

# Cognition des relations humaines dans *Intouchables* : une analyse psycholinguistique multimodale du langage, du geste et du regard

Matthew Adekunle

Bamidele Olumilua University of Education, Science and Technology,  
Ikere-Ekiti, Nigeria

**RÉSUMÉ.** Cette recherche examine comment langage, geste et regard interagissent pour façonner la cognition des relations humaines dans le film *Intouchables*. Alors que la psycholinguistique s'est surtout attachée à l'oral, à l'écrit et à l'acquisition des langues, le cinéma demeure un champ encore peu étudié malgré son potentiel multimodal. En mobilisant la grammaire générative transformationnelle de Chomsky et le modèle MDB-GSG (parole, geste, regard), l'étude analyse quinze séquences dialogiques du film, segmentées et annotées selon une grille de catégorisation multimodale, afin d'identifier les structures profondes et de surface, les transformations syntaxiques ainsi que les fonctions interactionnelles des indices non verbaux. Les résultats montrent que l'humour constitue un mécanisme cognitif et interactionnel majeur, capable de désamorcer la stigmatisation et de consolider la complicité entre Philippe, aristocrate tétraplégique, et Driss, aide-soignant issu de l'immigration. L'analyse révèle que l'ambiguïté syntaxique, lorsqu'elle est amplifiée par les gestes et les regards, produit des effets d'intimité et de solidarité qui transforment la différence en rapprochement. En articulant syntaxe et multimodalité, cette étude propose un cadre analytique transférable, contribuant à renouveler les débats contemporains en psycholinguistique et en études filmiques.

**MOTS-CLÉS :** psycholinguistique ; multimodalité ; grammaire générative ; humour ; *Intouchables*.

**ABSTRACT.** This research examines how language, gesture, and gaze interact to shape the cognition of human relationships in the film *Intouchables*. While psycholinguistics has primarily focused on speech, writing, and language acquisition, cinema remains a relatively underexplored field despite its multimodal potential. Drawing on Chomsky's transformational-generative grammar and the MDB-GSG model (speech, gesture, gaze), the study analyzes fifteen segmented dialogic sequences in order to

identify deep and surface structures, syntactic transformations, as well as the interactional functions of non-verbal cues. The findings reveal that humor serves as a major cognitive and interactional mechanism, capable of defusing stigmatization and reinforcing complicity between Philippe, a tetraplegic aristocrat, and Driss, a caregiver of immigrant background. The analysis further shows that syntactic ambiguity, when amplified by gestures and gazes, produces effects of intimacy and solidarity that transform difference into closeness. By bringing syntax and multimodality into dialogue, this study proposes a transferable analytical framework, contributing to the renewal of contemporary debates in psycholinguistics and film studies.

KEYWORDS: psycholinguistics; multimodality; generative grammar; humour; *Intouchables*.

## Introduction

La communication humaine constitue un phénomène complexe qui dépasse la simple transmission d'informations pour englober un ensemble d'opérations cognitives et interactionnelles où s'articulent langage, gestes, regard et intonation. Depuis plusieurs décennies, la psycholinguistique s'attache à analyser la manière dont les individus produisent, interprètent et négocient le sens à travers des processus mentaux qui lient le langage aux mécanismes cognitifs sous-jacents (Fernandez & Cairns, 2010 : 210-212 ; Harley, 2005 : 198 ; Levelt, 2012 : 21). Ce champ, situé à la croisée de la psychologie cognitive et de la linguistique, met en évidence le rôle de la syntaxe, de la mémoire, de l'attention et des affects dans la compréhension et la production de la parole (Altman, 2001 : 79-83 ; Harris et al., 2012 : 252). Or, si ces analyses ont largement été appliquées à la langue orale, écrite ou à l'acquisition des langues, elles demeurent encore relativement marginales dans l'étude des œuvres filmiques, alors même que le cinéma offre un terrain privilégié où se manifestent, de manière synchronisée, parole, geste et regard (Chare & Watkins, 2017 : 6 ; Mondada, 2024 : 12).

Le film *Intouchables* (Nakache & Toledano, 2011), inspiré d'une histoire réelle, constitue un exemple particulièrement fécond pour explorer cette articulation entre langage et cognition relationnelle. Le récit met en scène la rencontre improbable entre Philippe, riche aristocrate tétraplégique, et Driss, jeune homme d'origine immigrée issu des banlieues parisiennes. Leur amitié progressive, forgée au croisement de la différence sociale, culturelle et physique, illustre la puissance du langage; verbal et non verbal comme instrument de solidarité, d'humour et de désamorçage de la stigmatisation (Hernanda & Firmonasari, 2015 : 69 ; King, 2018 : 153-155). À travers leurs échanges, la parole devient non seulement un outil de communication mais aussi un vecteur de lien affectif. Le langage de Driss, souvent familier et spontané, entre en tension avec celui de Philippe, plus élaboré et distancié, révélant la plasticité du langage comme miroir de la cognition sociale et émotionnelle. Le langage de Driss inclut également des formes d'argot urbain, notamment le verlan, procédé consistant à

inverser les syllabes d'un mot. Ce registre contribue à marquer l'identité sociale du personnage et renforce l'effet humoristique dans certaines scènes.

Dans ce contexte, analyser les échanges dialogiques, mais aussi les gestes, postures et regards, permet de mieux comprendre comment se construisent les liens interpersonnels et comment la cognition s'exprime à travers des ressources multimodales (Skogmyr Marian, 2024 : 22-26 ; Tseng, Laubrock & Bateman, 2021 : 18). Le cinéma, par sa nature même, offre une fenêtre privilégiée sur ces dynamiques, en permettant d'observer comment la grammaire visuelle et la grammaire linguistique se répondent mutuellement. L'interprétation d'une scène ne dépend pas uniquement du texte prononcé mais aussi de la manière dont celui-ci s'incarne c'est-à-dire ton, rythme, prosodie, geste ou mouvement de caméra, révélant une interconnexion profonde entre cognition, syntaxe et perception.

L'originalité de la présente recherche réside dans l'articulation de la grammaire générative transformationnelle (Chomsky, 1957 : 65 ; Chomsky, 1965 : 15-17 ; Chomsky, 1995 : 201 ; Chomsky, 2006 : 103) avec un modèle multimodal intégrant la parole, le geste et le regard (Quek et al., 2002 : 96-97). La grammaire générative, en distinguant structures profondes et structures de surface, fournit un cadre pour analyser les transformations syntaxiques et leurs implications sémantiques (Van Gelderen, 2015 : 109 ; Yule, 2016 : 98). En d'autres termes, elle offre une grille de lecture permettant d'observer comment des structures sous-jacentes universelles se manifestent dans des contextes particuliers de communication. Appliquée au dialogue filmique, cette approche éclaire la manière dont les choix syntaxiques, souvent subtils, contribuent à la création d'effets émotionnels, humoristiques ou symboliques.

Ces structures, lorsqu'elles sont appliquées aux dialogues filmiques, révèlent comment l'ambiguïté, l'humour et l'intimité sont produits par la manipulation des règles syntaxiques et leur interaction avec des indices non verbaux (Lew, 1997 : 139-143 ; Canestrari & Bianchi, 2012 : 551-553). Dans la continuité des travaux de Raskin, Attardo (2001 : 155-159) développe la « General Theory of Verbal Humor », qui propose plusieurs paramètres permettant d'analyser les mécanismes humoristiques, tels que la structure narrative, les stratégies linguistiques et les oppositions sémantiques. Cette approche permet d'examiner comment différentes ressources linguistiques contribuent à la production de l'humour dans le discours. Ainsi, un simple déplacement de constituants ou une inversion thématique peut générer un effet de connivence ou d'ironie, renforcé par un regard ou un sourire complice. Loin d'être de simples ornements discursifs, ces mécanismes jouent un rôle crucial dans la construction cognitive des relations humaines (Mahlberg et al., 2014 : 381 ; Zasiékina, 2015 : 153). Ils montrent que la compréhension du langage ne saurait être dissociée de l'expérience sensorielle et affective du sujet parlant.

De plus, la littérature contemporaine en multimodalité insiste sur la nécessité d'intégrer les gestes et le regard à l'analyse linguistique. Les travaux de Kress et Van Leeuwen (2006 : 107-110) sur la grammaire visuelle, ceux de Forceville et Urios-Aparisi (2009 : 2850) sur la métaphore multimodale, ainsi que ceux de Bateman (2014 : 183) sur la sémiotique filmique montrent que le sens naît de la

convergence de plusieurs modalités. Dans le cas d'*Intouchables*, l'humour repose souvent sur cette combinaison : une réplique syntaxiquement ambiguë, renforcée par un geste ou un regard, produit un effet de surprise qui engage à la fois cognition et affect (Markhabayeva & Tseng, 2024 : 7 ; Ding & Chen, 2022 : 176).

Dans le domaine de la linguistique de l'humour, plusieurs approches ont cherché à expliquer les mécanismes qui produisent l'effet comique dans le discours. Raskin (1985 : 325-335) propose la « Semantic Script Theory of Humor », selon laquelle un énoncé devient humoristique lorsqu'il peut être interprété selon deux scripts sémantiques incompatibles mais coexistants ; le passage de l'un à l'autre génère un effet de surprise qui déclenche le rire. Attardo (2001 : 155-159) et d'autres auteurs ont élargi cette perspective en identifiant des stratégies discursives et des ressources linguistiques qui favorisent l'humour verbal. Dans une approche pragmatique et cognitive, Dynel (2009 : 123-131) montre que l'humour repose fréquemment sur des ambiguïtés interprétatives et des effets de surprise dans le traitement du discours.

Cette théorie peut être illustrée par des scènes du film *Intouchables*. Par exemple, la scène du rasage condense l'articulation entre humour verbal et indices non verbaux : le jeu de regards, la tonalité ironique de la voix et la gestuelle détendue de Driss créent une tension comique où le non-dit devient signifiant. Cette situation montre que l'humour ne se limite pas à la langue, mais intervient directement dans la construction des interactions sociales et des relations entre les individus. Comme l'a déjà souligné Henri Bergson (1900 : 178), le rire constitue un phénomène profondément social, souvent produit par le décalage entre attentes et comportements.

Cependant, malgré ces avancées, les recherches actuelles présentent plusieurs lacunes. D'une part, peu d'études relient systématiquement les concepts de la syntaxe générative aux interactions multimodales dans le cinéma (Mehler, 1969 : 10 ; Krysanova, 2023 : 45-63). D'autre part, si les dimensions pragmatiques de l'humour et de la stigmatisation ont été décrites (Lew, 1997 : 135 ; King, 2018 : 150), elles le sont rarement à partir d'un cadre opérationnel reposant sur une grille de catégorisation des structures syntaxiques, des gestes et des regards, permettant d'identifier leurs cooccurrences dans la production d'effets interactionnels. Enfin, les approches expérimentales telles que l'eye-tracking montrent que l'attention des spectateurs est guidée simultanément par des indices linguistiques et visuels (Thiele, Schmidt-Borcherding & Bateman, 2024 : 5), mais peu d'analyses appliquent cette perspective à l'étude des relations interpersonnelles représentées à l'écran.

Pour mieux cadrer l'analyse, il est essentiel de préciser les notions centrales mobilisées dans cette recherche : l'humour et la cognition des relations humaines.

## Définitions des concepts clés

Dans cette étude, l'humour est envisagé comme un phénomène interactionnel résultant d'un décalage entre attentes et interprétations possibles dans une situation communicative (Canestrari & Bianchi, 2012 : 557). Dans le contexte filmique, cet effet peut être produit par des ambiguïtés linguistiques, des transformations syntaxiques ou encore par la coordination entre parole, gestes et regards. L'humour est ainsi considéré non seulement comme un procédé discursif, mais aussi comme un mécanisme cognitif qui organise la relation entre les personnages.

La cognition des relations humaines renvoie à l'ensemble des processus par lesquels les individus interprètent, négocient et ajustent leurs interactions sociales. Dans le cadre du film *Intouchables*, cette cognition relationnelle se manifeste dans la manière dont les personnages construisent progressivement une relation de confiance et de solidarité à travers leurs échanges verbaux et non verbaux.

Ces deux notions servent de fil conducteur à l'analyse : l'étude cherche à montrer comment les structures syntaxiques et les ressources multimodales participent conjointement à la production de ces effets relationnels.

Dans cette optique, la présente étude vise à combler ce manque en proposant une approche intégrative où la grammaire générative devient un instrument d'analyse dynamique du sens en contexte filmique. En reliant les structures syntaxiques aux signaux visuels et kinésiques, elle cherche à montrer comment le film, en tant que texte multimodal, mobilise la cognition du spectateur de manière simultanée sur plusieurs plans : linguistique, perceptif et affectif. Ce croisement entre théorie linguistique et analyse sémiotique contribue à redéfinir les frontières de la psycholinguistique appliquée au cinéma, en soulignant le rôle du langage comme espace d'interaction cognitive, émotionnelle et esthétique.

## Démarche de la recherche : objectifs et hypothèses

### Objectifs de recherche

1. Analyser comment les structures syntaxiques profondes et de surface contribuent à la construction cognitive des relations interpersonnelles et à l'expression de l'humour, de l'intimité et de la solidarité dans *Intouchables*.
2. Examiner l'interaction entre parole, geste et regard et leurs fonctions pragmatiques dans la dynamique relationnelle entre Philippe et Driss.
3. Élaborer une catégorisation opérationnelle (ambiguïtés, transformations, fonctions interactionnelles) applicable et vérifiable dans d'autres contextes filmiques

## Questions de recherche et hypothèses

Afin de structurer l'analyse, cette recherche s'appuie sur deux questions principales :

1. comment les transformations syntaxiques dans les dialogues d'*Intouchables* contribuent-elles à la production d'effets humoristiques et relationnels ?
2. comment les gestes et les regards des personnages interagissent-ils avec ces structures linguistiques pour orienter l'interprétation des échanges par les interlocuteurs et par le spectateur ?

À partir de ces questions, l'étude formule l'hypothèse principale suivante : les effets humoristiques et relationnels observés dans le film ne reposent pas uniquement sur les structures linguistiques des énoncés, mais sur une coordination multimodale entre syntaxe, gestes et regards, qui guide l'interprétation cognitive des interactions.

## Méthodologie

Cette étude adopte une approche qualitative et analytique fondée sur l'articulation entre la grammaire générative transformationnelle et le modèle multimodal MDB-GSG (parole, geste, regard). Le corpus principal est constitué de séquences extraites du film *Intouchables*, sélectionnées en raison de leur forte valeur interactionnelle et de leur potentiel à illustrer les dynamiques relationnelles entre Philippe et Driss.

## Corpus et sélection des séquences

Le présent travail s'inscrit à l'intersection de la grammaire générative, de la multimodalité et de la psycholinguistique, en mobilisant notamment les travaux de (Chomsky, 1957 : 26-92 ; Chomsky, 1965 : 17 ; Chomsky, 1995 : 201), (Kress & Van Leeuwen, 2006 : 122), (Forceville & Urios-Aparisi, 2009 : 2849), (Bateman, 2014 : 120-125), (Tseng, Laubrock & Bateman, 2021 : 50) et (Attardo, 2001 : 157). Le corpus de l'étude est constitué de séquences dialogiques extraites du film *Intouchables* (Nakache & Toledano, 2011). Afin de garantir la cohérence analytique, l'étude s'appuie sur un sous-corpus de quinze scènes sélectionnées pour leur forte densité interactionnelle et leur potentiel à illustrer les dynamiques relationnelles entre Philippe et Driss.

La sélection des scènes repose sur trois critères principaux :

1. présence d'un échange dialogique direct entre les deux personnages principaux ;
2. occurrence d'un effet interactionnel identifiable (humour, tension, intimité ou solidarité) ;
3. présence simultanée d'indices verbaux et non verbaux (gestes, regards, expressions faciales ou prosodie).

Les scènes retenues incluent notamment : la scène du rasage, l'épisode du chocolat, l'échange dans la cuisine avec Yvonne, la scène du pichet d'eau chaude, ainsi que plusieurs interactions quotidiennes entre Driss et Philippe. Ces séquences ont été retenues car elles condensent des phénomènes linguistiques et multimodaux particulièrement saillants. La liste détaillée des séquences analysées, incluant minutage et justification, est présentée en fin d'article.

Chaque scène a été segmentée en unités interactionnelles courtes, correspondant généralement à un tour de parole ou à un événement multimodal saillant. Cette segmentation permet d'observer simultanément :

1. la structure linguistique de l'énoncé,
2. les gestes co-verbaux,
3. la direction du regard,
4. les indices prosodiques et expressifs.

La transcription suit un principe de synchronisation multimodale, inspiré des travaux de Quek et al. (2002 : 96) et de Tseng, Laubrock & Bateman (2021 : 19), où les éléments verbaux et non verbaux sont annotés dans la même unité analytique. Chaque segment comprend donc :

1. l'énoncé verbal produit par le personnage ;
2. la description du geste ou du mouvement corporel associé ;
3. la direction du regard (vers l'interlocuteur, vers l'objet, ou détourné) ;
4. le contexte interactionnel immédiat.

Cette méthode permet d'examiner les cooccurrences entre structures syntaxiques et indices multimodaux, condition essentielle pour analyser la production de l'humour et de l'intimité dans le film.

### Décodage, reproductibilité, articulation du modèle MDB-GSG

Dans un premier temps, les énoncés verbaux ont été décrits et classés selon les principes de la grammaire générative transformationnelle, afin d'identifier les opérations syntaxiques et transformations sémantiques à l'œuvre. Dans un deuxième temps, les ressources multimodales (regard, gestes, intonation) ont été codées et mises en relation avec les structures linguistiques correspondantes, en s'appuyant sur une grille de catégorisation opérationnelle. Enfin, les cooccurrences entre modalités verbales et non verbales ont été interprétées à la lumière des fonctions pragmatiques qu'elles remplissent dans la dynamique relationnelle des personnages. Cette méthodologie, ancrée dans la rigueur descriptive et l'interprétation contextuelle, vise à fournir un cadre transférable aux recherches psycholinguistiques et filmiques, en mettant en évidence l'apport conjoint de la syntaxe et de la multimodalité dans la cognition des relations humaines.

Le modèle MDB-GSG proposé par Quek et al. (2002 : 91-113) sert de cadre pour analyser l'articulation entre parole (*speech*), geste (*gesture*) et regard (*gaze*) dans l'interaction filmique. Dans cette perspective, les trois modalités ne sont pas considérées comme des éléments indépendants mais comme des ressources synchronisées participant à la production du sens. L'application du modèle consiste donc à examiner, pour chaque segment, la manière dont une transfor-

mation syntaxique donnée est accompagnée ou renforcée par un geste ou par une orientation du regard, produisant ainsi un effet interactionnel spécifique.

## Revue de la Littérature

### Grammaire générative : structures profondes et structures de surface dans l'analyse linguistique contemporaine

270

Depuis les travaux fondateurs de Chomsky (1957 : 19, 1965 : 15, 1995 : 201), le paradigme générativiste distingue clairement les structures profondes (*deep structures*, ou D-structure) et structures de surface (*surface structures*) pour rendre compte de la manière dont des transformations syntaxiques modifient la forme apparente des énoncés tout en maintenant, ou en altérant, leurs propriétés sémantiques (Van Gelderen, 2015 : 10). Ce bipartisme permet d'étudier comment des opérations comme la passivation, la mise en cause, ou encore l'inversion interrogative, se manifestent en surface mais tirent leur origine de relations syntaxiques sous-jacentes (Chomsky, 1995 : 205 ; Van Gelderen, 2015 : 11).

Dans les études filmiques, cependant, l'application de ces concepts reste parcellaire. Certains auteurs se concentrent sur des dialogues, variations de registre ou d'adresse (tu/vous) pour identifier des transformations syntaxiques implicites (King, 2018 : 154-156 ; Putri Wulandari, 2019 : 35). Mais peu relie systématiquement ces transformations à la structure profonde, telle que définie dans le générativisme, dans le contexte multimodal : la parole, le geste et le regard du film.

Le programme minimaliste, notamment, a atténué l'importance théorique des structures de surface, voire abandonné explicitement la terminologie D-structure/S-structure pour préférer des notions comme merge, échafaudage syntaxique, interfaces phonologiques et sémantiques (Chomsky, 1995 : 42-53 ; Van Gelderen, 2015 : 103). Malgré cela, plusieurs travaux contemporains en syntaxe expérimentale et linguistique cognitive réaffirment la valeur heuristique des contrastes entre ce qui est produit en surface et ce qui est opéré en profondeur particulièrement dans la compréhension, la production, et le traitement syntaxique implicite (par exemple dans la recherche sur l'acquisition, ou la variation dialectale).

### Ressources multimodales dans le langage filmique : parole, geste et regard

La littérature récente en psycholinguistique, en sémiotique cognitive et en études filmiques met en évidence que le verbal ne suffit pas à rendre compte des relations cognitives et interactionnelles visibles à l'écran (Kress & Van Leeuwen, 2006 : 70-75; Forceville & Urios-Aparisi, 2009 : 2848 ; Bateman, 2014 : 157 ; Amundrud et al., 2025 : 1-2 ; Chen et al., 2025 : 110). Le geste (*co-speech gestures*), le regard (*gaze*), l'expression faciale, ainsi que les indices prosodiques

sont aujourd'hui reconnus comme des ressources essentielles qui interagissent avec la syntaxe et la sémantique pour produire du sens. Par exemple, les études d'*eye-tracking* appliquées à des vidéos ou des scènes filmiques montrent comment l'attention visuelle est distribuée selon des indices visuels, auditifs et linguistiques (Tseng, Laubrock & Bateman, 2021 : 270 ; Thiele et al., 2024 : 5 ; Siegelman et al., 2025 : 11-13 ; Zou et al., 2025 : 3). Tseng, Laubrock & Bateman (2021 : 18) ont exploré *multimodal cohésion* dans les films, liant cooccurrences des indices visuels et auditifs à la compréhension narrative et aux transitions de scènes. Ces travaux montrent que les spectateurs reçoivent et traitent simultanément des signaux verbaux et non verbaux, qui interviennent notamment dans la construction de l'intention des personnages ou dans l'interprétation des actions implicites (Markhabayeva & Tseng, 2024 : 7). Des travaux expérimentaux récents confirment également que le traitement des gestes dépend de ressources cognitives spécifiques. Özer, Özyürek et Göksun (2025 : 1639-1653) montrent que la mémoire de travail spatiale joue un rôle déterminant dans l'intégration des gestes au discours verbal, en particulier lorsque le lien sémantique entre geste et parole varie. Ces résultats soutiennent l'idée que la multimodalité n'est pas simplement additive, mais cognitivement contraignante.

Par ailleurs, certaines études montrent comment des gestes pragmatiques et des marqueurs interactionnels accompagnent des tours de paroles complexes. Par exemple, Skogmyr Marian (2024 : 22) observe que le marqueur « tu vois » en français langue seconde s'accompagne souvent d'un regard mutuel et d'un geste orienté vers l'interlocuteur, ce qui renforce la fonction pragmatique de sollicitation ou de demande de *feed-back*. De même, dans les interactions face à face, la direction du regard et des gestes manuels modulent la transition entre locuteurs, accentuent certains actes de parole (directifs, expressifs) ou retardent des transitions lorsque les gestes sont incomplets ou le regard détourné (étude de la prise de parole multimodale). Dans une perspective développementale, Orakçı-Beyaztaş et Karadöller (2025 : 3-24) montrent que la manière dont les gestes sont présentés (perspective gestuelle) influence la performance spatiale des enfants. Bien que menée hors du contexte filmique, cette étude renforce l'idée que la configuration visuo-gestuelle affecte directement la cognition, ce qui demeure pertinent pour l'analyse des gestes filmiques.

### Humour, intimité, solidarité et stigmatisation : effets pragmatiques des ressources linguistiques et non linguistiques

Le film *Intouchables* est souvent analysé pour son usage de l'humour comme dispositif permettant de désamorcer la stigmatisation, en particulier vis-à-vis du handicap et de l'altérité (King, 2018 : 156). Le contraste initial entre Philippe et Driss par exemple les classes sociales, la culture, les expressions corporelles et le langage est graduellement estompé par l'interaction : dialogues, silences, gestes partagés, regards et moments non verbaux participent à la création d'une relation d'intimité et de solidarité.

Cependant, ces études tendent à se limiter à des niveaux descriptifs : elles cataloguent les stéréotypes, les moments comiques, les actes illocutoires, l'intonation (Hernanda & Firmonasari, 2015 : 52) mais rarement à combiner systématiquement les structures syntaxiques profondes, les transformations sémantiques, et les ressources multimodales en une analyse opérationnelle mesurable. Il y a, en particulier, peu de travaux qui transcrivent minutieusement à la fois la parole, le geste, le regard segment par segment, pour identifier les cooccurrences linguistiques avec les gestuelles et oculaires (regard) qui produisent l'humour, l'intimité, la solidarité ou la stigmatisation.

Des recherches liées, dans d'autres contextes filmiques ou conversationnels, montrent que l'intonation ou la prosodie servent à signaler l'ironie, la tension affective ou le conflit, et que ces signaux prosodiques interagissent avec le geste et le regard pour amplifier ou nuancer le message.

### Études empiriques liant syntaxe générative et multimodalité

Un nombre croissant d'études s'efforcent de croiser les perspectives syntaxiques formelles et l'analyse multimodale. Par exemple, des études d'eye-tracking dans le domaine de la langue seconde (L2) montrent que la charge cognitive liée à la compréhension syntaxique comme les constructions complexes et les transformations se manifeste dans des durées de fixation, retraits, etc. Ding & Chen (2022 : 174) examinent comment l'humeur des spectateurs module leur attention visuelle lors du visionnage de films, suggérant que les états affectifs influencent la distribution de l'attention mais aussi la sensibilité aux indices non verbaux (gestes, expressions) dans les scènes. Ces résultats montrent que les spectateurs ne réagissent pas seulement aux mots, mais aussi aux gestes, aux regards, et aux couplages multimodaux, ce qui est pertinent pour comprendre comment l'humour ou l'intimité sont perçus cinématographiquement.

La cohésion multimodale combinée à des méthodes expérimentales permet aussi de tester comment la synchronisation entre modalités verbales et non verbales affecte la compréhension narrative, l'interprétation des intentions des personnages et le sentiment de continuité ou de rupture dans le discours filmique (Tseng et al., 2021 : 51). Sur le plan méthodologique, l'essor de corpus oculométriques à grande échelle constitue un tournant majeur. La seconde vague du Multilingual Eye-Movement Corpus (MECO) proposée par Siegelman et al. (2025 : 11-13) fournit des données comparables sur le traitement visuel du langage à travers plusieurs langues, offrant un socle empirique robuste pour l'étude conjointe de la syntaxe, de l'attention visuelle et de la cognition multimodale.

### Justifications et directions pour notre cadre proposé

À partir de ce panorama, plusieurs lacunes se dégagent :

1. intégration pleine de la syntaxe générative et de la multimodalité : peu d'études appliquent les transformations syntaxiques (issues du générativisme) aux dialogues filmiques en les croisant avec le geste et le regard de

- manière segmentée ;
2. fonctions pragmatiques des ressources non verbales : alors que gestes, regard, intonation sont souvent décrits, leur rôle précis dans l'humour, dans l'intimité, dans la solidarité ou dans la stigmatisation est rarement codé, quantifié, ou mis en relation directe avec les structures syntaxiques profondes ;
  3. cadre opérationnel mesurable : les études expérimentales existent (*eye-tracking*, mesure d'attention, etc.), mais aucune, à notre connaissance, ne propose une catégorisation combinée des structures profondes et de surface, types d'ambiguïté, fonctions pragmatiques des gestes/regard qui soit falsifiable, transférable, et applicable dans l'analyse filmique, notamment pour *Intouchables*.

En mobilisant la grammaire générative transformationnelle (avec D-structure, S-structure ou leurs équivalents modernisés) et un modèle multi-modal parole-geste-regard, segmenté, mesurable, nous plaçons l'étude à l'intersection de plusieurs tendances actuelles : syntaxe expérimentale, psycholinguistique, études filmiques et multimodalité.

## Structures syntaxiques et cognition : le cas Driss

Les analyses qui suivent s'appuient sur plusieurs scènes<sup>1</sup> sélectionnées du film, chacune présentant une interaction significative entre Philippe et Driss ainsi qu'avec Yvonne. L'objectif n'est pas de couvrir l'ensemble du film, mais d'examiner des moments où les ressources linguistiques et multimodales se combinent de manière particulièrement saillante pour produire des effets relationnels tels que l'humour, l'intimité ou la solidarité.

Suivant les bases psycholinguistiques établies, la linguistique appliquée aux énonciations des personnages du film *Intouchables* nous aide à comprendre les implications cognitives et interactionnelles de leurs choix syntaxiques. L'un des exemples les plus significatifs se trouve dans le récit de Driss sur son histoire familiale :

1. c'est mon oncle et ma tante. Ils m'ont adopté à 8 ans ;
2. je m'appelle Bakari. C'est mon vrai nom ;
3. mais, ils m'ont appelé Idriss. C'est Driss qui est resté (1:27:38 - 1:27:58).

Chaque phrase présente une structure syntaxique particulière, articulée entre structure profonde et structure de surface. Par transformation grammaticale, elles pourraient être reformulées ainsi :

1. mon oncle et ma tante m'ont adopté à 8 ans ;
2. mon vrai nom est Bakari ;
3. ils m'ont enfin appelé Driss.

Le choix de Driss d'exprimer ces énoncés sous forme de structures plus complètes plutôt que d'opter pour des formulations condensées met en évidence

<sup>1</sup> Bateman (2014 : 1-292) développe une approche structurée de l'analyse multimodale des objets filmiques. Toutefois, ce cadre n'intègre pas explicitement les modèles générativistes de la syntaxe.

son état cognitif et affectif. L'analyse en termes de structure profonde permet d'interpréter l'intensité émotionnelle du moment où il évoque son passé. Comme l'indique Chomsky (1995 : 45), la structure profonde est le lieu où se joue l'interprétation sémantique essentielle ; ici, elle permet d'éviter toute ambiguïté dans un récit marqué par l'émotion.

Cette préférence peut être interprétée comme un trait caractéristique du personnage de Driss, tel qu'il est construit dans le scénario. Les dialogues du film sont en effet scénarisés et ne constituent pas des productions linguistiques spontanées : l'analyse ne porte donc pas sur la compétence linguistique réelle d'un locuteur, mais sur la manière dont le film met en scène certains styles discursifs.

Par exemple, dans une scène (43:02 – 45:00), Philippe demande à Driss s'il possède des compétences artistiques ou littéraires. Driss répond brièvement par la négative, ce qui peut être interprété comme un rejet implicite des formes d'expression perçues comme trop intellectuelles. Cette réponse met en évidence l'écart entre les univers culturels des deux personnages et éclaire la simplicité volontaire de ses choix linguistiques.

Cet écart apparaît dès l'entretien initial, lorsque Driss détourne la question des références en citant des artistes populaires « Kool and the Gang, Earth, Wind and Fire », révélant une reconfiguration pragmatique de la notion même de compétence culturelle (10:27 – 10:36). La compétence linguistique de Driss n'est pas académique, mais pragmatique : elle assure la transmission claire de son vécu et renforce l'impact affectif de son discours.

Cette dynamique se prolonge enfin dans la scène de découverte de sa chambre, où le regard de Driss se fixe sur les objets luxueux, notamment la baignoire, traduisant une confrontation perceptive entre habitus social et nouvel environnement (20:57 – 21:11).

### La structure de l'humour et la mentalité

L'analyse cognitive et syntaxique de Driss peut être prolongée vers la compréhension de l'humour dans *Intouchables*. L'humour, loin d'être un simple ornement discursif, apparaît comme un mode de communication structuré. Il suppose l'interaction de plusieurs participants et repose sur des attentes partagées.

Un exemple marquant est la scène où Driss, manipulant un pichet d'eau chaude, feint de masser Philippe :

Driss : *Putain, pardon* (il renverse de l'eau chaude sur Philippe).

Philippe : *Qu'est-ce qui se passe ?* (il ne sent rien).

Driss : *Rien.*

Philippe : *Sûr ?*

Driss : *Ouais. Je masse là, je continue* (27:48 - 28:32).

Dans cette scène, l'effet humoristique ne repose pas uniquement sur le contenu verbal, mais sur la coordination entre parole, geste et regard. Driss ren-

verse accidentellement de l'eau chaude sur Philippe tout en prononçant « Putain, pardon », ce qui introduit déjà un décalage entre la gravité potentielle de l'action et la relative légèreté de la réaction verbale. Cependant, la situation devient véritablement comique lorsque Philippe demande : « Qu'est-ce qui se passe ? », révélant qu'il ne ressent aucune douleur en raison de sa condition physique.

Le geste de Driss — qui interrompt brièvement son mouvement, puis reprend le massage comme si rien ne s'était produit — renforce l'ambiguïté de la situation. Son regard furtif vers l'infirmière qui entre dans la pièce accentue encore cet effet de connivence. La scène met ainsi en place une véritable synchronisation multimodale : la question sérieuse de Philippe, la réponse minimaliste de Driss (« Rien ») et la continuité du geste produisent un contraste générateur de rire.

Dans cette configuration, la transformation interactionnelle ne relève pas seulement de la syntaxe, mais d'une véritable orchestration multimodale. L'humour émerge précisément de la dissociation entre l'événement physique (l'eau renversée), l'énoncé verbal « Rien » et le comportement gestuel de Driss, qui poursuit le massage comme si la situation était parfaitement normale.

Cette scène illustre une logique interactionnelle fondée sur un savoir partagé : Philippe sait que Driss est joueur, et le spectateur accède progressivement à ce savoir implicite. L'humour apparaît alors comme une forme de communication codée, enracinée dans une complicité cognitive entre les deux personnages, complicité à laquelle le spectateur est progressivement associé.

Des mécanismes similaires apparaissent dans d'autres scènes. Lorsqu'il interrompt un couple dans le jardin, Driss transforme une situation intime en interaction comique par intrusion verbale et renversement pragmatique « Essaie dans tes veuchs » (27:05 – 27:17). L'effet humoristique repose également sur une désynchronisation attentionnelle lorsqu'il détourne son regard vers Magalie, au point de perturber l'acte alimentaire, révélant l'interaction entre perception visuelle et action motrice (29:14 – 29:27). Cette désynchronisation atteint un point critique lorsqu'il tend machinalement le téléphone à Philippe, illustrant une rupture momentanée dans le traitement cognitif de la situation interactionnelle (29:30 – 29:39).

La scène du parking met en évidence une validation implicite d'une norme alternative : l'approbation de Philippe « Bonne méthode » transforme un acte agressif en stratégie interactionnelle légitimée (31:45 – 32:20). Cette logique se prolonge dans une série de transformations visuelles (figures caricaturales), où le geste tend progressivement à se substituer à la parole dans la production de l'humour (1:43:15 – 1:45:12).

La transition musicale lors de l'anniversaire illustre, quant à elle, une reconfiguration collective de l'interaction : le passage à une musique dansante redéfinit la dynamique relationnelle et favorise une synchronisation affective entre les participants (1:08:56 – 1:12:55).

## Ambiguïté et transformations syntaxiques

L'ambiguïté est l'un des ressorts majeurs de l'humour dans le film. Prenons l'échange entre Driss et Yvonne autour du concombre :

Driss : *Et vous ? Pas d'amoureux ? Même pas le jardinier ?*

Yvonne : *Pas du tout.*

Driss : *Il vous regarde souvent. J'ai raison ou pas ?*

Yvonne : *Arrêtez maintenant, c'est ridicule.*

Driss : *Il a mis ou pas ?*

Yvonne : *Quoi ?*

Driss : *Son concombre (Yvonne le regarde avec étonnement, semblant hésiter sur le sens de l'énoncé).*

Driss : *Quoi ? (Riant de se moquer d'elle).*

Yvonne : *Je rêve ?*

Driss : *Albert! Albert! (Il part). Bientôt.*

Yvonne : *Oh ! (38:42 - 39:15).*

L'explication psychologique du comportement de Driss et d'Yvonne ne réside pas seulement dans leur cognition, mais aussi dans le lieu où se déroule l'échange, à savoir la cuisine. Cela éclaire le jeu de mots de Driss lorsqu'il compare la virilité d'Albert à un concombre. Ce n'est donc pas uniquement le cadre de l'interaction qui suscite un effet comique, mais également la manière dont la cognition de l'acteur en nourrit le sens. Chez Driss, ce type de comportement n'est pas nouveau : il a l'habitude de s'exprimer de façon provocatrice avec ses collègues. Ce type d'humour peut être interprété comme provocateur dans le contexte interactionnel, bien que Driss ne manifeste aucun signe d'embarras. On peut ainsi souligner la capacité d'un plaisantin à produire une blague à travers sa cognition et sa personnalité, comme l'illustre le cas de Driss que nous analysons ici en tant qu'un personnage caractérisé par un style interactionnel direct et provocateur. L'ambiguïté se retrouve également dans l'évocation du soulagement, où Philippe mobilise un geste plutôt qu'un énoncé explicite, déplaçant l'interprétation vers un implicite partagé que Driss reformule humoristiquement (42:45 – 42:57). Son développement cognitif le conduit à mobiliser un humour reposant sur des sous-entendus sexuels et des implicites pragmatiques, qui déplace l'échange vers un registre autre :

Driss : *Magalie, tu as deux minutes ?*

Magalie : *Non.*

Driss : *Je vais te montrer un truc.*

Magalie : *Quoi ?*

Driss : *Flippe pas, viens voir.*

Magalie : *Bon, une minute. Et donc ?*

Driss : *J'ai une baignoire, c'est tout.*

Magalie : *C'est intéressant ! Et alors ?*

Driss : *Alors, on peut prendre un bain. Y a de la place, y a du sel, de la mousse.*

Magalie : *(souriant d'une manière trompeuse). Ok, après tout pourquoi pas ? Vas-y, commence à te déshabiller.*

Driss : *T'es comme ça ? Un peu coquin ... J'aime bien. Et juste se rincer ! (Elle part sans bruit lorsque Driss déshabille joyeusement) (36:18 -3714).*

Le précédent exemple illustre un développement structuré autour d'un contexte humoristique dans lequel l'interlocuteur est amené à formuler des hypothèses sur les intentions du locuteur. Toutefois, ce dernier encode son sens réel à travers un jeu de mots ou une expression indirecte, générant ainsi un décalage interprétatif. Ce décalage produit un effet de surprise qui constitue le ressort comique de l'échange. Le contexte d'une blague ou d'une émission humoristique repose la plupart du temps sur un jeu de mots conscient ou une expression indirecte. Cela peut amener l'auditeur à un moment de dilemme quant au choix d'un sens spécifique parmi plusieurs interprétations possibles générées par la communication, le discours ou le contexte humoristique. C'est ce que l'on appelle l'ambiguïté, en raison de la complexité de sens qu'une phrase ou un processus communicationnel peut comporter. Ainsi, avant d'aborder la compréhension de l'humour, il convient d'analyser l'ambiguïté à partir d'une approche linguistique, ici la théorie générative transformationnelle de Noam Chomsky, qui distingue la structure de surface et la structure profonde.

### La structure profonde et la structure de surface

L'explication transformationnelle<sup>2</sup> montre que la structure profonde constitue un niveau abstrait d'organisation dans lequel tous les éléments nécessaires à l'interprétation sont représentés. Elle détermine le sens primaire et les catégories de phrase. Dans les contextes humoristiques, Driss produit des énoncés ambigus en transformant la structure profonde en un arrangement linéaire de mots et de phrases articulées, ce que l'on appelle la structure de surface (38:52). Par exemple : « Il l'a mis ou pas ? Son concombre ! » constitue une représentation phonologique issue de l'application des règles transformationnelles. Pour comprendre cette explication, il faut noter que la structure profonde renvoie à une interprétation sémantique où se trouve toute l'information nécessaire, tandis que la structure de surface est une phrase dérivée qui sert de représentation phonologique. Autrement dit, l'application des règles transformationnelles (ajouter, supprimer, remplacer ou modifier l'ordre des constituants de la struc-

2 Chomsky (1957 : 28-29, 1965 : 15, 1995 : 153) fonde le cadre de la grammaire générative transformationnelle mobilisé ici pour analyser les structures profondes et de surface. Toutefois, ce modèle, initialement conçu pour des données décontextualisées, ne prend pas en compte les dimensions interactionnelles et multimodales.

ture profonde) produit une phrase simple, utilisée au quotidien, appelée structure de surface.

Cependant, comme on l'a observé, la structure de surface ne désigne pas un sens unique. Le résultat peut être une ambiguïté structurelle, organisée autour d'une grammaire qui ouvre la possibilité de plusieurs interprétations. Prenons l'exemple de : « Il l'a mis ou pas ? Son concombre », et tentons d'en reconstruire la structure profonde :

Albert a mis son concombre → S + V + O / PN + PV + PN.

Ici, S désigne le sujet, V le verbe, PN la phrase nominale et PV la phrase verbale.

Dans cette scène, l'ambiguïté syntaxique fonctionne comme un déclencheur de l'effet humoristique. La structure interrogative « Il l'a mis ou pas ? » introduit volontairement une indétermination référentielle, qui n'est résolue qu'au moment où Driss ajoute « son concombre ». Ce déplacement progressif du sens crée un effet de surprise. L'interprétation comique est renforcée par les indices multimodaux : le regard perplexe d'Yvonne, le sourire de Driss et son ton ironique orientent la compréhension vers une lecture sexuelle implicite.

Il ne fait pas de doute que le pronom « il » réfère à Albert, car ce nom est déjà évoqué dans la discussion de la scène. La question d'Yvonne n'est donc pas « qui », mais bien « quoi ». Elle connaît déjà le sujet.

La première phrase représente la structure de base (*kernel structure*) à partir de laquelle dérivent d'autres séries par transformation. En grammaire transformationnelle, la structure de base désigne généralement une phrase simple, affirmative, active et sans négation. À partir de cette phrase de base, plusieurs transformations<sup>3</sup> sont possibles :

- i. Tn (transformation de négation) ;
- ii. Tq (transformation interrogative) ;
- iii. Tb (transformation de but).

### La transformation de négation

Considérons l'application de Tn aux séries sous-jacentes :

- ia. Albert n'a pas mis son concombre;
- ib. Albert n'a jamais mis son concombre

Ces transformations illustrent la manière dont une même structure de base peut produire plusieurs interprétations possibles. Dans le contexte humoristique de la scène, cette variabilité syntaxique contribue à créer une ambiguïté interprétative, que les indices gestuels et les réactions des personnages permettent ensuite de résoudre.

3 Tseng et al. (2021 : 50) analysent la synchronisation entre parole, geste et regard dans l'interaction. Leur modèle est opératoire pour l'annotation multimodale, bien qu'il ne traite pas directement des transformations syntaxiques.

### La transformation interrogative

À partir de la négation, appliquons la transformation interrogative :

iiia. Albert, a-t-il mis son concombre ?

iiib. Qu'est-ce qu'Albert a mis ?

iiic. Qui a mis son concombre ?

Intuitivement, (iiia) est une interrogative totale qui appelle une réponse par « oui » ou « non », tandis que (iiib) et (iiic) sont des interrogatives partielles. Le Tn peut aussi s'appliquer au Tq, donnant des formes interrogatives négatives :

iiid. Albert n'a pas mis son concombre ?

iiie. Albert n'a jamais mis son concombre ?

iiif. Albert n'a que mis son concombre ?

Toutes ces phrases dérivent de la phrase de base « Albert a mis son concombre », vers l'application successive de transformations. Bien que les phrases conservent un ordre syntaxique comparable, elles produisent des sens différents, ce qui laisse à l'interlocuteur la responsabilité de choisir l'interprétation la plus pertinente.

### La transformation de but

La transformation de but (Tb) permet d'introduire des finalités. Parmi les dérivés possibles :

iiia. Albert a mis son concombre dans sa poche (pour le garder).

iiib. Albert a mis son concombre dans sa bouche (pour le manger).

iiic. Albert a mis son concombre dans Yvonne (pour « la manger »).

La dernière interprétation (iiic) apparaît la plus proche du sens visé, car elle est appuyée par d'autres indices discursifs (« pas d'amour ? », « il vous regarde souvent ») et par le geste moqueur de Driss. Ici, le concombre est un euphémisme pour désigner l'organe génital, ce qui inscrit l'énoncé dans un registre grivois.

Relevons brièvement l'analyse d'une deuxième phrase, tirée d'une autre scène, où Driss introduit l'ambiguïté en se moquant de Philippe, son patron :

Philippe : *Allez, donnez-moi un chocolat (il ouvre la bouche).*

Driss : *Non.*

Philippe : *Donnez-moi chocolat (Driss fait un geste pour que Philippe se fâche contre lui).*

Driss : *Pas de bras, pas de chocolat. (Philippe semble confus. Driss se rend compte de ce qui se passe). C'est une vanne oh, je déconne.*

Philippe : *Ah c'est une blague ?*

Driss : *Elle est bien !*

Philippe : *C'est une vanne, très bonne blague !*

Driss : *Elle est chan-mé! Elle est très bonne. Vous n'avez pas de bras (Il rit bruyamment en se moquant de lui). Dommage qu'il n'y ait pas de public (33 : 23 - 34 : 00).*

Sans doute, Philippe est troublé par l'ambiguïté contenue dans l'énoncé « pas de bras, pas de chocolat ». L'analyse transformationnelle permet de préciser les différentes interprétations possibles.

La phrase de base est : « *Philippe a des bras* ». À partir de cette structure, plusieurs transformations sont envisageables. La transformation de négation peut donner : « *Philippe n'a pas du tout de bras* » ou encore « *Les bras de Philippe ne fonctionnent pas* ». Dans ce dernier cas, il ne s'agit pas d'une amputation mais d'une paralysie.

Cette négation peut ensuite s'étendre à la transformation interrogative, produisant des formes telles que : « *Qui n'a pas de bras ?* », « *Philippe, a-t-il des bras ?* » ou encore « *Les bras de Philippe fonctionnent-ils ?* ». La transformation de but ouvre également la possibilité suivante : « *Philippe a des bras pour prendre le chocolat* ».

L'ambiguïté se manifeste principalement au niveau des transformations négatives et interrogatives. Voilà pourquoi Philippe se trouve confus : en réalité, il possède bien des bras, mais une condition médicale l'empêche de les utiliser pleinement. Finalement, cette situation déclenche le rire, car Philippe ne comprend qu'après coup qu'il s'agissait d'une blague. Driss, quant à lui, s'amuse de la situation et souligne que c'est dommage que d'autres n'aient pas assisté à ce jeu d'ambiguïté dans lequel Philippe est piégé.

### Production et compréhension de l'humour

Ce passage expliqué ci-dessus illustre la fonction sociale de l'humour : désamorcer la tension et rééquilibrer les rôles interactionnels (King, 2018 : 147). Comme le rappelle Chomsky (2006 : 27-30), la transformation syntaxique ne se limite pas à la structure de surface ; elle engage aussi des effets pragmatiques et contextuels. Ici, la force de l'humour réside dans l'écart entre la réalité profonde (Philippe n'est pas amputé) et la surface de l'énoncé (qui suggère l'inverse).

L'humour de Driss peut être qualifié comme rude et vulgaire, mais il prend racine dans sa cognition. Les représentations morphologiques et phonologiques activées dans son esprit sont traduites en énoncés par ses organes vocaux. Cette dynamique illustre que l'humour est d'abord une production cognitive avant d'être une performance sociale.

La compréhension de l'humour, en revanche, repose sur un travail interprétatif. L'interlocuteur doit décoder une structure de surface ambiguë en choisissant le sens le plus pertinent parmi les options transformationnelles (Canestrari & Bianchi, 2012 : 548-560). Ce processus confirme l'idée que la compétence linguistique (au sens chomskyen) est centrale dans l'accès au comique.

## Grille de catégorisation multimodale (syntaxique, gestuelle et oculaire)

Afin de rendre l'analyse reproductible, les données ont été codées à l'aide d'une grille de catégorisation combinant trois dimensions principales : la structure syntaxique, les ressources multimodales (geste et regard) et les fonctions interactionnelles. Cette grille permet d'identifier les cooccurrences entre transformations syntaxiques et indices non verbaux dans la production d'effets humoristiques ou relationnels.

**Tableau 1 - Cadre de codage multimodal des données**

Catégorie	Sous-catégorie	Description analytique	Exemple dans le corpus
Structure syntaxique	Structure de base	Phrase simple affirmative servant de structure de base transformationnelle	« Albert a mis son concombre »
Transformation syntaxique	Interrogation / négation / déplacement	Modification de la structure de base produisant un effet d'ambiguïté	« Il l'a mis ou pas ? »
Geste	Geste iconique / déictique / expressif	Mouvement manuel accompagnant la parole	geste moqueur de Driss
Regard	Regard direct / détourné / partagé	Orientation visuelle vers l'interlocuteur ou ailleurs	regard perplexe d'Yvonne
Fonction interactionnelle	Humour / intimité / solidarité / tension	Effet pragmatique produit par la combinaison multimodale	rire partagé

**Tableau 2 - Séquences analysées**

No	Scène	Minutage	Justification
1.	Pichet d'eau chaude	27:48 – 28:32	Séquence à forte densité multimodale où le décalage entre sensation physique et réponse verbale produit un effet humoristique fondé sur la dissociation cognition/action.
2.	Cuisine – concombre (Yvonne)	38:42 – 39:15	Ambiguïté syntaxique et double interprétation sémantique renforcées par le regard, générant un humour à connotation sexuelle.
3.	Bain – Magalie	36:18 – 37:14	Construction d'un implicite pragmatique et d'un malentendu intentionnel reposant sur la coordination parole/geste.
4.	Récit personnel de Driss	1:27:38 – 1:27:58	Séquence marquée par une structuration syntaxique explicite révélant une forte charge émotionnelle et cognitive.

5.	Philippe interroge Driss (compétences)	43:02 – 45:00	Contraste sociolinguistique entre registres discursifs, révélant une asymétrie culturelle et cognitive.
6.	« Vous n’avez pas de bras » (chocolat)	33:23 – 34:00	Exemple d’humour direct fondé sur la transgression des normes sociales et la désacralisation du handicap.
7.	Références musicales (entretien)	10:27 – 10:36	Jeu sur la notion de « références » produisant un décalage sémantique entre attente institutionnelle et interprétation personnelle.
8.	Découverte de la chambre luxueuse	20:57 – 21:11	Mise en scène visuelle de la surprise cognitive face au luxe, avec focalisation sur l’objet (baignoire) et accompagnement musical.
9.	Jardin – interaction avec le couple	27:05 – 27:17	Intrusion interactionnelle produisant un renversement de situation et un humour fondé sur la rupture de script social.
10.	Regard sur Mégalie (distraction)	29:14 – 29:27	Désalignement entre attention visuelle et action motrice, révélant l’impact du regard sur l’interaction.
11.	Téléphone donné à Philippe	29:30 – 29:39	Inadéquation cognitive momentanée illustrant une automatisation du geste non adaptée au contexte du handicap.
12.	Conflit du parking	31:45 – 32:20	Affirmation d’une norme sociale par l’action physique, validée implicitement par Philippe, révélant une solidarité émergente.
13.	« Rien ne peut vous soulager ? »	42:45 – 42:57	Humour fondé sur une interprétation détournée du terme « soulager », appuyée par le regard et le geste.
14.	Scène du rasage	1:43:15 – 1:45:12	Transformation identitaire ludique reposant sur le geste et la référence culturelle, renforçant l’intimité.
15.	Anniversaire – transition musicale	1:08:56 – 1:12:55	Rupture de registre culturel (opéra à danse) illustrant une reconfiguration relationnelle par la multimodalité sonore et corporelle.

## Conclusion

L’analyse linguistique ne saurait, à elle seule, rendre compte de l’effet humoristique : gestes, regards et intonations y contribuent de manière décisive. Dans la scène du concombre, l’expression perplexe d’Yvonne et l’intonation ironique de Driss orientent l’interprétation vers une lecture humoristique, illustrant le rôle structurant des indices multimodaux.

Plus généralement, la compréhension des énoncés dans *Intouchables* repose sur la coordination de ressources verbales, visuelles et prosodiques. Le

modèle MDB-GSG (Quek et al., 2002 : 91-113) permet d'en rendre compte en concevant parole, geste et regard comme des systèmes interdépendants. Les travaux récents (Tseng, Laubrock & Bateman, 2023 : 261-285 ; Skogmyr Marian, 2024 : 28) confirment le rôle central de la synchronisation multimodale dans la construction du sens filmique.

Dans cette perspective, l'humour apparaît comme un mécanisme cognitif et interactionnel majeur. Ancré dans l'ambiguïté syntaxique, structuré par des configurations profondes et renforcé par des indices multimodaux, il participe à la construction de l'intimité et de la solidarité entre les personnages. L'articulation entre grammaire générative transformationnelle et modèle MDB-GSG montre ainsi que les structures linguistiques constituent des outils opératoires pour analyser l'engagement social à l'écran.

L'apport principal de cette recherche réside dans l'intégration d'une approche générativiste et d'un cadre multimodal, ouvrant des perspectives pour les études filmiques et psycholinguistiques. Des prolongements empiriques notamment par l'*eye-tracking* ou l'analyse des réactions spectatoriennes permettraient de tester et d'affiner ce modèle dans divers contextes culturels et génériques.

Enfin, certaines limites doivent être soulignées : le corpus restreint, la prédominance d'une approche qualitative et la focalisation sur l'humour au détriment d'autres dynamiques interactionnelles (conflit, autorité, silence). Ces éléments constituent autant de pistes pour de futures recherches.

## Bibliographie

- AMUNDRUD, T., AZAUIO, B., CREMONA, G., *et alii*, « Multimodality in face-to-face teaching and learning: Contemporary re-evaluations in theory, method, and pedagogy », *Frontiers in Communication*, 10, 2025 : <https://doi.org/10.3389/fcomm.2025.1656681>.
- ALTMAN, G. T. M., « Thematic role assignment in context », *Journal of Memory and Language*, 45, 1, 2001 : p. 68-92.
- ATTARDO, Salvatore, *Humorous Texts: A Semantic and Pragmatic Analysis*, Berlin, Mouton de Gruyter, 2001.
- BATEMAN, John A., *Text and image: A critical introduction to the visual/verbal divide*, London, Routledge, 2014.
- BATEMAN, John A., TSENG, Chiao-I., « Multimodal discourse analysis as a method for revealing narrative strategies in news videos », *Multimodal Communication*, 12, 3, 2023 : p. 261-285.
- BERGSON, Henri, *Le rire : essai sur la signification du comique*, Paris, Alcan, 1900.
- CANESTRARI, Carla, BIANCHI, Ivana, « Perception of contrariety in jokes », *Discourse Processes*, 49, 7, 2012 : p. 539-564.
- CHARE, Nicholas, WATKINS, Liz, dir., *Gesture and film: Signalling new critical perspectives*, London, Routledge, 2017.
- CHEN, L., ABDULLAH, M. A. R., *et alii*, « A bibliometric survey on multimodal discourse analysis (2015-2024): Looking behind to look ahead », *English Language and Literature Studies*, 15, 3, 2025 : p. 99-114.

- CHOMSKY, Noam, *Syntactic Structures*, The Hague, Mouton, 1957.
- CHOMSKY, Noam, *Aspects of the Theory of Syntax*, Cambridge, MIT Press, 1965.
- CHOMSKY, Noam, *The Minimalist Program*, Cambridge, MIT Press, 1995.
- CHOMSKY, Noam, *Language and Mind*, Cambridge, Cambridge University Press, 2006.
- DING, Ning, CHEN, Yi, « Mood modulates attention in film viewing: An eye-tracking study », *International Communication of Chinese Culture*, 9, 4, 2022 : p. 169–180.
- DYNEL, Marta, *Humorous Garden-Paths: A Pragmatic-Cognitive Study*, Newcastle, Cambridge Scholars Publishing, 2009.
- FERNANDEZ, Eva M., CAIRNS, Helen Smith, *Fundamentals of Psycholinguistics*, Oxford, Wiley-Blackwell, 2010.
- FORCEVILLE, Charles, URIOS-APARISI, Eduardo, dir., *Multimodal Metaphor*, Berlin, Mouton de Gruyter, 2009.
- HARLEY, Trevor, *The Psychology of Language: From Data to theory*, Hove, Psychology Press, 2005.
- HARRIS, Richard J., DUBITSKY, Tony M., BRUNO, John, « Psycholinguistic studies of misleading advertising », *Journal of Consumer Psychology*, 22, 2, 2012 : p. 241–262.
- HERNANDA, K. G. L., FIRMONASARI, A., *Intonasi dialog dalam film Intouchables*, Universitas Gadjah Mada, 2015.
- KING, Andrew, « Humour, disability and the film Intouchables », *Journal of Language and Popular Culture*, 21, 2, 2018 : p. 145–162.
- KRESS, Gunther, VAN LEEUWEN, Theo, *Reading images: The grammar of visual design*, London, Routledge, 2006.
- KRYSANOVA, Tetiana, « Psycholinguistic and cognitive-semiotic dimensions of constructing fear in horror films: A multimodal perspective », *East European Journal of Psycholinguistics*, 10, 1, 2023 : p. 45–63.
- LEW, Robert, « Towards a taxonomy of linguistic jokes », *Studia Anglica Posnaniensia*, 31, 1997 : p. 123–152.
- LEVELT, Willem J. M., *A History of Psycholinguistics: The Pre-Chomskyan Era*, Oxford, Oxford University Press, 2012.
- MAHLBERG, Michaela, CONKLIN, Kathy, BISSON, Marie J., « Reading Dickens's characters: Employing psycholinguistic methods to investigate the cognitive reality of patterns in texts », *Language and Literature*, 23, 4, 2014 : p. 369–388.
- MARKHABAYEVA, D., TSENG, Chiao-I., « Multimodal cohesion and viewers' comprehension of scene transitions in film: An empirical investigation », *Frontiers in Communication*, 9, 2024 : <https://doi.org/10.3389/fcomm.2024.1347788>.
- MEHLER, Jacques, *Psycholinguistique et grammaire générative*, Paris, Larousse, 1969.
- MONDADA, Lorenza, « Multimodal gestalts in interaction: Demandes commerciales et langues parlées », *Languages*, 9, 1, 2024 : <https://doi.org/10.3390/languages9010001>.
- ORAKÇI-BEYAZTAŞ, E., KARADÖLLER, D. Z., « Exploring the relation between gesture presentation perspective and children's spatial performance », *Gesture*, 2025 : <https://doi.org/10.1075/gest.25016.ora>.
- ÖZER, D., ÖZYÜREK, A., GÖKSUN, T., « Spatial working memory is critical for gesture processing: Evidence from gestures with varying semantic links to speech », *Psychonomic Bulletin & Review*, 32, 2025 : p. 1639–1653.
- PUTRI WULANDARI, « Analisis tindak tutur ilokusi dalam film Intouchables karya Olivier Nakache dan Éric Toledano », Universitas Pendidikan Indonesia, 2019 : <https://repository.upi.edu/45127/>.

- QUEK, Francis, MCNEILL, David, BRYLL, Robert, *et alii*, « A multimedia system for temporally situated perceptual psycholinguistic analysis », *Kluwer Academic Publishers*, 18, 2, 2002 : p. 91–113.
- RASKIN, Victor, *Semantic Mechanisms of Humor*, Dordrecht, Reidel, 1985.
- SIEGELMAN, N., SCHROEDER, S., BAO, Y. B., *et alii*, « Wave 2 of the Multilingual Eye-Movement Corpus (MECO): New text reading data across languages », *Scientific Data*, 12, 2025 : <https://doi.org/10.1038/s41597-025-05453-3>.
- SKOGMYR MARIAN, Klara, « Longitudinal change in linguistic resources for interaction: The case of tu vois », *Interactional Linguistics*, 4, 2024 : p. 3–37.
- THIELE, L., SCHMIDT-BORCHERDING, F., BATEMAN, John A., « All eyes on the signal? Mapping cohesive discourse structures with eye-tracking data », *Frontiers in Communication*, 9, 2024 : <https://doi.org/10.3389/fcomm.2024.1356495>.
- TSENG, Chiao-I., LAUBROCK, Joerg, BATEMAN, John A., « The impact of multimodal cohesion on attention and interpretation in film », *Discourse, Context & Media*, 44, 2021 : <https://doi.org/10.1016/j.dcm.2021.100544>.
- VAN GELDEREN, Elly, *An Introduction to Minimalism*, Amsterdam, John Benjamins, 2015.
- YULE, George, *The Study of Language*, Cambridge, Cambridge University Press, 2016.
- ZASIEKINA, Larisa, « Psycholinguistic approach to diagnostics and modification of individual negative core beliefs », *East European Journal of Psycholinguistics*, 2, 2, 2015 : p. 151–157.
- ZOU, Q., ZHENG, W., SU, Z., *et alii*, « The neural impact of editing on viewer narrative cognition in virtual reality films », *Frontiers in Psychology*, 16, 2025 : <https://doi.org/10.3389/fpsyg.2025.1584250>.

**Film cité :**

NAKACHE, Olivier, TOLEDANO, Éric, *Intouchables*, film, Quad Productions/Gaumont, 2011.

