publisher : Elsevier Science, p. 190-201, DOI : [10.1016/j.jml.2004.04.003](https://doi.org/10.1016/j.jml.2004.04.003).
- MASLOW (A. H.), « A theory of human motivation. » *Psychological Review*, 50-4 (1943), Place : US Publisher : American Psychological Association, p. 370-396, DOI : [10.1037/h0054346](https://doi.org/10.1037/h0054346).
- MCCRAE (Robert R.) et COSTA (Paul T.), « The structure of interpersonal traits : Wiggins's circumplex and the five-factor model. » *Journal of Personality and Social Psychology*, 56-4 (1989), Place : US Publisher : American Psychological Association, p. 586-595, DOI : [10.1037/0022-3514.56.4.586](https://doi.org/10.1037/0022-3514.56.4.586).
- MEHTA (Yash), FATEHI (Samin), KAZAMEINI (Amirmohammad), STACHL (Clemens), CAMBRIA (Erik) et EETEMADI (Sauleh), « Bottom-Up and Top-Down : Predicting Personality with Psycholinguistic and Language Model Features », dans *2020 IEEE International Conference on Data Mining (ICDM)*, Sorrento, Italy, 2020, p. 1184-1189, DOI : [10.1109/ICDM50108.2020.00146](https://doi.org/10.1109/ICDM50108.2020.00146).
- MYHILL (John), « Change and continuity in the functions of the American English modals », *ling*, 33-2 (1995), p. 157-212, DOI : [10.1515/ling.1995.33.2.157](https://doi.org/10.1515/ling.1995.33.2.157).
- NA (Jinkyung) et CHOI (Incheol), « Culture and first-person pronouns. » *Personality and Social Psychology Bulletin*, 35-11 (2009), Place : US Publisher : Sage Publications, p. 1492-1499, DOI : [10.1177/0146167209343810](https://doi.org/10.1177/0146167209343810).
- NERBONNE (John), « The Secret Life of Pronouns. What Our Words Say About Us », *Literary and Linguistic Computing*, 29 (avr. 2014), p. 139-142, DOI : [10.1093/llc/fqt006](https://doi.org/10.1093/llc/fqt006).
- NERIA (Yuval), « Encyclopedia of Personality and Individual Differences », dans 2016, p. 1-3, DOI : [10.1007/978-3-319-28099-8_2026-1](https://doi.org/10.1007/978-3-319-28099-8_2026-1).
- NESSELHAUF (Nadja), « Mechanisms of language change in a functional system : The recent semantic evolution of English future time expressions », *Journal of Historical Linguistics*, 2-1 (25 juill. 2012), p. 83-132, DOI : [10.1075/jhl.2.1.06nes](https://doi.org/10.1075/jhl.2.1.06nes).

- OBERLANDER (Jon) et NOWSON (Scott), « Whose thumb is it anyway? : classifying author personality from weblog text », dans *Proceedings of the COLING/ACL on Main conference poster sessions* -, Sydney, Australia, 2006, p. 627-634, DOI : [10.3115/1273073.1273154](https://doi.org/10.3115/1273073.1273154).
- PALMER (F.), *Mood and Modality*, 1986, URL : [doi:10.1017/CB09781139167178](https://doi.org/10.1017/CB09781139167178).
- PAYEN (Jean-Charles), *L'humanisme médiéval et la redécouverte de l'individu en Occident du XIIIe siècle à la fin du XIIIe*, 1985.
- PENNEBAKER (James) et CHUNG (Cindy), « Expressive Writing, Emotional Upheavals, and Health » (, janv. 2007).
- PENNEBAKER (James W) et KING (Laura A), « Linguistic Styles : Language Use as an Individual Difference » ().
- PENNEBAKER (James W.), CHUNG (Cindy K.), FRAZEE (Joey), LAVERGNE (Gary M.) et BEAVER (David I.), « When Small Words Foretell Academic Success : The Case of College Admissions Essays », *PLOS ONE*, 9–12 (déc. 2015), Publisher : Public Library of Science, p. 1-10, DOI : [10.1371/journal.pone.0115844](https://doi.org/10.1371/journal.pone.0115844).
- PIETRASZKIEWICZ (Agnieszka), FORMANOWICZ (Magdalena), GUSTAFSSON SENDÉN (Marie), BOYD (Ryan L.), SIKSTRÖM (Sverker) et SCZESNY (Sabine), « The big two dictionaries : Capturing agency and communion in natural language », *European Journal of Social Psychology*, 49–5 (août 2019), p. 871-887, DOI : [10.1002/ejsp.2561](https://doi.org/10.1002/ejsp.2561).
- PORTNER (Paul), *Modality*, 2010, URL : [doi:10.1017/S0022226709990351](https://doi.org/10.1017/S0022226709990351).
- ROSENWEIN ({} and {} C. S. Heppleston. Barbara H.), “Y avait-il un «moi» au haut Moyen Âge?”, 2005, URL : <http://www.jstor.org/stable/40957516>..
- ROTHSTEIN (Susan), « What are Incremental Themes? » ().
- ROUHIZADEH (Masoud), JAIDKA (Kokil), SMITH (Laura), SCHWARTZ (H. Andrew), BUFFONE (Anneke) et UNGAR (Lyle), « Identifying Locus of Control in Social Media Language », dans *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, 2018, p. 1146-1152, DOI : [10.18653/v1/D18-1145](https://doi.org/10.18653/v1/D18-1145).
- RYMES (Betsy R.), « The Construction of Moral Agency in the Narratives of High-School Drop-Outs », *Discourse & Society*, 6 (1995), p. 495-516, URL : <https://api.semanticscholar.org/CorpusID:145568427>.
- SCHWARTZ (H. Andrew), EICHSTAEDT (Johannes C.), KERN (Margaret L.), DZIURZYNSKI (Lukasz), RAMONES (Stephanie M.), AGRAWAL (Megha), SHAH (Achal), KOSINSKI (Michal), STILLWELL (David), SELIGMAN (Martin E. P.), *et al.*, « Personality, Gender, and Age in the Language of Social Media : The Open-Vocabulary Approach », dir. Tobias Preis, *PLoS ONE*, 8–9 (25 sept. 2013), e73791, DOI : [10.1371/journal.pone.0073791](https://doi.org/10.1371/journal.pone.0073791).

- SELIGMAN (Martin), « Psychological history and predicting the future », *Possibility Studies & Society*, 1–1 (2023), _eprint : <https://doi.org/10.1177/27538699221128224>, p. 206-210, DOI : [10.1177/27538699221128224](https://doi.org/10.1177/27538699221128224).
- SIEGEL (Eric V.) et MCKEOWN (Kathleen R.), « Learning Methods to Combine Linguistic Indicators : Improving Aspectual Classification and Revealing Linguistic Insights », *Computational Linguistics*, 26–4 (déc. 2000), p. 595-628, DOI : [10.1162/089120100750105957](https://doi.org/10.1162/089120100750105957).
- SIMS (Matthew), PARK (Jong Ho) et BAMMAN (David), « Literary Event Detection », dans *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy, 2019, p. 3623-3634, DOI : [10.18653/v1/P19-1353](https://doi.org/10.18653/v1/P19-1353).
- SMITH (Nicholas), « Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English », dans *Modality in Contemporary English*, dir. Roberta Facchinetti, Frank Palmer et Manfred Krug, Berlin, Boston, 2003, p. 241-266, DOI : [doi:10.1515/9783110895339.241](https://doi.org/10.1515/9783110895339.241).
- SNEFJELLA (Bryor), GÉNÉREUX (Michel) et KUPERMAN (Victor), « Historical evolution of concrete and abstract language revisited », *Behavior Research Methods*, 51–4 (1^{er} août 2019), p. 1693-1705, DOI : [10.3758/s13428-018-1071-2](https://doi.org/10.3758/s13428-018-1071-2).
- STAJNER (Sanja) et YENIKENT (Seren), « A Survey of Automatic Personality Detection from Texts », dans *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain (Online), 2020, p. 6284-6295, DOI : [10.18653/v1/2020.coling-main.553](https://doi.org/10.18653/v1/2020.coling-main.553).
- STRIK LIEVERS (Francesca), BOLOGNESI (Marianna) et WINTER (Bodo), « The linguistic dimensions of concrete and abstract concepts : lexical category, morphological structure, countability, and etymology », *Cognitive Linguistics*, 32–4 (25 nov. 2021), p. 641-670, DOI : [10.1515/cog-2021-0007](https://doi.org/10.1515/cog-2021-0007).
- TALMY (Leonard), « Path to Realization : A Typology of Event Conflation », dans 1991, URL : <https://api.semanticscholar.org/CorpusID:60877805>.
- TROPE (Yaacov) et LIBERMAN (Nira), « Construal-level theory of psychological distance. » *Psychological Review*, 117–2 (2010), Place : US Publisher : American Psychological Association, p. 440-463, DOI : [10.1037/a0018963](https://doi.org/10.1037/a0018963).
- TWENGE (Jean M.), « The age of anxiety? The birth cohort change in anxiety and neuroticism, 1952–1993. » *Journal of Personality and Social Psychology*, 79–6 (2000), Place : US Publisher : American Psychological Association, p. 1007-1021, DOI : [10.1037/0022-3514.79.6.1007](https://doi.org/10.1037/0022-3514.79.6.1007).
- *Generation Me : Why today's young Americans are more confident, assertive, entitled—and more miserable than ever before*. Pages : vii, 292, New York, NY, US, 2006 (Generation Me : Why today's young Americans are more confident, assertive, entitled—and more miserable than ever before.).

- TWENGE (Jean M.), CAMPBELL (W. Keith) et GENTILE (Brittany), « Generational increases in agentic self-evaluations among American college students, 1966–2009. » *Self and Identity*, 11–4 (2012), Place : United Kingdom Publisher : Taylor & Francis, p. 409-427, DOI : [10.1080/15298868.2011.576820](https://doi.org/10.1080/15298868.2011.576820).
- « Changes in pronoun use in American books and the rise of individualism, 1960–2008. » *Journal of Cross-Cultural Psychology*, 44–3 (2013), Place : US Publisher : Sage Publications, p. 406-415, DOI : [10.1177/0022022112455100](https://doi.org/10.1177/0022022112455100).
- VALLACHER (Robin R.) et WEGNER (Daniel M.), « Levels of personal agency : Individual variation in action identification. » *Journal of Personality and Social Psychology*, 57–4 (1989), Place : US Publisher : American Psychological Association, p. 660-671, DOI : [10.1037/0022-3514.57.4.660](https://doi.org/10.1037/0022-3514.57.4.660).
- VENDLER (Zeno), « “Verbs and Times.” » *The Philosophical Review*, vol. 66, no. 2, 1957, pp. 143–160 (), DOI : <https://doi.org/10.2307/2182371>.
- VON FINTEL (Kai), « Modality and language », dans *Encyclopedia of Philosophy*, dir. D. Borchert, 2006, t. 2, p. 20-27.
- WALKOVÁ (Milada), « Particle verbs in English : Telicity or scalarity ? », *Linguistics*, 55–3 (2017), p. 589-616, DOI : [doi:10.1515/ling-2017-0005](https://doi.org/doi:10.1515/ling-2017-0005).
- WÄRNSBY (Anna), *(De)coding modality : the case of must, may, måste and kan*, Lund, 2006 (Lund studies in English, 113).
- WIGGINS (Jerry S.), « A psychological taxonomy of trait-descriptive terms : The interpersonal domain. » *Journal of Personality and Social Psychology*, 37–3 (1979), Place : US Publisher : American Psychological Association, p. 395-412, DOI : [10.1037/0022-3514.37.3.395](https://doi.org/10.1037/0022-3514.37.3.395).
- « Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. » Dans *Thinking clearly about psychology : Essays in honor of Paul E. Meehl*, Vol. 1 : Matters of public interest ; Vol. 2 : Personality and psychopathology. Minneapolis, MN, US, 1991, p. 89-113.
- WOIKE (Barbara A.), « The use of differentiation and integration processes : Empirical studies of "separate" and "connected" ways of thinking. » *Journal of Personality and Social Psychology*, 67–1 (1994), Place : US Publisher : American Psychological Association, p. 142-150, DOI : [10.1037/0022-3514.67.1.142](https://doi.org/10.1037/0022-3514.67.1.142).
- YARKONI (Tal), « Personality in 100,000 Words : A large-scale analysis of personality and word use among bloggers », *Journal of Research in Personality*, 44–3 (juin 2010), p. 363-373, DOI : [10.1016/j.jrp.2010.04.001](https://doi.org/10.1016/j.jrp.2010.04.001).
- YU (Feng), PENG (Theodore), PENG (Kaiping), TANG (Shi), CHEN (Chuan Shi), QIAN (Xiaojun), SUN (Pei), HAN (Tingting) et CHAI (Fangyuan), « Cultural value shifting in pronoun use. » *Journal of Cross-Cultural Psychology*, 47–2 (2016), Place : US Publisher : Sage Publications, p. 310-316, DOI : [10.1177/0022022115619230](https://doi.org/10.1177/0022022115619230).

- ZARCONI (Alessandra) et LENCI (Alessandro), « Computational Models for Event Type Classification in Context », dans *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, 2008, URL : http://www.lrec-conf.org/proceedings/lrec2008/pdf/315_paper.pdf.
- ZHAO (Yiyun), NGUI (Jian Gang), HALL HARTLEY (Lucy) et BETHARD (Steven), « Do pretrained transformers infer telicity like humans? », dans *Proceedings of the 25th Conference on Computational Natural Language Learning*, Online, 2021, p. 72-81, DOI : [10.18653/v1/2021.conll-1.6](https://doi.org/10.18653/v1/2021.conll-1.6).

Table des figures

1.1	Typologie de l'agentivité s'appuyant sur l'horizon temporel de l'action - Hitlin et Edler (2007)	11
2.1	Elements constitutifs de la transitivité selon Hopper et Thompson (1980)	15
2.2	Eléments participant de l'individuation de l'objet grammatical selon Hopper et Thompson (1980)	18
3.1	La Pyramide des besoins de Maslow (1943) intégrée à un paradigme évolutionnaire	26
4.1	Le Circomplexe interpersonnel de Wiggins (1995) et le Modèle à 5 facteurs ou <i>Big Five</i>	32
5.1	Rouhizadeh et al. (2018) - Coefficients de régression logistique entre les catégories de verbes et le contrôle (les valeurs positives correspondent à la pertinence du contrôle et au contrôle externe). ** : $p < 0,01$ après correction de Benjamini-Hochberg. Les caractéristiques ayant des coefficients positifs sont en bleu et celles ayant des coefficients négatifs sont en rouge	39
6.1	Représentation des dénominations personnelles en fonction du degré d' <i>animation</i> qu'elles évoquent - Ahearn (2001)	51
10.1	Evolution du score agentif moyen des personnages haut et bas statut, selon l'année de publication	80
2	Le « formulaire d'identification des comportements » utilisé par Vallacher (1989) afin d'estimer le degré d'agentivité de ses participants	89
3	Résultats de la régression linéaire prédisant le score agentif par le statut - modèle 1	90
4	Résultats de la régression linéaire prédisant le score agentif par le statut - modèle 2	90
5	Résultats de la régression logistique prédisant le statut par le score agentif - modèle 1	91
6	Résultats de la régression logistique prédisant le statut par le score agentif - modèle 2	91
7	Résultats du modèle linéaire hiérarchique intégrant l'effet de l'année de publication à la prédiction du score agentif par le statut - modèle 1	92
8	Résultats du modèle linéaire hiérarchique intégrant l'effet de l'année de publication à la prédiction du score agentif par le statut - modèle 2	92

Liste des tableaux

4.1	Récapitulatif des marqueurs linguistiques retenus	33
5.1	Dictionnaire de l'Agentivité de Pietraszkiewicz	38
5.2	Les mots agentifs retenus par Seligman	39
5.3	Expressions adverbiales à valeur intentionnelle retenues	40
5.4	Expressions adverbiales à valeur de doute retenues	42
5.5	Expressions adverbiales à valeur d'affirmation retenues	42
7.1	Expressions à valeur de but retenues	58
7.2	Expressions adverbiales de manière retenues	61
7.3	Expressions adverbiales de lieu retenues	61
7.4	Expressions adverbiales temporelles retenues	62
7.5	Lexique univoquement temporel retenu en complément des expressions adverbiales précédentes pour l'analyse regex	63
8.1	Récapitulatif des marqueurs linguistiques distinctifs identifiés pour les deux caté- gories de traits d'intérêt du modèle des <i>Big Five</i>	66
9.1	Score agentif par statut - modèle 1	72
9.2	Score agentif par statut - modèle 2	73
9.3	Résultats de l'évaluation des modèles logistiques relatifs au statut en validation croisée	74
9.4	Score agentif par genre	74
9.5	Résultats de l'évaluation du modèle logistique relatif au genre en validation croisée	75