

MÉMOIRE DE RECHERCHE

*Master 1 Sciences sociales - Parcours quantifier en sciences sociales
2025/2026*

Avoir voix au chapitre

**Pratiques langagières et positions sociales à
l'Assemblée Nationale**

Soutenu par

LYLIA BOUGUEDOURA

Encadré par

M. ETIENNE PENISSAT

Table des matières

Introduction.....	2
I. Quantifier la parole.....	8
A. Que valent les comptes-rendus de séance ?.....	8
1. Des conditions de production des comptes-rendus.....	9
2. Oralité et corporéité.....	10
3. Précision des verbatim et imprécision des locuteurs.....	11
4. Locuteurs ou interlocuteurs ?.....	12
B. Faire parler la parole.....	14
1. Construction du corpus.....	14
2. Du (non) prétraitement.....	16
3. De la construction au traitement : les indicateurs.....	18
II. Situer socialement la parole.....	20
A. Se classer pour se présenter : le Répertoire national des élus.....	20
1. Classements négociés dans une nomenclature imposée.....	22
2. Mise en récit et trajectoire biographique.....	23
a) Parler de soi, parler des autres ?.....	24
b) La députation : un milieu social qui s'ignore ?.....	26
B. Le RNE : potentiel et prolongements.....	27
1. Les variables sociodémographiques, substrat de l'analyse.....	27
2. Prolongements et enrichissements.....	31
III. Parler et dire : une ébauche de la domination par les pratiques langagières.....	33
A. Parler : Distributions de la parole et interactions.....	33
1. Une distribution inégale de la parole.....	34
2. La parole et l'auditoire : configurations d'interactions.....	36
B. Dire : des pratiques langagières et de leur légitimité.....	43
Conclusion.....	49
Bibliographie.....	51
Sources.....	52
Annexes.....	54

Introduction

« *Les hommes parlent, parlent entre eux, pour se dire des choses, des choses qui ont un rapport avec leur propre être-au-monde* »¹ : en soulignant l'omniprésence du langage dans les relations sociales, Claude Javeau rappelle combien celui-ci constitue une dimension ordinaire et pourtant centrale du monde social. En effet, « *toutes les relations sociales, quelle qu'en soit la nature, mettent en œuvre du langage. [...] Le langage n'a pas tout à fait le même statut que d'autres « objets » tels que l'école, la famille, la ville, etc., car, dans chacun de ces objets, le langage est présent* »².

Le recours au langage étant à la fois quasi-universel et omniprésent, il est un objet de connaissance indigène par excellence : il est communément perçu comme social au sens où il fait l'objet de processus d'apprentissage reconnu comme tel par les acteurs, et sa pratique même laisse entrevoir, dans l'interaction, la diversité de ses usages. Tantôt pensées comme intimes et personnelles, tantôt perçues par un prisme normatif (ce qui se dit et ce qui ne se dit pas), les pratiques langagières sont un terrain particulièrement fertile pour la sociologie puisqu'elles se prêtent merveilleusement à son geste fondateur : dénaturer et désindividualiser des pratiques, en les resituant dans les structures sociales qui les ont produites.

Dans le même temps, cela impose d'établir une définition claire du langage qui ne se heurte pas au polysémisme ou au flottement définitionnel qu'implique le sens commun. Nous entendons ici le langage ou les pratiques langagières comme « *acte effectif, en relation avec l'ensemble des actes (langagiers ou non) dont il fait partie* »³. Cette définition se distingue de celles qu'en ont les linguistes, qui tendent à la définir comme « *systèmes de formes porteuses de sens* »⁴. Notre sociologie nomme donc langage ce que la linguistique qualifie de discours, c'est-à-dire les pratiques d'usage et de non-usage de modalités particulières d'expression verbale.

Cela implique de ne pas traiter le langage comme un simple vecteur d'opinions qu'il s'agirait de comparer ou d'informations qu'il faudrait évaluer, mais comme une pratique socialement différenciée. Ce qu'elle dit des locuteurs ne se limite donc pas à *ce qu'ils disent* ou entendent dire, mais inclut également leur *façon* de dire ou de taire les choses. Notre

¹ Claude Javeau, « Sociologies du langage », *Raisons politiques*, 2001, vol. 2, n° 2, p. 79.

² Bernard Lahire, « Achard Pierre, La sociologie du langage. », *Revue française de sociologie*, 1994, n° 35-1, p. 144.

³ Pierre Achard, *La sociologie du langage*, Paris, Presses universitaires de France, 1993, p. 10.

⁴ *Ibid.*, p. 121.

analyse porte donc sur les usages différenciés du langage, plutôt que sur le contenu des prises de parole.

Si le caractère socialement construit – et donc socialement situé – du langage est communément admis, les connaissances indigènes du langage tendent à ignorer sa capacité créatrice ou « *originaire* »⁵. Les paroles ne sont pas des « *énoncés dérivés, constatifs, simples enregistrements d'un donné préexistant* »⁶, elles façonnent le monde social et construisent des réalités en les rendant socialement reconnues. Cela peut émaner explicitement de la volonté du locuteur qui en est à l'origine : c'est là l'objectif des ordres, des prières, des insultes, des bénédictions... Dans d'autres cas moins explicites, la parole entérine ou renouvelle les catégories de pensée et de perception des acteurs qui la produisent ou la reçoivent.

Toutefois, cette capacité à faire reconnaître ses paroles comme une réalité n'est pas également distribuée : elle est d'autant plus importante que le locuteur détient une quantité importante de capitaux, ou *a minima* qu'il s'exprime conformément à la norme de son milieu, c'est-à-dire de façon légitime.

Le droit en est un exemple particulièrement seyant : on “prononce” une peine, “déclare” coupable... Il s'agit de l'énoncé performatif par excellence, c'est-à-dire d'une parole qui « *fait surgir à l'existence ce qu'elle énonce* »⁷. Cette capacité à dire le droit, le juste, et dans une certaine mesure le vrai, est monopolisée par des agents investis « *d'une compétence inséparablement sociale et technique* »⁸. Cette capacité est donc double : il s'agit d'une part du *pouvoir* socialement reconnu à dire le droit, du fait d'une position institutionnelle, d'une trajectoire professionnelle ou de diplômes ; et d'autre part de la *détention des connaissances requises* pour produire un discours conforme aux attentes du champ et toujours suffisamment différent de ceux qui lui sont extérieurs pour justifier ce monopole.

Il en va de même pour ceux qui font le droit : ils monopolisent la capacité à dire ce qui doit ou non provoquer une action de l'Etat et, le cas échéant, quelle doit être cette action. Légitimer ce monopole requiert pour ces agents de se positionner en porte-parole, et par là même, de faire exister la volonté qu'ils entendent exprimer⁹. Ainsi, à l'Assemblée Nationale, la parole des députés occupe une place centrale dans l'exercice de leurs fonctions : elle est

⁵ Pierre Bourdieu, *Langage et pouvoir symbolique*, Paris, Éditions Fayard, 2001, p. 66.

⁶ *Ibid.*

⁷ *Ibid.*

⁸ Pierre Bourdieu, « La force du droit », *Actes de la recherche en sciences sociales*, 1986, n° 64, p. 4.

⁹ P. Bourdieu, *Langage et pouvoir symbolique*, *op. cit.*, p. 224.

particulièrement encadrée, tant par des règles formalisées (règlement) que par l'instauration d'une hiérarchie en charge de leur application. Les temps et ordres de parole sont ainsi définis selon la position d'un membre dans son groupe politique, le poids dudit groupe, la nature de l'intervention et sa finalité.

La possibilité de s'exprimer dans ce contexte est déjà la manifestation d'une position dominante dans la société puisqu'elle témoigne de l'appartenance à un champ dominant et particulièrement cloisonné, auquel l'accès est conditionné à la détention de capitaux symboliques, culturels, sociaux, et dans une moindre mesure, économique, particulièrement importants. Cette concentration de capitaux en fait l'objet d'attention provenant de nombreux autres champs, y compris universitaire, qui prend l'arène parlementaire comme objet, en étudient la composition, les discours, les thématiques centrales, etc. A cet égard, la sociologie du Parlement est particulièrement riche : les discours des députés ont déjà été finement analysés, tant par le biais de méthodes qualitatives ou quantitatives.

Ces démarches avaient pour point commun d'analyser le sens des prises de parole, par exemple pour comprendre les registres de légitimation de soi quand on s'exprime sur le "social"¹⁰, ou pour déceler les conceptions différentes du rôle de député lors d'une réforme portant sur le lien entre politique et justice¹¹. Néanmoins, ainsi que le dit Pierre Achard, « *le langage comme activité tend à être ignoré tant qu'on ne voit que le sens et non la forme* »¹². A rebours de ce courant d'analyse de discours, nous entendons donc faire une sociologie du langage qui cherche à « *saisir les règles et mécanismes de cette production signifiante, articulée aux situations sociales et aux rapports sociaux dans les champs étudiés* »¹³.

Louis Pinto disait ainsi au sujet des intellectuels que l'on parle peu de leur oralité, « *comme s'ils existaient seulement par leurs textes et que leur enseignement n'en était qu'un simple complément* »¹⁴. Un constat similaire s'applique aux députés, qui sont usuellement appréhendés par leurs prises de positions politiques. Or, celles-ci ne sont pas abstraites : elles

¹⁰ Brigitte Gaïti et Annie Collovald, « Discours sous surveillances : Le social à l'Assemblée » dans *Le « social » transfiguré. Sur la représentation politique des préoccupations « sociales »*, Presses universitaires de France., s.l., 1990, p. 9-54.

¹¹ Claire de Galember, Olivier Rozenberg et Cécile Vigour, « Chapitre 7. Réflexivité, rôles parlementaires et discours de légitimation. Les apports du logiciel Alceste à l'analyse lexicale d'un débat parlementaire relatif aux rapports entre justice et politique » dans *Faire parler le parlement: méthodes et enjeux de l'analyse des débats parlementaires pour les sciences sociales*, Paris, LGDJ-Lextenso, 2013, p. 127-149.

¹² P. Achard, *La sociologie du langage, op. cit.*, p. 121.

¹³ Johannes Angermüller et Marc Glady, « La sociologie du langage. Perspectives d'un champ émergent », *Langage et société*, 24 mai 2017, vol. 160161, n° 2, p. 168.

¹⁴ Louis Pinto, « La communication orale en milieu académique – Note programmatique », *Regards sociologiques*, 2024, n° 63, p. 12.

se construisent, se véhiculent et se manifestent par des voies concrètes, autrement dit par des pratiques, entre autres langagières.

De fait, cette position dominante de député et l'encadrement institutionnel des conditions d'exercice de la profession ne suffisent pas à uniformiser pleinement les pratiques langagières entre les députés. Ceux-ci évoluent dans ce que nous nommons l'espace parlementaire, sous-espace du champ politique dans son acception bourdieusienne, c'est-à-dire entendu comme un microcosme social relativement autonome régi par une loi qui est commune à tous les agents qui le composent (*nomos*), et au sein duquel se joue une lutte pour se rapprocher du centre, c'est-à-dire pour accumuler les capitaux qui y sont valorisés¹⁵. Ainsi, si le cloisonnement du champ politique précédemment établi peut conduire à l'émergence d'une norme commune, l'on peut aisément supposer que cela rend, dans le même temps, les stratégies de distinction encore plus nécessaires et subtiles.

Capter cette distinction à travers les pratiques langagières permet de voir ce que l'analyse de discours occulte, c'est-à-dire de situer les agents dans un espace des positions sociales, qui ne se superpose pas nécessairement à l'espace des prises de positions politiques. Cela se fait par l'attention à des particularités du discours des individus : en somme, à comment ils parlent. Par exemple, Pierre Achard, dans la lignée des travaux de Labov sur la prononciation du "r"¹⁶, expliquait qu'« *aucun locuteur ne fait toutes les liaisons, mais aucun ne les omet toutes. Chaque locuteur a une propension inégale à les faire* »¹⁷. Il en découle que cette variation des pratiques langagières « *ne peut être observée que statistiquement* »¹⁸. Lorsque les résultats de cette démarche variationniste sont rapportés à des positions sociales, ils permettent de capter les structurations sociales sous-jacentes à une pratique langagière donnée, ainsi que l'a fait Basil Bernstein s'agissant du lien entre vocabulaire et classe sociale¹⁹. Ces pratiques n'étant pas aléatoirement distribuées au sein d'une population, il est possible que leurs articulations spécifiques produisent des styles de paroles (que nous nommerons également des façons de parler).

La théorie bourdieusienne du langage, dans la lignée de laquelle nous nous inscrivons, permet de concilier cette approche structuraliste qui se concentre sur les « *dispositions*

¹⁵ Pierre Bourdieu, *Choses dites*, Paris, Éd. de Minuit, 1987, p. 126.

¹⁶ William Labov, *The social stratification of English in New York City*, 2nd ed., s.l., Cambridge University Press, 1966, 498 p.

¹⁷ P. Achard, *La sociologie du langage*, op. cit., p. 7.

¹⁸ *Ibid.*

¹⁹ Basil Bernstein, *Langage et classes sociales: codes socio-linguistiques et contrôle social*, Paris, Éd. de Minuit, 1975, 347 p.

socialement façonnées de l'habitus linguistique »²⁰ avec le rôle que jouent les configurations interactionnelles précises dans lesquelles se trouvent les acteurs, c'est-à-dire les « *système[s] de sanctions et de censures spécifiques* »²¹ qui forment un « *marché linguistique* »²² donné, et que la sociologie interactionniste a largement exploré.

En traitant les députés non seulement comme des députés mais comme des locuteurs socialement situés, et en étudiant la distribution et la variation des pratiques langagières entre eux, nous entendons mettre à jour les logiques de hiérarchisation qui structurent ce champ (qui domine ?), et leur expression incorporée (comment parlent ceux qui dominent ?). Cela soulève la question de savoir comment se distinguent les dominants par la parole dans un espace *a priori* codifié et homogène, et inversement, cela interroge sur les stratégies d'expression mises en œuvre par les dominés.

Ici, nous entendons la domination au sens bourdieusien, c'est-à-dire comme la dialectique qui se joue entre la détention des capitaux et des ressources valorisées dans un champ donné (qui est ainsi un champ de force) et la capacité à les imposer comme légitimes (ce qui en fait un champ de luttes), le tout permettant à certains d'y occuper une position centrale, sans toutefois que l'on puisse préjuger de la nature de ces ressources. Certaines propriétés dominantes dans le macrocosme social pourraient être retraduites spécifiquement dans le champ parlementaire, mais il est également possible que certaines ressources perdent ou gagnent de la valeur dans ce champ spécifique. Il conviendra de faire émerger cela en dégagant les caractéristiques individuelles qui distinguent le plus les individus et qui se rapportent le plus à leurs pratiques linguistiques : il peut s'agir du genre, de l'âge, de la profession antérieure à l'élection, de l'ancienneté... Initialement indéfinies, les conditions de la domination au sein du champ parlementaire ne sont pas un angle mort pour autant : leur mise au jour est à l'inverse l'un des objectifs de notre recherche, et ne pas les présupposer permet de percevoir les logiques de hiérarchisation propres à l'espace étudié qui, si elles reflètent dans une certaine mesure les structurations plus larges du macrocosme social (genre, race, classe), les déclinent certainement de manière spécifique.

Symétriquement, les pratiques langagières se déclinent à travers « *mille signes expressifs* »²³ : faire émerger des styles de parole requiert donc « *un travail de construction d'indicateurs permettant la mise en relation aussi systématique que possible des propriétés*

²⁰ P. Bourdieu, *Langage et pouvoir symbolique*, op. cit., p. 60.

²¹ *Ibid.*

²² *Ibid.*

²³ L. Pinto, « La communication orale en milieu académique – Note programmatique », art cit, p. 12.

élocutoires et des propriétés des positions et des dispositions des agents »²⁴. Cette construction d'indicateurs recèle le cœur de notre question : déterminer par quelles pratiques les agents se distinguent ou se conforment, de même que comprendre les logiques de domination propres au champ analysé, requiert une objectivation empirique aussi rigoureuse que possible. Pour l'étude de notre corpus, nous construisons un ensemble d'indicateurs statistiques fondés sur des expressions régulières, c'est-à-dire sur la correspondance éventuelle entre une chaîne de caractères définie par le chercheur et le contenu de chaque locution. Ces indicateurs peuvent ensuite être agrégés à différentes échelles : pour toutes les locutions d'une même séance, d'un type de loi donné, d'un groupe parlementaire, etc. Dans le cas présent, nous agrégeons les indicateurs à l'échelle du locuteur afin de déceler les variations interindividuelles et de les rapporter à des caractéristiques sociodémographiques.

En somme, il s'agit de comprendre comment les pratiques langagières participent et reflètent les hiérarchies propres au champ parlementaire. Si celui-ci constitue un espace fortement codifié, au sein duquel les conditions de prise de parole sont encadrées, les façons de parler y demeurent socialement différenciées. Certaines manières de parler peuvent ainsi fonctionner comme des ressources de distinction et de légitimité, permettant à certains agents d'occuper des positions dominantes dans cet espace. Il convient alors de déterminer quels styles de paroles correspondent aux différentes positions possibles dans le champ. Nous cherchons donc à comprendre dans quelle mesure les pratiques langagières des députés participent des logiques de hiérarchisation propres au champ parlementaire, et dans quelle mesure elles les reflètent ?

Autrement dit : si parler est une pratique socialement différenciée, comment se distribuent les styles de parole dans un espace où tous les locuteurs occupent formellement la même position, et quelles propriétés sociales structurent ces variations ?

Afin de répondre à cette question, nous nous fondons sur le traitement statistique des retranscriptions de l'ensemble des séances publiques ayant eu lieu à l'Assemblée Nationale entre le 18 juillet 2024 et le 30 janvier 2026 inclus. En parallèle, et afin de caractériser nos locuteurs, nous employons les données issues du Répertoire national des élus (RNE) dans sa version du 1er trimestre 2026. Nous nous pencherons dans un premier temps sur les possibilités qu'offre le traitement statistique de nos données sur la parole à l'Assemblée Nationale ainsi que sur les limites de telles données pour une sociologie du langage (I), puis

²⁴ *Ibid.*, p. 24.

nous évaluerons dans un second temps différentes façons de situer nos locuteurs dans l'espace social ainsi que les implications analytiques de chacune d'elles (II). Enfin, nous présenterons une première ébauche de ce que signifie parler et dire à l'Assemblée en termes de position sociale (III).

I. Quantifier la parole

S'agissant des sociologues et des statisticiens étudiant le suicide, Baudelot et Establet écrivaient : « *qu'ils s'y résignent, l'objet dont ils traitent n'est rien d'autre que le suicide que médecins et policiers s'accordent à appeler ainsi. Mieux vaut le savoir que l'ignorer ; mieux vaut connaître ce suicide-là que rien du tout* »²⁵. Sans doute en va-t-il ainsi de l'ensemble des données dont la sociologie peut traiter sans les avoir produites, et nos locutions n'échappent guère à la règle. Transcrites et publiées par l'institution parlementaire, leur usage interroge : que masquent-elles (A) et que peut-on tout de même en dire (B) ?

A. Que valent les comptes-rendus de séance ?

« Depuis plusieurs années, la chambre des députés a reconnu la nécessité de donner aux débats parlementaires une publicité qui leur a, jusqu'à ce jour, manqué complètement. »

Anonyme, De la publicité des débats parlementaires, 1846.

C'est ainsi que l'auteur.e anonyme d'une brochure intitulée « De la publicité des débats parlementaires » et publiée en 1846 entamait son propos²⁶, à l'aube de la fin du contrat confiant à un journal, le Moniteur universel, la retranscription et la publicisation des débats au sein de l'Assemblée. L'auteur.e y prenait pour point de départ l'objectif de transparence et de publicité des débats qui avait dominé (quoique non sans débats) depuis la Révolution, et déplorait le manque d'efficacité et de visibilité qu'apportait l'organisation alors en vigueur : les rédactions étant déjà effectuées par des sténographes professionnels rattachés à l'Assemblée, il proposait en outre de centraliser l'impression au sein même de l'Assemblée afin de permettre la diffusion immédiate à tous les organes de presse, à l'inverse du quasi-monopole qu'entretenait le Moniteur universel jusqu'alors. Cette proposition prend en réalité part à de longues oppositions entre les tenants d'une information produite et véhiculée exclusivement par les organes de presse, quitte à ce que celle-ci soit "plurielle" ou variable

²⁵ Christian Baudelot et Roger Establet, « 2. Que valent les statistiques du suicide ? » dans *Durkheim et le suicide*, Paris cedex 14, Presses Universitaires de France, 2011, vol.8e éd., p. 53.

d'un journal à l'autre, et ceux en faveur de transcriptions uniques, officielles et gratuites des débats²⁷.

1. Des conditions de production des comptes-rendus

Aujourd'hui, la seconde proposition domine. La transcription des débats découle de dispositions constitutionnelles, qui prévoient que « *Les séances des deux Assemblées sont publiques. Le compte-rendu intégral des débats est publié au Journal officiel.* »²⁸. Celles-ci sont prises en charge par des rédacteurs appartenant au Service des comptes-rendus de l'Assemblée Nationale, à laquelle ils sont donc pleinement intégrés. Se relayant toutes les quinze minutes, ils prennent en note ce qui se dit en séance publique, y ajoutent des détails et remettent en forme ces prises de parole *a posteriori*. Les comptes-rendus de chaque journée sont publiés le lendemain.

Cet enregistrement et cette publicisation ne sont pas sans effets sur les pratiques que nous cherchons à mesurer : tous les indicateurs que nous serons amenés à construire porteront en eux les conséquences de la conscience qu'ont les acteurs du fait d'être filmés, transcrits, observables et mesurables. Cette conscience est d'autant plus forte que les deux dernières décennies ont vu émerger puis se multiplier des classements faits de multiples indicateurs supposés mesurer l'activité des parlementaires²⁹. Vraisemblablement confrontés à ces classements, en particulier dans la sphère médiatique, les députés s'adaptent : « *les plus réservés déclarent « faire ce qu'il faut » pour ne pas mal figurer dans les statistiques du site, quand d'autres reconnaissent clairement déposer amendements et questions afin de rester dans le haut des classements. Leurs collaborateurs ont en général moins de pudeur et racontent volontiers (ou montrent directement à l'enquêteur) leurs stratégies pour multiplier amendements et questions rapidement* »³⁰. Ayant posé ce constat, E. Ollion en déduit que, conformément à la loi de Goodhart, « *quand un indicateur devient un objectif, il cesse d'être un indicateur pertinent* »³¹. Cette conclusion s'applique manifestement lorsque l'on cherche à établir des indicateurs d'activité. Toutefois, sa pertinence est plus nuancée lorsqu'il s'agit des seules transcriptions des débats. En effet, la parole est plus difficile à feindre que des dépôts

²⁶ Anonyme, « De la publicité des débats parlementaires », 1846.

²⁷ Yves Lavoine, « Publicité des débats et espace public », *Études de communication. langages, information, médiations*, 1999, n° 22, p. 115-132.

²⁸ « Constitution du 4 octobre 1958 » art. 33, al. 1.

²⁹ Étienne Ollion, *Les candidats: novices et professionnels en politique*, Paris, Presses universitaires de France, 2021, p. 136.

³⁰ *Ibid.*, p. 137.

³¹ *Ibid.*

d'amendements ou de questions, car elle est moins externalisable et constitue une manifestation incarnée de l'habitus. En outre, dans le cas où elle serait particulièrement investie par certains acteurs pour gagner en visibilité ou en légitimité, il ne s'agirait pas d'une « malformation » des données qu'il faudrait corriger, mais précisément du témoin qu'elle est un outil mobilisable dans le jeu des positions et des repositionnements sociaux³².

Les locuteurs ne sont toutefois pas notre seule source de biais potentiels. Ainsi, il résulte de la brève présentation que nous avons faite des raisons d'être des transcriptions qu'elles ne sont pas en tous points fidèles aux *verbatim*. Puisqu'elles visent à « *produire un document à la fois fiable, exhaustif et aisément lisible par tout un chacun* »³³, elles sont des versions volontairement remises en forme des locutions. Le caractère public et volontaire de cette transformation ne doit toutefois pas conduire à en surestimer l'effet : tout produit de la transcription par un tiers de la parole d'un individu porte nécessairement en lui les traces du transcripteur, ainsi qu'en témoigne l'importance classiquement accordée aux processus de catégorisation et de recodage dans enquêtes sociologiques³⁴.

Sans faire pleinement obstacle à l'analyse, cela implique tout de même des précautions. La teneur précise de ces dernières nous est apparue par l'observation en présentiel d'une séance publique, puis par la confrontation de nos observations aux comptes-rendus rédigés.

2. Oralité et corporéité

Tout d'abord, puisque le langage est « *un matériel singulier caractérisé par des nuances subtiles et infinitésimales* »³⁵, il semble manifeste que la simple transcription des mots prononcés ne suffit pas à rendre compte de l'ensemble de ces informations. À titre d'exemple, Rémi Lenoir disait que la transcription des cours de Pierre Bourdieu au Collège de France ne pouvait restituer « *tout ce qui tient à la personne, la voix, son timbre et ses modulations, les intonations, les variations prosodiques et articulatoires, les rythmes verbaux et, surtout, le langage des gestes, les mimiques, bref l'énergie, la passion et toute la posture corporelle qui est constitutive de la personne même* »³⁶. Ce même constat s'applique aux

³² *Ibid.*, p. 173.

³³ Fiche n°30 - Les comptes rendus, <https://www.assemblee-nationale.fr/dyn/synthese/organisation-assemblee-nationale/les-comptes-rendus>, (consulté le 25 mai 2026).

³⁴ Pierre Bourdieu, Jean-Claude Chamboredon et Jean-Claude Passeron, *Le métier de sociologue: préalables épistémologiques*, Nouvelle éd., Paris, Éditions EHESS, 2021, p. 142-156.

³⁵ L. Pinto, « La communication orale en milieu académique – Note programmatique », art cit, p. 12.

³⁶ *Ibid.*, p. 12-13.

députés sur lesquels porte notre étude, dont diverses dimensions de l'oralité et de la corporéité sont perdues dans le processus transcritoire. Ainsi, tout en considérant le langage comme « *une dimension de l'hexis corporelle où s'expriment tout le rapport du monde social et tout le rapport socialement instruit au monde* »³⁷, son aspect sémantique est le seul dont nous disposons.

Les formes d'expression non verbale peuvent néanmoins être consignées dans les comptes-rendus, à deux conditions : que celles-ci n'émanent pas du locuteur, et qu'elles soient collectives. L'on lira par exemple :

« Eh, Pinocchio ! C'est fini, les mensonges ? (Exclamations sur les bancs des groupes EPR et DR.) »

Locution d'Elisa Martin, le 31 octobre 2025

Ces incises ne sont pas rares, l'on en compte environ 18 990 dans le corpus brut³⁸. Si l'accès à la dimension incorporée du langage des locuteurs nous est impossible, l'on peut tout de même se projeter dans les réactions que sa parole suscite.

3. *Précision des verbatim et imprécision des locuteurs*

Par la suite, il convient de dire que, malgré la volonté affichée d'« *éliminer les imprécisions et imperfections de l'expression* »³⁹, la transcription conserve une bonne part des inexactitudes qui font le propre de l'expression orale, et demeure globalement fidèle aux mots employés par les locuteurs, ce qui permet de préserver les régularités caractéristiques de chaque individu. Cette préservation est d'autant plus importante que le discours prononcé est lu.

Pour autant, il s'agit là d'une autre limite possible de nos données : que faire si les paroles du locuteur ne sont en réalité pas les siennes ? Il semblerait que cet obstacle n'en soit pas tout à fait un : d'abord, car tout processus de mise à l'oral d'un texte écrit implique un certain nombre d'adaptations ou de réajustements qu'opère le locuteur. Ensuite, et y compris dans les cas où ces ajustements seraient minimes, le parler ne perd pas de sa valeur heuristique. En effet, l'intérêt de notre recherche n'est pas d'évaluer une compétence ou une capacité à produire une parole conforme à certaines normes, évaluation que la lecture viendrait fausser. L'objectif étant de déceler des styles de parole en lien avec la position

³⁷ Pierre Bourdieu, *Ce que parler veut dire: l'économie des échanges linguistiques*, Paris, Fayard, 1982, p. 89.

³⁸ Voir à ce titre la section I. B.

³⁹ Fiche n°30 - Les comptes rendus,

<https://www.assemblee-nationale.fr/dyn/synthese/organisation-assemblee-nationale/les-comptes-rendus>, art cit.

(actuelle ou désirée) dans l'espace parlementaire, la lecture directe de discours (qui donne un style écrit à l'oral) en est autant un indicateur que l'improvisation, et doit être analysée en ce sens. Elle peut indiquer un rapport du locuteur à son espace, à son travail, à ses compétences orales et précisément à ce qu'il estime être la norme ou l'attendu du champ. C'est par exemple ce que fait L. Pinto dans son analyse de l'oralité des intellectuels, repérant la lecture à voix haute et la conversation profane comme « *deux écueils symétriques* »⁴⁰ dont le locuteur, qu'il nomme ici « *animateur* » au sens d'E. Goffman, doit s'éloigner. La proximité excessive à l'un ou à l'autre de ces styles de parole (trop spontané ou trop dicté) est ainsi analysée au regard de la trajectoire biographique du locuteur car, loin de masquer un style plus véritable qu'il s'agirait de déceler, elle met au contraire en évidence sa conception de *ce* qu'il faut dire et de la *façon* dont il faut le dire.

4. *Locuteurs ou interlocuteurs ?*

Enfin, la forme même que prend la transcription rend certaines opérations particulièrement ardues : en effet, elle est organisée sous la forme d'un enchaînement linéaire de blocs correspondant chacun à un point de l'ordre du jour, et qui se subdivise lui-même en plusieurs blocs, un par prise de parole. Chacune de ces sous-unités comporte des balises spécifiant le nom du locuteur, son identifiant, et le texte de sa prise de parole. Cette organisation linéaire préserve la chronologie nécessaire à la compréhension mais ne permet que très difficilement de déduire l'identité de ceux à qui s'adressent les locuteurs.

Cette lacune apparente tient à plusieurs choses : d'abord, l'organisation de la parole au sein de l'Assemblée présente des spécificités. Le locuteur s'adresse tantôt à la personne qui préside la séance, tantôt à l'ensemble des députés présents⁴¹, mais le plus souvent aux deux simultanément. Plus rarement, elle peut également s'adresser à un certain nombre de députés identifiés mais non nécessairement mentionnés nommément, en particulier lorsque le locuteur souhaite répondre à un élément spécifique de ce qui a été dit antérieurement. Cela résulte de l'esprit même de l'organisation de la parole dans cet espace, y compris dans sa dimension architecturale : Jean-Philippe Heurtin soutient à cet égard que la forme de l'hémicycle, où ont lieu toutes les séances publiques que nous analysons, participe à la

⁴⁰ L. Pinto, « La communication orale en milieu académique – Note programmatique », art cit, p. 15.

⁴¹ Et, de façon plus indirecte, aux profanes qui pourraient être amenés à être témoins de ces paroles, par les transcriptions, la diffusion télévisuelle, etc. C'est ce que Cécile Vigour nomme le double-auditoire (v. Claire de Galemberg, Olivier Rozenberg et Cécile Vigour, « Chapitre 7. Réflexivité, rôles parlementaires et discours de légitimation. Les apports du logiciel Alceste à l'analyse lexicale d'un débat parlementaire relatif aux rapports entre justice et politique » dans *Faire parler le parlement: méthodes et enjeux de l'analyse des débats parlementaires pour les sciences sociales*, Paris, LGDJ-Lextenso, 2013, p. 127-149).

construction d'une « *grammaire de l'activité parlementaire* »⁴² distincte de l'interaction de type théâtrale en ce que « *le lieu de la parole [est placé] à chacune des places de la salle* »⁴³, ce qui est pensé par les concepteurs de cette architecture comme une façon de « *figurer cette légitimité des expressions polyphoniques des opinions particulières que l'on peut avoir sur un un problème à résoudre* »⁴⁴, et qui a pour effet concret que chaque acteur est tout à tour locuteur et auditeur.

Dans cette configuration, la position des acteurs qui ne parlent pas à un moment donné est floue. E. Goffman écrivait à ce titre « *qu'il existe, grosso modo, trois sortes d'auditeurs : ceux qui, participants non ratifiés, ne font qu'entendre, par inadvertance ou non, qu'on les y ait ou non encouragés ; ceux qui (dans une conversation à plus de deux) sont bien des participants ratifiés, mais auxquels le locuteur ne s'adresse pas spécialement ; ceux enfin, participants ratifiés, auxquels le locuteur s'adresse en effet, c'est-à-dire vers lesquels il s'oriente de manière à indiquer que ses paroles leur sont particulièrement destinées* »⁴⁵, précisant en outre que « *toutes sortes de variations mineures [de ces configurations] sont possibles* »⁴⁶. Il en résulte que l'interlocution n'est jamais précisément délimitée : elle est davantage l'extrémité d'un continuum de la parole, où deux locuteurs successifs s'adressent expressément l'un à l'autre et attendent de l'autre une réponse. Dans toutes les autres configurations possibles, et dès lors qu'il y a plus de deux personnes, les frontières entre auditeurs auxquels le locuteur s'adresse et simple auditeurs ratifiés se brouillent. Lorsque ces derniers viendront à s'exprimer à leur tour, il sera difficile de déterminer dans quelle mesure ils répondent au locuteur précédent tant qu'ils ne le désignent pas nommément.

Les données dont nous disposons permettent donc de déceler des co-locutions davantage que des inter-locutions. Elles ouvrent la voie à l'analyse de paroles prononcées dans un même contexte, sur un même sujet, sans pour autant pouvoir affirmer à qui s'adresse chaque locuteur, d'autant que ces derniers ne s'adressent que rarement à une personne définie.

Ce fait est d'autant plus saillant que l'hémicycle est en réalité un espace fluctuant, dont les députés entrent et sortent sans cesse, où ils se déplacent pour s'asseoir près de ceux

⁴² Jean-Philippe Heurtin, « Architectures morales de l'Assemblée nationale », *Politix, Revue des sciences sociales du politique*, 1994, p. 110.

⁴³ *Ibid.*, p. 128.

⁴⁴ *Ibid.*

⁴⁵ Erving Goffman, *Façons de parler*, Paris, Ed. de Minuit, 1987, p. 15-16.

⁴⁶ *Ibid.*, p. 16.

dont ils sont proches⁴⁷, entretiennent mille et une conversations parallèlement au discours que produit l'orateur du moment. Ainsi, la séance du matin que nous avons observée comptait une cinquantaine de députés présents à l'ouverture, ce nombre augmentant au fur et à mesure que la séance passait pour atteindre plus de 115 députés aux alentours de 11 heures. Il chute à nouveau après chaque interruption de séance. Cette variation, quoique fréquente, n'a pas la même ampleur selon les séances, l'assiduité étant vraisemblablement plus importante lors de séances où les enjeux (de politique, de visibilité...) sont perçus comme importants. La transcription moyennise ces dynamiques à l'échelle d'une seule séance, où l'on peut estimer le nombre de locuteurs et la répartition de la parole entre eux (éclatement ou concentration), et masque les nombreuses interactions "non publiques", c'est-à-dire qui se tiennent dans l'hémicycle mais à destination d'un nombre restreint de personnes, souvent proches les unes des autres, tant physiquement que politiquement. En d'autres termes, il existe, au sein de l'hémicycle, des paroles informelles, non ratifiées, et au contexte plus variables, qui ne nous parviennent pas.

Tous ces éléments invitent donc à interpréter les résultats avec prudence sans pour autant les disqualifier. Si la transcription parlementaire ne donne accès qu'à une version partielle et stabilisée de l'activité de l'hémicycle, cette version est à certains égards le reflet des dynamiques qui la structurent, tant s'agissant de l'absence d'interlocutions (qui s'apparentent davantage à des colocations) que concernant les prises de parole lues et dont l'auteur n'est pas toujours connu. Sur d'autres aspects, la transcription occulte des informations comme la dimension corporelle et orale du langage, mais elle permet néanmoins d'analyser certaines régularités des échanges et de dégager des tendances quant à la répartition de la parole et à la structuration des interactions visibles.

B. Faire parler la parole

Les spécificités de nos données exigent donc d'être attentifs tant dans la construction que dans le traitement de notre corpus.

1. Construction du corpus

Tout d'abord, nous avons porté notre choix sur la législature actuelle à des fins d'homogénéité contextuelle : en effet, les prises de parole dépendent de la configuration politique et des équilibres partisans au sein de l'Assemblée, qui sont dans ce cas-ci

⁴⁷ Y compris lorsque cela exige d'accourir à son siège pour chaque scrutin public (qui ne s'effectue pas à main-levée mais par des boutons, assurant la traçabilité et la transparence du vote).

relativement stables et faciles à appréhender. La période d'un an et huit mois sélectionnée (entre le 18 juillet 2024, début de la législature, et le 30 janvier 2026) nous semble donc satisfaire ce critère, tout en ayant lieu dans le contexte quelque peu inédit d'une tripartition politique équilibrée entre la gauche, le centre-droit et l'extrême droite^{48, 49}. Cette délimitation du corpus permet ainsi d'englober une période suffisamment courte pour demeurer comparable mais suffisamment longue pour que les variations observées ne soient pas le fait d'un contexte spécifique, par exemple le fruit de la division du travail entre députés selon les thématiques des lois abordées, qui privilégierait mécaniquement ceux dont la loi étudiée constitue une thématique de prédilection. Par ailleurs, la période choisie s'arrête avant les élections municipales, moment de renouvellement partiel de la composition de l'Assemblée lorsque des députés acquièrent de nouveaux mandats. Cela a pour effet de stabiliser les locuteurs. Le résultat brut de l'extraction des données est un tableau comportant 80 410 lignes correspondant chacune à une prise de parole. À chacune de ces lignes est associée un identifiant de séance (il y en a 415 au total), le nom du locuteur ainsi que le contenu de sa locution.

Cette première version compte 705 locuteurs, dont certains en doublon car la personne présidant la séance est nommée "le président" ou "la présidente" lorsqu'elle préside, et par son nom le reste du temps. Sur les 80 410 lignes, 3 225 ne présentent pas de locuteur, parfois parce que celui-ci n'a pas pu être identifié, notamment dans le cas de locutions courtes et impromptues :

« À cause de vous ! » ou encore « Ce serait pas mal de parler des Français aussi ! »

Locuteurs non identifiés, le 02 juillet 2025

Toutefois, la majorité de ces lignes (2 302 sur les 3 225) ne sont pas des locutions à proprement parler, mais des enregistrements de l'état de l'avancement, comme l'adoption ou le rejet d'un amendement ou d'un article, ou l'interruption d'une séance. Celles-ci sont immédiatement évincées du corpus car elles ne sont de toute évidence pas le produit de l'oralité d'un quelconque acteur, et ne présentent donc aucun intérêt pour notre analyse.

Par ailleurs, il arrive que certaines locutions ne puissent être attribuées à une unique personne.

⁴⁸ Bien qu'aucun de ces « blocs » ne soit homogène ou exempt de tensions.

⁴⁹ Nous ne nous attardons pas sur la composition de l'Assemblée, qui est abordée plus tard.

Cyrielle Châtelain : « Aujourd’hui, au Sénat, les Républicains ont remis en cause la possibilité pour les étrangers de se marier avec des Français. (« La honte ! » sur les bancs du groupe EcoS.) »

Fabien di Filippo et plusieurs députés du groupe RN : « Quand ces étrangers sont clandestins ! »

Le 20 février 2025

Dans le cadre de notre analyse, nous faisons le choix de retirer toutes les locutions dont le locuteur n’est pas défini, qu’il soit pluriel ou simplement non-identifié, car l’impossibilité de rattacher des éléments d’oralité au locuteur rend inopérant notre processus analytique. Nous nous réservons toutefois la possibilité de revenir à la transcription intégrale d’un débat, c’est-à-dire en incluant les formes d’expressions orales ou non, lorsque nous cherchons à resituer des paroles données dans leur contexte d’énonciation.

En outre, nous faisons le choix de retirer l’ensemble des locutions émanant de la personne qui préside une séance donnée, car celle-ci se trouve dans une position différente du reste des locuteurs : elle annonce l’examen de chaque texte, invite les locuteurs à s’exprimer, déclare le résultat des votes, etc. ; autant de paroles fortement encadrées et répétitives qui déséquilibreraient le corpus par leur nombre. En contrepartie, ce retrait, qui concerne 17 942 lignes, implique de tenir compte de la sous-représentation des personnes assurant la présidence de séance dans le corpus.

2. *Du (non) prétraitement*

Une fois l’ensemble de ces opérations effectuées, le corpus contient 59 241 lignes correspondant chacune à une locution. Au final, en croisant les locuteurs avec les députés présents dans le RNE, nous retirons également tous les locuteurs qui ne sont pas députés et qui occupent donc une position particulière au sein de l’hémicycle (il s’agit en majorité de ministres ou de secrétaires d’Etat, dans le cadre de questions au gouvernement ou de l’examen d’un projet de loi par lequel ils sont concernés).

Il nous semblerait inopportun d’appliquer à ce corpus quelque forme de “nettoyage” supplémentaire : en effet, là où l’analyse de discours cherche à repérer des thèmes récurrents ou des mots saillants, nous cherchons à observer des *façons* de parler, ce qui requiert de mettre à distance certaines pratiques tellement récurrentes en analyse de discours qu’elles y sont considérées comme de simples opérations techniques. Nous faisons par exemple le choix de ne pas “nettoyer” le corpus de ses ponctuations, qui sont parfois considérées comme des “parasites” lorsque l’on cherche à étudier les thèmes du langage,

mais qui seront dans notre cas très parlantes (qui s'interroge ? qui s'exclame ? qui saccade son discours et pourquoi ?).

Similairement, il nous semble nécessaire de travailler sur une version non-lemmatisée et non-stemmatisée du corpus, car si “évidemment” et “évident” renvoient à la même idée, le passage du premier au second nous fait perdre des informations cruciales (par exemple, sur une personne ou une catégorie de personnes qui auraient fréquemment recours aux longs mots ou aux adverbes pour rendre son discours plus formel).

Enfin, et plus important encore, la suppression des “mots-vides” (ou stopwords) est également considérée comme un passage obligé lorsque le but de l'analyse est par exemple de montrer l'évolution des thèmes au cours d'une période, ou les divergences de thématiques entre différentes personnes. Fastidieuse et traitée comme purement technique, cette étape est souvent effectuée à partir de listes préconstituées. Dans notre cas, l'approche de sociologie du langage dans laquelle nous souhaitons nous inscrire appelle à la prudence : nous ne pouvons guère nous permettre de supprimer aveuglément des mots sous prétexte qu'ils seraient vides de sens. En témoigne une liste de mots usuellement supprimés (extrait de la liste stopwords fr du package *Quanteda*, très utilisé en analyse textuelle sur R) : "moi", "notre", "nous", "votre", "vous"⁵⁰. Sans doute ces mots ne semblent-ils pas participer à la construction du sens pour certaines analyses, toutefois, il semble manifeste que nul locuteur n'utiliserait “nous” et “vous” de manière interchangeable. Dès lors, dire l'un ou l'autre est porteur d'un sens sociologique, à défaut d'être topique. Sans doute certains locuteurs sont-ils plus dans la confrontation, la désignation de l'altérité, tandis que d'autres préféreront parler d'eux-mêmes et de leur groupe. A l'inverse, il est possible que tous les députés disent à peu près de façon équivalente “vous” et “nous” dans un jeu d'opposition. Quelle que soit l'hypothèse qui se vérifie, il en ressort que la suppression de mots est un choix méthodologique plus que technique, injustifié au regard de l'analyse que nous entendons faire.

Nous n'aurons donc recours à aucune des ces méthodes (stopwords, processus de lemmatisation ou de stemmatisation) qui permettent de faire abstraction de la forme pour mieux atteindre le fond, puisque nous cherchons précisément à observer des pratiques langagières aussi brutes que possible.

⁵⁰ Kenneth Benoit et al., « *Quanteda: An R package for the quantitative analysis of textual data* », *Journal of Open Source Software*, 6 octobre 2018, vol. 3, n° 30, p. 774.

3. De la construction au traitement : les indicateurs

Arlette Farge disait au sujet des archives que « *Quel que soit le but, la recherche dans ce cas s'effectue à partir du même, de l'identique apparent, et la collection de textes recueillis sera ensuite traitée en essayant de briser le jeu des ressemblances pour trouver du dissemblable, voire du singulier* »⁵¹. Nos locutions ne manquent pas d'être "apparemment identiques" : elles émanent d'acteurs occupant tous la même fonction, sont toutes prononcées dans un même contexte institutionnel, dans un laps de temps court (selon les parties analysées, d'une minute à l'autre), portent sur les mêmes thèmes, ont *a priori* la même finalité (exprimer sa volonté et/ou légiférer selon sa volonté). Notre objectif sera alors de rechercher en quoi elles se distinguent, pour ensuite comprendre en quoi ces distinctions de paroles se rapportent à des écarts dispositionnels, à des caractéristiques proprement sociologiques des individus.

D'ailleurs, dans un espace aussi normé que le champ parlementaire, le semblable semble s'inviter partout et le singulier se nicher dans des détails. C'est là tout l'intérêt des méthodes quantitatives, qui permettent de déceler des régularités à partir de caractéristiques langagières *a priori* anodines. La linguiste Véronique Magri écrivait à ce titre que « *la massification des données est un atout indéniable dès qu'une variation linguistique doit être observée et analysée. L'émergence de régularités permet alors, au prix d'une simplification heuristique, de construire un modèle. Les singularités ne sont pas occultées pour autant, comme on pourrait le penser, mais elles servent justement de contrepoint au général pour affiner la modélisation* »⁵². Afin de tirer profit de cet atout, encore faut-il construire des indicateurs permettant de résumer l'information recherchée.

Notre démarche est ici inspirée de celle déployée par L. Boltanski dans son travail sur les lettres de dénonciation⁵³ : plus que la codification de variables, l'enjeu est de soumettre nos locutions et à travers elles, nos locuteurs, à une sorte de questionnaire permettant de les interroger concrètement sur la nature de leurs pratiques. Nous avons choisi pour cela d'avoir recours à des expressions régulières, ce qui revient à définir une chaîne de caractères donnée, puis de la comparer à chaque locution à la recherche d'une éventuelle correspondance. Le résultat de toutes ces "interrogations" successives est ensuite agrégé à l'échelle d'un individu,

⁵¹ Arlette Farge, *Le goût de l'archive*, Paris, Éd. du Seuil, 1998, p. 81.

⁵² Véronique Magri, « La linguistique et le nombre », *Le Français moderne - Revue de linguistique française*, 2020, p. 4.

⁵³ Luc Boltanski, Yann Darré et Marie-Ange Schiltz, « La dénonciation », *Actes de la recherche en sciences sociales*, 1984, p. 3-40.

comme l'on compilerait les réponses à différentes questions d'un questionnaire. L'avantage de cette méthode réside dans sa simplicité : elle permet en effet de construire une multitude d'indicateurs très précis qui, pris individuellement, ne sauraient être de bons descripteurs d'une locution, mais dont la simplicité est le garant de l'objectivation des caractéristiques, ce que les méthodes plus avancées telles que les grands modèles de langage ne garantissent pas. Elle présente en outre un second avantage : nonobstant la faible portée interprétative des indicateurs lorsqu'ils sont pris séparément, leur articulation permet de brosser un tableau précis et fidèle des styles de parole de nos locuteurs.

L'analyse de L. Boltanski comptait initialement 160 variables, dont certaines furent évincées car non pertinentes dans l'analyse. Dans le cas présent, et puisque la construction des variables est un enjeu central de notre recherche en ce qu'elle implique de déterminer quels sont les traits pertinents pour distinguer des façons de parler, nous procédons similairement, c'est-à-dire en construisant un nombre important d'indicateurs, inspirés dans leur majorité par le visionnage des prises de parole en séance publique, parmi lesquels nous pourrions par la suite isoler ceux dont l'apport heuristique est le plus important.

Les variables que nous avons construites sont nombreuses et ne sauraient toutes être détaillées ici ; elles peuvent toutefois être divisées en trois catégories : certaines sont des indicateurs de quantité (1), comme le nombre de locutions prononcées par chaque locuteur, leur longueur moyenne, la longueur de la locution la plus courte et de celle la plus longue chez chacun, le nombre de mots prononcés, le nombre de mots uniques, etc. D'autres variables portent davantage sur la forme des phrases (2) : leur ponctuation (interrogative, exclamative ou affirmative), leur rythme (avec le nombre de virgules), etc. D'autres encore nous renseignent plus précisément sur le contenu (3) : le nombre d'adverbes, de pronoms personnels parlant de soi (nous, je, notre...) ou des autres (vous, votre, vos...; eux, leur, leurs...)⁵⁴.

Les locutions ainsi interrogées et caractérisées, il est opportun de rappeler que « *tout acte de langage intervient dans l'incessant jeu de structuration/restructuration du social et ce*

⁵⁴ Pour chacun de ces indicateurs se pose la question de son agrégation à l'échelle de l'individu : par exemple, faut-il calculer la longueur moyenne de ses prises de parole ou se référer à la médiane ? Ou encore, faut-il calculer le nombre d'exclamations ou le rapporter au nombre total de prises de parole à des fins de comparaison ? Dans certains cas, l'un des indicateurs permet une meilleure distribution des effectifs ou une lecture plus simple des résultats : il est alors sélectionné. Dans d'autres, en particulier lorsqu'il s'agit du nombre d'occurrences d'un phénomène donné et de leur proportion, il peut être difficile de n'interpréter que l'un des deux, car toute comparaison entre individus exige qu'ils se trouvent dans des situations comparables. Ainsi, l'on ne rapportera pas la proportion de phrases exclamatives d'un député ayant moins de 10 locutions en séance publique avec quelqu'un qui en a des centaines.

jeu, il [le locuteur] ne peut le mener qu'à partir d'un lieu structuralement situé »⁵⁵. Cela appelle à mettre en relation les caractéristiques du langage jusque là développées avec des propriétés ayant trait aux positions et aux dispositions des agents.

II. *Situer socialement la parole*

« Aux questions « qui parle ? » et « que dit-il/elle ? », s'ajoute nécessairement « d'où parle-t-il/elle ? » ; ce « où » étant référé à une position sociale déterminante dans toute prise de position à l'égard du monde »⁵⁶. Dans notre analyse, nous faisons le choix de répondre à cette question par deux voies : d'abord par les données que contient le Répertoire national des élus et qui exigent certaines précautions d'usage (A), puis par la construction d'autres indicateurs permettant de compléter ou de renforcer les informations que contient ce dernier (B).

A. *Se classer pour se présenter : le Répertoire national des élus*

Le répertoire national des élus (ou RNE)⁵⁷, dans sa version accessible au public, est créé par décret en 2001⁵⁸ : les données qu'il contient sont issues de déclarations remplies par les candidats et déposées auprès des Préfectures afin de se présenter à une élection donnée. Il concerne tous les mandats électifs accessibles par une élection au suffrage universel, et regroupe des informations dites d'état civil, telles que le nom, le prénom, le sexe, le lieu et la date de naissance des élus; ainsi que le contenu du dossier de dépôt de candidature. Dans le cas particulier des députés, ceux-ci doivent remplir un formulaire⁵⁹ qui formalise leur candidature à l'élection législative : si certaines informations peuvent être directement traitées, comme le sexe et l'âge, d'autres méritent davantage d'attention. Tel est le cas de la catégorie socioprofessionnelle renseignée. Sa récolte s'effectue par le biais d'un encart dans le formulaire de candidature, qui se présente comme suit⁶⁰ :

2. SITUATION DU CANDIDAT
Profession ⁽¹⁾ : Numéro CSP correspondant ⁽²⁾ : _ _ Étiquette politique déclarée (facultatif) : Êtes-vous actuellement député ? oui ☐ non ☐

⁵⁵ C. Javeau, « Sociologies du langage », art cit, p. 84.

⁵⁶ *Ibid.*

⁵⁷ Ministère de l'Intérieur, « Répertoire national des élus », septembre 2025 et janvier 2026.

⁵⁸ « Décret n°2001-777 du 30 août 2001 portant création au ministère de l'intérieur d'un fichier des élus et des candidats aux élections au suffrage universel ».

⁵⁹ Cerfa N° 16110*02.

⁶⁰ L'ensemble de ces informations est également exigé pour les suppléants.

Cet encart est suivi des notes de bas de page suivantes :

- ⁽¹⁾ Pour les fonctionnaires, indiquer l'intitulé exact du poste occupé.
⁽²⁾ Voir la nomenclature des catégories socioprofessionnelles (CSP) figurant dans la notice du CERFA ou en annexe du mémento à l'usage des candidats disponible sur le site internet du ministère de l'Intérieur. Pour plus de précisions vous pouvez également consulter le site Internet de l'Insee.

Un complément y est adjoint : une page intitulée “Nomenclature des catégories socioprofessionnelles” qui présente un tableau à deux colonnes : “Code” et “Libellé”. Celui-ci liste les 42 postes du niveau détaillé de la nomenclature INSEE de 2003.

A ce stade de notre développement, il convient de préciser que le dépôt de candidature à la Préfecture n'est pas le seul moment où les députés sont amenés à déclarer une profession. En effet, suite à l'élection, ils leur revient déterminer ce qui sera inscrit sur leur page de présentation sur le site de l'Assemblée Nationale⁶¹ : ils ont alors le choix entre la nomenclature et un champ de réponse libre. Cette seconde source de données permet certes d'être au plus proche de ce que les acteurs disent d'eux-mêmes, mais elle rend les comparaisons particulièrement ardues, puisque son exploitation statistique impliquerait un travail de regroupements et de classification nécessairement biaisé et fondé sur des informations très parcellaires. Il nous semble donc préférable d'utiliser principalement les données issues du RNE, et de mobiliser cette seconde source de données pour des mises en perspective ponctuelles. Pour autant, le RNE n'est pas non plus exempt de biais.

Dans un premier temps, se pose la question de ses raisons d'être. Le texte de loi créant le RNE fait état de la nécessité d'organiser le suivi des candidats afin d'enrayer toute potentielle inéligibilité ou incompatibilité entre les activités professionnelles et le mandat⁶², ce qui explique que les fonctionnaires doivent indiquer l'intitulé exact de leur poste plutôt que leur catégorie socioprofessionnelle (CSP).

Toutefois, si cela nous informe (quoique sans doute de façon partielle) sur les logiques présidant la constitution de cette base de données administrative du côté des pouvoirs publics, cela ne renseigne guère sur les processus entrant en jeu dans l'élaboration de son contenu, qui incombe, elle, aux élus⁶³. Le terme de “déclaration” employé jusqu'ici se révèle alors trompeur : les candidats aux élections ne se contentent pas de répéter une information déjà sue et à laquelle ils adhèrent pleinement. L'inscription de la catégorie socioprofessionnelle est au contraire un fait qui participe de leur identité sociale, témoigne de leurs représentations du

⁶¹ Assemblée Nationale, « Liste des députés en exercice ».

⁶² « Décret n°2001-777 du 30 août 2001 portant création au ministère de l'intérieur d'un fichier des élus et des candidats aux élections au suffrage universel », art cit.

⁶³ Qui ne sont que candidats au moment où ils renseignent les informations demandées.

monde social et de leur place en son sein, mais également d'une mise en récit particulière qui se construit dans la perspective d'un tournant biographique : l'élection.

1. Classements négociés dans une nomenclature imposée

Tout d'abord, bien que l'objectif affiché du RNE soit le contrôle et le désamorçage des incompatibilités et inéligibilités, on peut imaginer que cette information n'a qu'un faible effet sur les pratiques déclaratives des candidats du fait de la multiplicité d'acteurs qui détiennent ou pourraient détenir cette information, et donc contrer une potentielle falsification. Ainsi, il est invraisemblable qu'un parti politique investisse un candidat dont la profession constituerait un obstacle à l'élection, de même qu'il est invraisemblable qu'un candidat indépendant ne tente de le masquer car en cas de victoire, les recours formés par les adversaires sont fréquents et peuvent entraîner de nouvelles élections s'ils s'avèrent fondés. Outre le caractère fortement désincitatif de ce phénomène, les candidats élus devront informer les autorités administratives (et plus particulièrement la Haute Autorité pour la Transparence de la Vie Publique) de leur situation professionnelle de façon extensive, en déclarant employeur, revenus, durée d'exercice, etc.

L'enjeu de l'auto-classification ne réside donc pas dans de potentielles falsifications qu'il s'agirait de mettre au jour, mais davantage dans les opérations subtiles de découpage du monde social et de positionnement en son sein que cet exercice requiert de la part des députés.

L'on sait à cet égard que les individus, lorsqu'ils sont amenés à exprimer leurs représentations du monde social, tendent spontanément à se référer à des espaces qui ne sont pas caractérisés par des frontières nettement délimitées, mais où ils situent des figures archétypales. C'est ce qu'affirment Luc Boltanski et Laurent Thévenot, qui écrivent que « *tout se passe comme si le sens social ordinaire s'orientait non par référence à un espace homogène, segmenté et orienté (que supposent les nomenclatures savantes) [...] mais par référence à un espace diversifié, avec ses points saillants et représentés, faciles à repérer dans le paysage social, à identifier et à classer* »⁶⁴. Cette non-segmentation de l'espace social se double d'une non-exhaustivité, ces deux caractéristiques n'émanant pas spontanément des représentations communes mais étant, au contraire, le produit d'un travail de taxinomie. Ainsi, les agents sociaux « *se situent plutôt les uns par rapport aux autres, dans des espaces plus restreints. Leur action obéit à des « logiques pratiques ». Celles-ci n'ont pas besoin de*

⁶⁴ Luc Boltanski et Laurent Thévenot, « Comment s'orienter dans le monde social », *Sociologie*, 31 mars 2015, Vol. 6, n° 1, p. 13.

*l'exhaustivité ni de la cohérence, qui sont le propre du scientifique ou du taxinomiste. »*⁶⁵.

Dans le cas que nous étudions, les individus interrogés ont à se référer à une nomenclature préexistante : ils doivent accommoder leurs conceptions propres à une taxinomie qui leur est imposée, et ce sans doute à partir d'images « *paradigmatiques* »⁶⁶ et « *stylisées* »⁶⁷ des différents groupes sociaux qui la composent. La large diffusion des catégories d'État depuis la création de la nomenclature INSEE⁶⁸ facilite cet appariement, mais celui-ci n'est pas une évidence pour autant : lorsqu'ils ont le choix de déclarer leur profession autrement que par la CSP (c'est-à-dire après l'élection), seuls 143 députés sur les 577 se réfèrent volontairement à la nomenclature.

Dans nombre d'enquêtes, la CSP est le produit de l'articulation de plusieurs informations par des professionnels du codage. Confrontés à des cas limites ou « *décristallisés* »⁶⁹, ces derniers tendent à user d'une vision globale de l'individu dont ils doivent coder la situation et à la rapprocher de leurs représentations paradigmatiques, mêlant ainsi connaissances savantes et indigènes du monde social. Les députés quant à eux, occupent dans ce processus le double rôle de classés et classeurs : ne disposant que rarement d'une connaissance savante ou d'une représentation totalisante du monde social, ils se référeront d'autant plus à des informations générales les concernant, quand bien même celles-ci ne sont pas parfaitement en cohérence avec leur représentation du poste qu'ils occupent. Ils tendront de ce fait à reconstruire leur CSP en articulant représentation de leur emploi et autres aspects potentiellement décristallisés de leur profil.

L'effet de cette articulation peut varier : il sera plus limité pour des professions fortement instituées, dont la délimitation repose sur des critères strictement définis (un unique diplôme d'État, l'inscription à un ordre professionnel...) tels que les médecins ou avocats, et plus fort pour des groupes sociaux aux contours moins nets ou faisant l'objet de luttes.

2. Mise en récit et trajectoire biographique

Une limite subsiste : jusqu'à présent, nous avons cherché à appréhender le rapport entre deux informations (la profession exercée et la CSP déclarée) en ne les considérant que comme des données, uniques et fixes. Les processus que nous décrivons s'inscrivent pourtant dans des trajectoires et doivent à ce titre être analysés de façon dynamique : d'abord car le

⁶⁵ Alain Desrosières, Alain Goy et Laurent Thévenot, « L'identité sociale dans le travail statistique : la nouvelle nomenclature des professions et catégories socioprofessionnelles », *Economie et statistique*, 1983, vol. 152, n° 1, p. 55-81.

⁶⁶ L. Boltanski et L. Thévenot, « Comment s'orienter dans le monde social », art cit, p. 15.

⁶⁷ *Ibid.*, p. 14.

⁶⁸ A. Desrosières, A. Goy et L. Thévenot, « L'identité sociale dans le travail statistique », art cit.

⁶⁹ L. Boltanski et L. Thévenot, « Comment s'orienter dans le monde social », art cit, p. 23.

parcours professionnel que les élus doivent classer n'est pas nécessairement linéaire et peut exiger de faire des choix, par exemple entre les différents postes occupés lorsque l'on cumule les activités. Ensuite, car cette "déclaration" de rattachement à un groupe social se fait dans la perspective d'une élection : elle participe à ce titre à la mise en récit que fait tout un chacun en donnant une cohérence à son parcours de vie *a posteriori*⁷⁰, et cette mise en récit se fait dans un contexte où l'individu cherche à monopoliser le pouvoir de représentation d'un groupe donné, ou à acquérir la capacité d'en dire les limites et *in fine* de parler en son nom. Pierre Bourdieu disait ainsi que l'action politique vise « *à faire ou à défaire les groupes [...] en produisant, en reproduisant ou en détruisant les représentations qui rendent visibles ces groupes pour eux-mêmes et pour les autres* »⁷¹. Les candidats désignent alors une appartenance au sein d'une catégorisation à laquelle ils peuvent ou non adhérer, et ce dans l'optique d'intégrer le champ politique, qui se veut « *un des lieux privilégiés de l'exercice du pouvoir de représentation ou de manifestation qui contribue à faire exister pleinement, c'est-à-dire à l'état objectivé, directement visible de tous, public, publié, officiel, donc autorisé, ce qui existait à l'état pratique, tacite ou implicite* »⁷². Au sein de ce champ, ils pourront donc investir plus ou moins un aspect sélectionné et mis en récit de leur position antérieure pour fonder leur action politique.

A cet égard, les catégories les plus poreuses sont celles d'agriculteurs, de professions libérales et de chefs d'entreprises. Des personnes qui pourraient se déclarer "professions libérales" (tels que des pharmaciens) se déclarent par exemple "profession scientifique"⁷³. Ces catégories sont donc particulièrement sujettes aux mises en récits et aux représentations individuelles des élus.

a) *Parler de soi, parler des autres ?*

La population des députés est particulière : le cloisonnement du champ implique de disposer de certains capitaux pour y accéder. Les professions libérales, chefs d'entreprise, cadres et anciens cadres y sont surreprésentés en comparaison avec la population active. Ces cinq groupes représentent à eux seuls 65% de l'hémicycle, tandis que d'autres professions, globalement moins dotées en capitaux, y sont moins représentées que dans la population générale (graphique n°1).

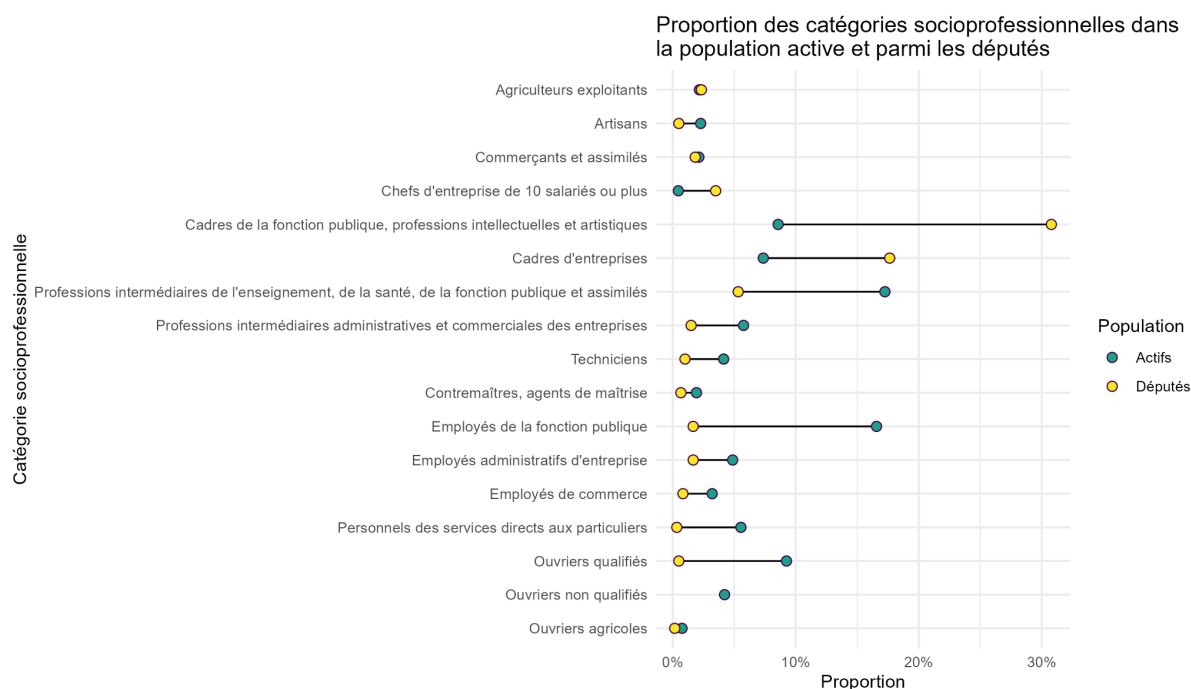
⁷⁰ Pierre Bourdieu, « L'illusion biographique », *Actes de la recherche en sciences sociales*, 1986.

⁷¹ Pierre Bourdieu, « Décrire et prescrire », 1981, p. 69.

⁷² P. Bourdieu, *Langage et pouvoir symbolique*, *op. cit.*, p. 224.

⁷³ Voir à ce titre l'exemple présenté page 26.

Graphique n°1



Source 1 : Dares-Drees-DGAFP-Insee, enquête Conditions de travail et Risques psychosociaux 2016.

Champ : France hors Mayotte, ensemble des salariés.

Source 2 : Répertoire national des élus, 2025 et 2026.

Champ : Ensemble des députés (hors retraités et nactifs) en 2025 et 2026.

Note de lecture : Les techniciens représentent 4% des actifs (en 2016) et 1% des députés de la législature en cours.

Pour mieux comprendre les possibilités d'interprétation qu'ouvre la CSP déclarée dans le RNE, il nous faut étudier les usages que font nos locuteurs de la catégorie à laquelle ils ont dit se rattacher lors du dépôt de candidature, et la mettre en relation avec les autres qualifications qu'ils ont pu retenir.

Tout d'abord, il convient de noter que pour appuyer leurs positions, les députés se présentent le plus souvent comme des « élus »⁷⁴, mais ils articulent cette source de légitimité qui leur est commune avec d'autres aspects plus personnels de leur positionnement social : s'agissant de la catégorie socioprofessionnelle, l'usage qui en est fait varie.

Le locuteur peut employer une catégorie techniquement inadaptée mais dont il se fait une représentation valorisante et suffisamment proche de la réalité de sa profession :

« *En tant que pharmacien d'officine*, je suis profondément attaché à notre maillage officinal de proximité, désormais gravement menacé... »

Jean-Pierre Bataille, le 1er juillet 2025. Se déclare « Professeur, profession scientifique » dans le RNE, et « Pharmacien » dans sa présentation sur le site de l'Assemblée Nationale. Maire de 1999 à 2025 (succédant à son père) et élu député en 2024

A l'inverse, le locuteur peut spontanément se désigner par un qualificatif qui l'inverse

⁷⁴ Dans notre corpus, ils disent par exemple « en tant que législateurs », « en tant que députés », « en tant que président de [telle commission] »...

ne recouvre qu'un aspect de la CSP, et de ce fait désigner un groupe plus large que celui déclaré, selon la dimension de sa trajectoire professionnelle qu'il cherche à valoriser dans l'échange :

« Il y a presque vingt-cinq ans déjà, nous connaissons l'essentiel des conséquences du changement climatique : inondations, montée des eaux, canicules, incendies de forêts. **En tant que scientifiques, nous** avons fait le job. **Nous** avons approfondi les connaissances et n'avons cessé de vous alerter – et ce malgré des salaires qui ont perdu 20 % de leur valeur et l'enfer bureaucratique entraîné par **vos** appels d'offres. »

Hendrik Davi, le 12/11/2024. Directeur de recherche à l'INRAE et élu depuis 2022, se déclare profession scientifique.

« **En tant que salarié**, j'ai connu en 2006 l'offre publique d'achat hostile de Mittal Steel sur Arcelor ; en 2014, **salarié et militant syndical**, j'ai assisté à la vente forcée d'une grande partie d'Alstom. »

Lionel Duparay, le 19/11/2025, élu depuis 2025, se déclare cadre d'entreprise dans le RNE et ingénieur dans sa présentation sur le site de l'Assemblée Nationale.

Les facteurs qui semblent structurer cet usage différencié de la CSP déclarée sont d'une part le degré d'institution du groupe (les « scientifiques », les « pharmaciens »...), et d'autre part la volonté d'investir le label qui résulte de la classification comme un groupe d'appartenance.

b) La députation : un milieu social qui s'ignore ?

Ce dernier aspect est fortement corrélé à la trajectoire même du député, qui peut faire montre d'un rapport plus ou moins professionnel à la politique. La cumulation et l'accumulation des mandats sont autant de facteurs qui participent de la construction d'un habitus professionnel d'élu et donc d'une forme d'obsolescence de la CSP déclarée.

Ainsi, l'on peut traiter d'une catégorie d'État par laquelle on s'est désigné lors du dépôt de candidature tout en la mettant à distance, c'est-à-dire sans véritablement s'y rattacher, comme le montre le “*nous*” de Nicole le Peih, qui la positionne en tant qu'élue et l'oppose aux agriculteurs, bien qu'elle se soit déclarée comme telle tant dans le RNE que sur le site de l'Assemblée Nationale, et qu'elle soit à la tête d'une exploitation agricole fondée avec son époux en 1990 :

« Cette proposition de résolution traduit des préoccupations que **nous** entendons tous : celles de **nos** agriculteurs et de **nos** filières agroalimentaires »

Nicole le Peih, le 27/11/2025, élue depuis 2001 et députée depuis 2017.

Cette professionnalisation de la politique est parfois même explicitée par les élus, qui

peuvent user de leur capacité susmentionnée « à faire ou à défaire les groupes »⁷⁵ pour faire valoir leur position d'élus comme une identité professionnelle. C'est par exemple ce que fait Laurent Mazaury au sujet de la création d'un statut d'élus local, déclarant :

« J'en sais quelque chose, *étant moi-même élu local* – comme nombre d'entre vous – depuis plus de quinze ans. ».

Laurent Mazaury, le 07/07/2025, élu depuis 2008.

A l'inverse, il ne mentionne pas, dans le corpus retenu, sa position de cadre dans une entreprise de communication, qu'il quitte lors de son élection en 2024 (soit seulement un an auparavant). En ne renseignant que sur la profession antérieure, la catégorie socioprofessionnelle issue du RNE masque donc les évolutions subséquentes aux élections et en particulier la construction d'un habitus professionnel propre aux élus.

En somme, l'analyse des catégories socioprofessionnelles renseignées dans le Répertoire national des élus appelle à une certaine prudence interprétative : loin de refléter fidèlement la profession d'un élu, celle-ci renseigne toutefois sur les représentations et les mises en récit opérées par les candidats, et ce dans le contexte particulier que constitue l'élection. Ce tournant biographique peut d'une part nécessiter une mise en cohérence dudit récit, et mène d'autre part à une reconfiguration de la position professionnelle en entraînant un éloignement plus ou moins prononcé (selon le nombre de mandats) à la profession déclarée. Cela implique deux usages possibles : le premier consiste à ne l'interpréter que comme ce que les députés disent d'eux-mêmes, et non comme ce qu'ils sont ; ce que nous faisons ici le plus souvent, en considérant qu'il s'agit d'un vague proxy de la position sociale effective ou désirée. Le second consiste à mobiliser la CSP issue du RNE conjointement à d'autres variables telles que celle déclarée sur le site de l'Assemblée Nationale ou encore le nombre de mandats afin d'objectiver la position sociale des députés, ce que nous ferons ici ponctuellement, et de façon plus extensive dans le cadre du mémoire de M2.

B. Le RNE : potentiel et prolongements

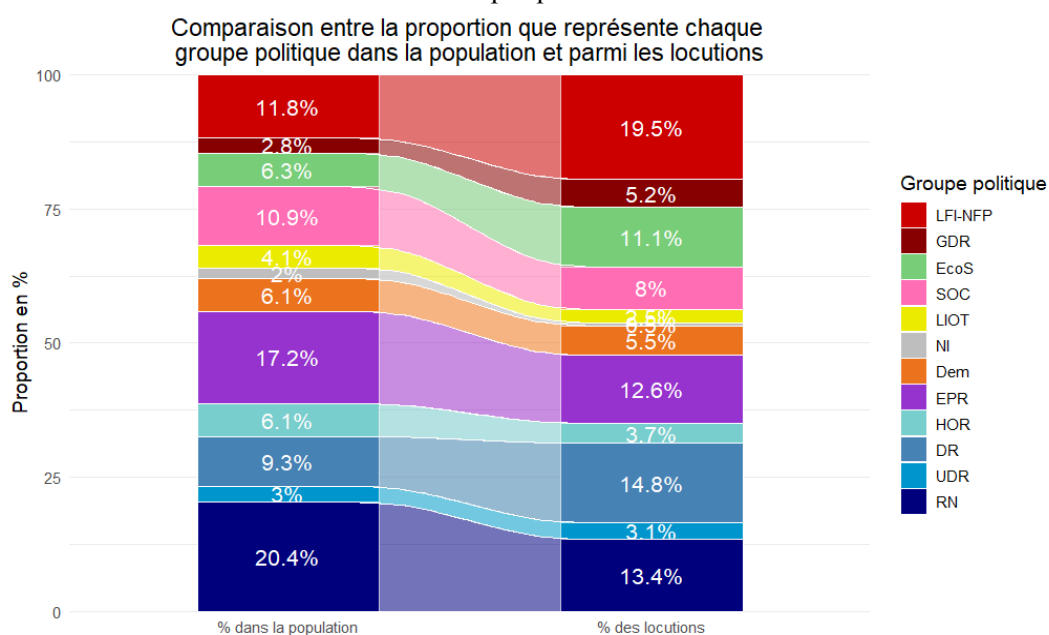
Les limites que nous avons trouvées au RNE sont donc nombreuses et exigent une multitude de précautions analytiques. Pour autant, elles constituent une première étape nécessaire pour situer les locuteurs dans l'espace social en nous fournissant quelques informations sociodémographiques.

1. Les variables sociodémographiques, substrat de l'analyse

⁷⁵ P. Bourdieu, « Décrire et prescrire », art cit, p. 69.

Tout d’abord, il est commun d’évaluer la composition de l’Assemblée en fonction des groupes politiques d’appartenance. Les logiques d’opposition et d’alliance qui la structurent sont saillantes et particulièrement investies tant par les acteurs en son sein que par les observateurs. En outre, elles fondent la répartition théorique des temps de parole (régies par le Règlement de l’Assemblée Nationale)⁷⁶ : elles sont donc le premier qualificatif auquel l’on peut recourir pour situer les agents dans ce champ et mettre leur position en relation avec leurs pratiques langagières. Cela permet d’observer que la composition de l’Assemblée en termes de groupes parlementaires n’est pas symétrique à la distribution de la parole (graphique n°2). Par exemple, les députés appartenant ou apparentés à la France Insoumise représentent 12% des députés mais 20% des locutions. Similairement, ceux de la Droite Républicaine sont à l’origine de 14% des prises de parole alors qu’ils ne sont que 9% de l’Assemblée. A l’inverse, certains groupes parlent moins souvent⁷⁷ que si la distribution avait été parfaitement égale : tel est le cas des députés du Parti Socialiste (8% des locutions mais 10.9% de la population), du Rassemblement National (13% des prises de parole pour 20% de la population), ou encore du groupe Ensemble pour la République (12.6% des locutions contre 17.2% des députés).

Graphique n°2.



Source 1 : Répertoire National des élus, Ministère de l’Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : Les députés du groupe LFI-NFP représentent 11.8% des députés et 19.5% des locutions dans la législature en cours.

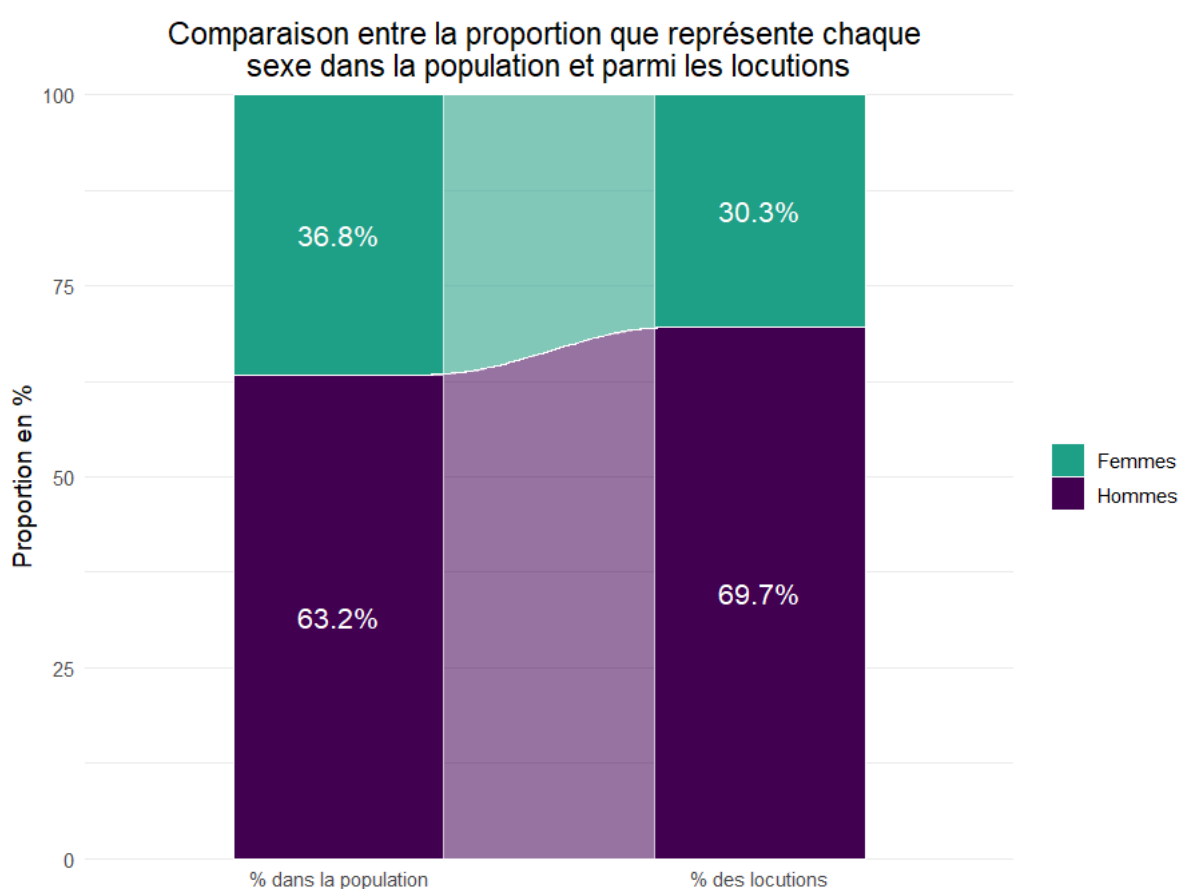
⁷⁶ *Règlement de l’Assemblée Nationale*, 2025, 336 p.

⁷⁷ Nous disons ici « moins souvent » plutôt que simplement « moins » car les données présentées sont celles du nombre de locutions par groupe rapporté au nombre total de locutions. Il s’agit donc d’un indicateur de fréquence qui ne tient pas compte de la longueur de ces locutions.

Il en ressort que la seule appartenance politique ne suffit pas à expliquer la distribution de la parole : ces variations trouvent donc vraisemblablement leur source dans d'autres caractéristiques des individus.

Le RNE fournit à cet égard une caractéristique sociodémographique simple mais heuristique : le sexe. Elle permet d'appréhender un autre aspect de la constitution de l'hémicycle : les hommes y sont surreprésentés à hauteur de 63% du total. Rapportée à la proportion de locutions qui émanent d'eux, cette information gagne en profondeur : ils sont en effet à l'origine de 70% des prises de parole de notre corpus (graphique n°3).

Graphique n°3.



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : Les femmes représentent 36.8% des députés et 30.3% des locutions dans la législature en cours.

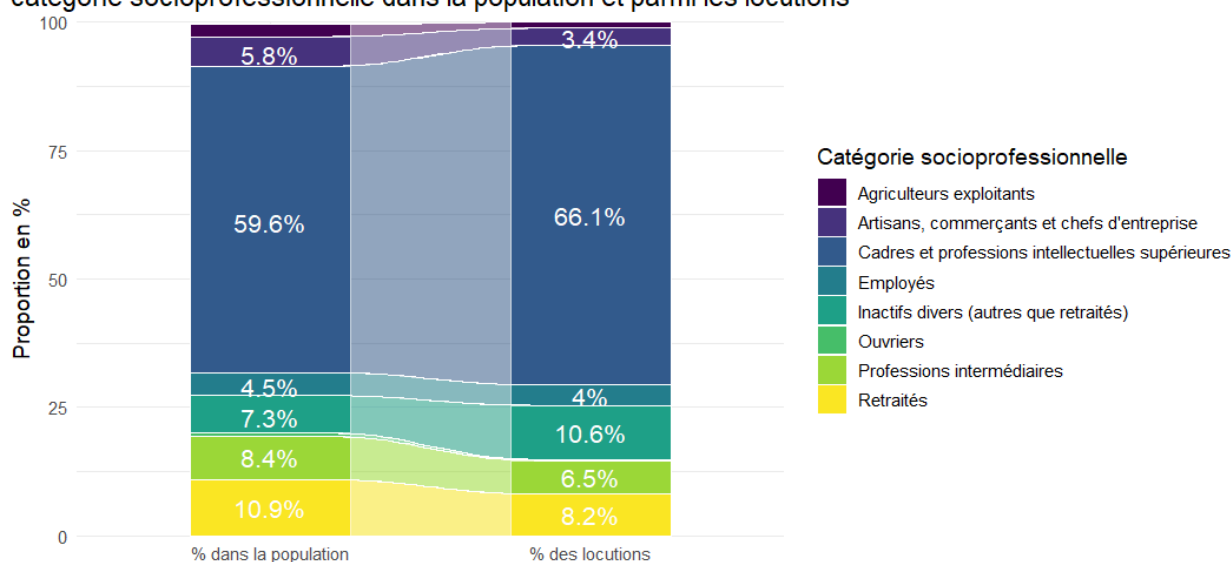
Ainsi, non seulement les hommes sont plus nombreux, mais ils interviennent également davantage dans les échanges. Cette surreprésentation peut être interprétée de manière exploratoire comme le produit de mécanismes de socialisation genrée, ou encore comme le résultat d'une division sexuée du travail parlementaire entre commissions et

séances publiques, et ces seules données ne permettent ni de confirmer ni d'infirmer ces hypothèses. Elles corroborent toutefois l'hypothèse selon laquelle la parole, y compris dans ses aspects les plus aisément saisissables comme sa distribution, participe des logiques de structuration et de hiérarchisation d'un milieu social. L'accès inégal à la parole, malgré un cadre institutionnel formellement égalitaire, confirme donc que les pratiques langagières reflètent et participent de logiques de domination, ici de genre.

Un constat similaire peut être fait concernant la catégorie socioprofessionnelle. Les groupes majoritaires sont aussi ceux qui s'expriment le plus souvent, à l'instar des cadres et des professions intellectuelles supérieures qui sont la première catégorie en termes d'effectifs (59%), et sont à l'origine de 66% des locutions, soit 7 points de pourcentage de plus que si la parole était répartie également entre tous les acteurs (graphique n°4). Un phénomène similaire, quoique moins marqué, s'observe chez les inactifs divers (à l'exclusion des retraités), ce qui concerne donc des étudiants, des personnes non occupées mais non chômeuses, etc. Autrement dit, les individus dont les dispositions et les capitaux permettent le mieux d'accéder au champ parlementaire, et qui y sont surreprésentés, sont aussi ceux qui y concentrent la parole.

Graphique n°4.

Comparaison entre la proportion que représente chaque catégorie socioprofessionnelle dans la population et parmi les locutions



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : Les artisans, commerçants et chefs d'entreprises représentent 5.8% des députés et 3.4% des locutions dans la législature en cours.

2. Prolongements et enrichissements

Ces caractéristiques sociodémographiques, au-delà d'être des indicateurs permettant de cerner la position de nos locuteurs dans le champ parlementaire, confirment que les pratiques langagières peuvent être fiablement corrélées à la position dans l'espace social. Pour autant, elles ne permettent de capter qu'un aspect limité de ces positions (CSP, âge, sexe, groupe parlementaire). Afin de la compléter, plusieurs voies sont possibles. Une première consiste à croiser notre base de données avec d'autres préexistantes et permettant d'obtenir des informations supplémentaires : il est envisageable de faire cela, à l'occasion d'un approfondissement pour le M2, avec les Portraits de circonscriptions législatives de l'INSEE, publiés en 2022, et constitués par le croisement du recensement de la population, du dispositif sur les revenus localisés sociaux et fiscaux (Filosofi) et de la base permanente des équipements (BPE)⁷⁸. Cette base nous renseigne sur la population de chaque circonscription, avec des indicateurs tels que la distribution entre les communes rurales, urbaines, ou appartenant à l'aire d'attraction d'une ville; la distribution des âges, du statut d'emploi (actif, retraité, étudiant, au chômage, autre inactif...), des diplômes et des catégories socioprofessionnelles dans la population. Ces informations pourraient nous informer sur les conditions d'élection d'un député et sur la configuration qui a permis son accession au champ que nous étudions.

Dans le même temps, elles ne permettent pas de rendre compte des logiques de structuration interne au champ : l'on ignore toujours ce qui distingue un député "légitime", de premier plan, central dans son champ, d'un député périphérique dans l'espace ou dominé. C'est là l'intérêt de la seconde voie possible d'enrichissement des données, qui consiste à entrer manuellement des informations collectées à différents endroits, à les systématiser selon une grille de codage afin d'assurer leur comparabilité. Cette méthode, quoique chronophage, permet d'obtenir des informations répondant à des questionnements spécifiques et donc de déceler plus finement des caractéristiques qui participeraient à définir la position d'un agent dans son champ. Par exemple, pour capter le degré de médiatisation des députés, l'on pourrait construire des indicateurs fondés sur le nombre d'articles de presse les mentionnant sur une période donnée, ou sur le nombre de personnes les suivant sur les réseaux sociaux, ou encore sur le nombre de modifications apportées à leur page Wikipédia : autant d'informations qui montreraient qu'un député est plus ou moins connu en dehors de l'espace parlementaire. De même, il serait possible d'approximer la trajectoire biographique des députés à partir de leur

⁷⁸ INSEE, « Portraits des circonscriptions législatives », 2022.

niveau et domaine d'études, mais ces informations seraient nécessairement issues de données de presse et non vérifiables. D'autres indicateurs plus objectivants sont envisageables, tels que la date du premier mandat, le nombre de mandats à l'Assemblée et le fait d'avoir ou non eu un mandat local ou national autre que la députation. Ces derniers permettraient de mettre en exergue le degré de professionnalisation. Ces variables auraient le double effet de situer le locuteur dans son champ d'une part et de contextualiser les autres données telles que la CSP déclarée d'autre part, puisque celle-ci peut correspondre à un aspect plus ou moins structurant de la réalité d'un individu. Par exemple, un député élu pour la première fois en 2024 après plusieurs échecs électoraux et un passage à la retraite n'est pas dans la même situation qu'un autre dont il s'agit également du premier mandat mais qui était étudiant et assistant parlementaire auparavant, alors même que les deux sont catégorisés comme "inactifs". En ce sens, l'ancienneté est à la fois une donnée interprétable en tant que telle comme indicateur d'homologie entre les dispositions d'un individu et les attendus de son champ, et un contrepoint permettant la mise en perspective des autres variables dont nous disposons déjà. Elle est donc l'un des principaux leviers d'approfondissement de notre analyse au M2. Similairement, il serait intéressant d'estimer systématiquement, par des indicateurs objectivants tels que les échecs électoraux ou l'expérience en tant qu'auxiliaire politique (membre du cabinet d'un élu, assistant.e parlementaire), le temps passé dans « la file d'attente »⁷⁹ permettant d'accéder à la position de professionnel de la politique, dont nous verrons dans la suite du développement qu'elle constitue un élément socialisateur important dans l'espace parlementaire. A défaut d'avoir pu entrer manuellement ces informations pour l'ensemble des députés dans le cadre du mémoire exploratoire, nous nous servirons de la collecte partielle déjà effectuée comme élément de contextualisation ponctuelle lorsque nous illustrons un profil archétypal ou un cas particulier, en la mobilisant à l'échelle individuelle plutôt que systématique.

Les données dont nous disposons déjà ne sont pas inexploitable pour autant : elles suffisent en effet pour dégager des régularités dans les pratiques langagières et les associer à des caractéristiques sociodémographiques générales.

⁷⁹ É. Ollion, *Les candidats*, op. cit., p. 78.

III. Parler et dire : une ébauche de la domination par les pratiques langagières

Les pratiques auxquelles nous nous intéressons peuvent être approchées par différents aspects : d'une part, tous les députés n'ont pas la même propension à parler, certains parlent plus ou moins souvent, plus ou moins longtemps, et devant plus ou moins de personnes (A). D'autre part, lorsqu'ils le font, ils ne disent pas les mêmes choses et ne les disent pas de la même façon (B).

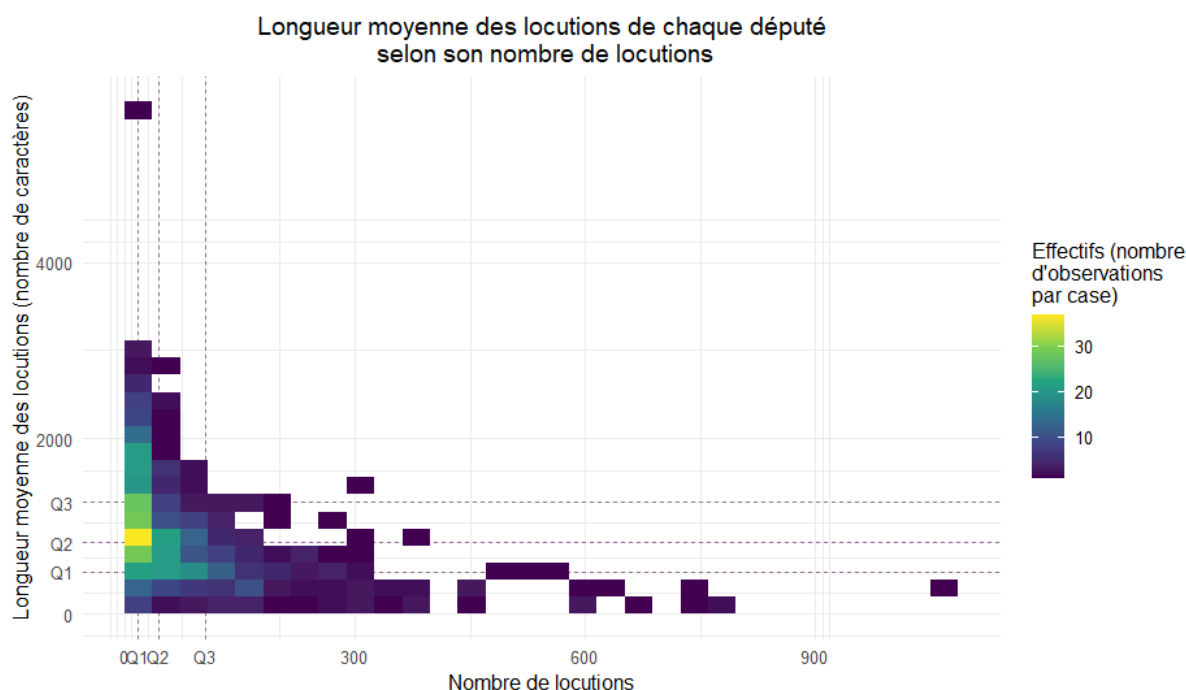
A. Parler : Distributions de la parole et interactions

« *Je veux devenir acteur mais pour être acteur il faut connaître les rôles et le jeu de rôle* »⁸⁰ tels sont les propos d'un député novice rapportés par E. Ollion dans son enquête sur la députation et les effets des trajectoires qui mènent à cette fonction. Il constate que ces novices, qu'il définit comme ceux dont la distance au champ politique était la plus importante (primo-députés, non militants, sans ancrage local, sans expérience en cabinet...) sont « *relégués au second plan de l'Assemblée nationale* ». Ils n'y occupent pas de positions de premier plan, sont moins souvent invités dans les médias, etc. Ainsi, il semblerait que « *du fait de la rareté des places, du fait des logiques de concentration des richesses sur une minorité, quelques élus récoltent beaucoup* ». Cette concentration de l'activité, de la visibilité et des ressources autour de certains élus correspond-elle à une concentration de la parole ?

⁸⁰ *Ibid.*, p. 156.

1. Une distribution inégale de la parole

Graphique n°5.

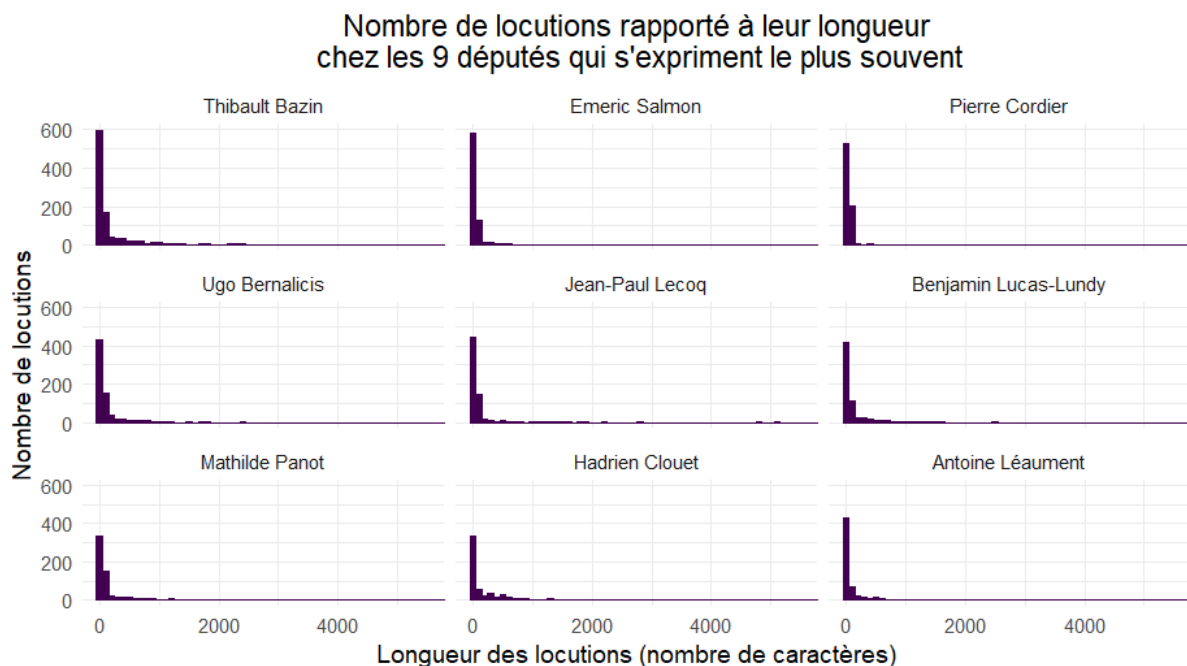


Le graphique n°5 montre que la majorité des députés se concentrent entre 0 et 125 locutions au cours de la période étudiée. La moitié des députés ont parlé 45 fois ou moins, et seuls 25% d'entre eux ont parlé plus de 105 fois. La longueur moyenne des prises de parole (pour chaque personne) se concentre similairement en deçà des 2000 caractères, avec une médiane à 810. Si le nombre de caractères nous paraît approximer la longueur d'une prise de parole de façon plus précise que le nombre de mots, cet indicateur peut néanmoins être abstrait. A titre d'exemple, voici une intervention longue de 800 caractères :

« La légèreté avec laquelle la gauche parle de l'article 33 nous inquiète. Pour les jeux olympiques de 2024, des centaines de milliers de contrôles ont été effectués, dont 100 000 environ sur des agents de sécurité privée. Ces 100 000 contrôles ont permis d'établir qu'une centaine de personnes fichées S avaient envoyé leur CV pour être agent de sécurité privée. On peut s'attendre à trouver parmi eux des personnes dangereuses pour la sûreté des sites olympiques, voire des individus qui avaient candidaté dans l'intention de commettre un délit ou, pire, un attentat. Nous le disons depuis longtemps et cette légèreté le confirme : j'espère qu'en 2027, les françaises et les français feront le bon choix et ne porteront pas au pouvoir les défenseurs des délinquants, présents sur les bancs d'en face. »

L'importante dispersion le long de l'axe des abscisses montre toutefois une hétérogénéité des députés dont la longueur moyenne des locutions est faible : certains ont peu de locutions (à gauche du graphique), d'autres au contraire interviennent très souvent : nombre de ces interventions sont d'ailleurs courtes, ce qui explique une moyenne basse. C'est ce qu'indique le graphique n°6 montrant, pour les 9 députés⁸¹ ayant le nombre de locutions le plus élevé, la distribution des longueurs de leurs locutions. L'écrasante majorité de leurs locutions se concentrent en dessous des 500 caractères, et le nombre d'interventions décroît systématiquement à mesure que leur longueur croît. Cela suggère que les députés qui s'expriment le plus souvent se livrent certes à un certain nombre de prises de paroles longues (à droite des graphiques), mais que leur spécificité réside dans cette multitude de locutions très courtes.

Graphique n°6.



Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : 600 locutions émanant de Thibault Bazin ont une longueur de 100 caractères ou moins.

Les phrases courtes ne sont donc pas caractéristiques des députés les plus enclins à parler, puisqu'elles constituent la majorité des locutions de l'ensemble des députés (graphique n°6), mais leur particulière abondance chez certains, chez qui elles se comptent

⁸¹ Nous nous concentrons sur les 9 premiers à des fins de visualisation, mais le même constat peut s'étendre à tout le premier quart des locuteurs ayant le plus de locutions.

par centaines, est ce qui leur permet de distinguer ou, à tout le moins, de s'imposer. Cela interroge : qui sont ces députés, à qui s'adressent-ils, ou, à du moins, dans quelles configurations s'expriment-ils ?

2. La parole et l'auditoire : configurations d'interactions

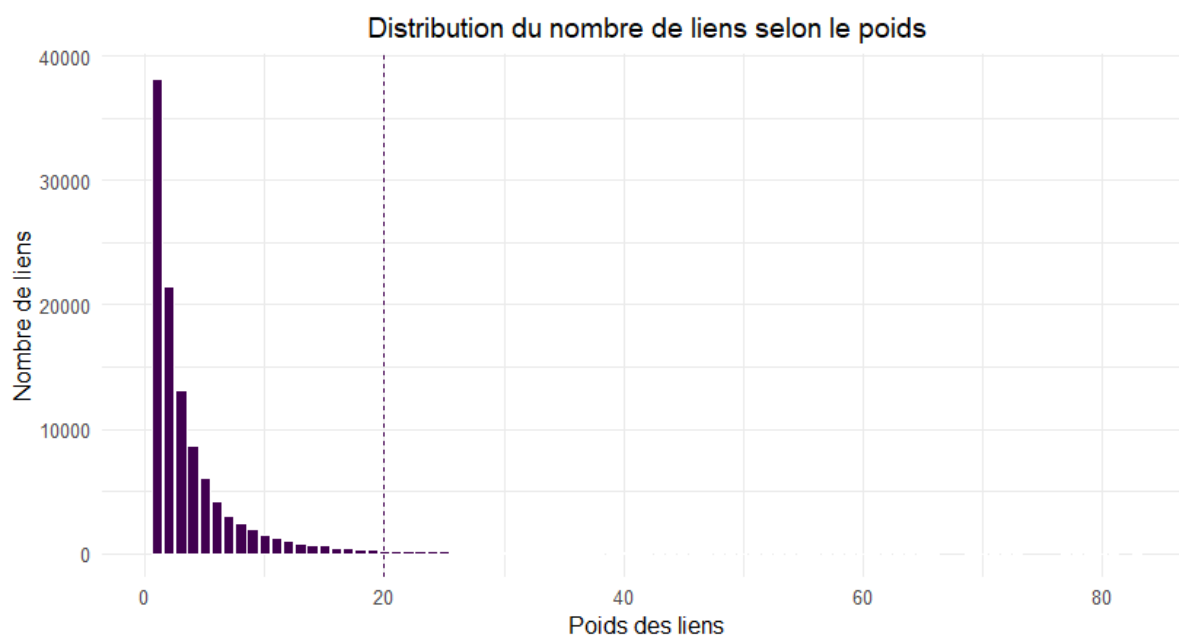
Nous avons expliqué précédemment en quoi la structure de nos données, ainsi que l'organisation même de l'Assemblée Nationale comme institution et comme espace évolutif faisaient obstacle à la reconstitution d'interactions au sens strict. Pour autant, il nous est possible d'observer des configurations de colocutions : en somme, déterminer qui s'exprime devant qui, et par là-même, déterminer qui s'exprime lors de séances de premier plan ou lors de séances moins centrales. Modéliser un réseau de colocutions nous a paru être la meilleure méthode : nous procédons en créant un graphe biparti, c'est-à-dire en reliant chaque locuteur à chaque séance où il intervient. Les locuteurs sont ensuite indirectement liés les uns aux autres s'ils interviennent tous deux lors d'une même séance. Cette méthode permet d'aller au-delà du comptage du nombre ou de la longueur des interventions en les situant dans des configurations où la teneur de l'auditoire et les enjeux de la prise de parole peuvent être approximés. Compte tenu de l'acception large que nous donnons aux liens, et du nombre importants de sommets (autant qu'il n'y a de députés qui parlent)⁸², les liens se démultiplient rapidement : il y en a 109 023⁸³. En ressort un graphe très dense où la majorité des agents sont reliés entre eux (62,1% des liens possibles sont réalisés)⁸⁴. Toutefois, tous ces liens ne sont pas d'égale importance en termes de fréquence d'apparition (graphique n°7) et, par extension, pour notre analyse.

⁸² 593, ce qui est supérieur au nombre de sièges à l'Assemblée car certains députés ont été remplacés (du fait de décès, de nominations à un ministère, etc.).

⁸³ Compte-tenu du fait que les auto-liens, c'est-à-dire les liens qui relient une personne à elle-même lorsqu'elle parle plusieurs fois au cours d'une même séance, sont supprimés.

⁸⁴ La densité du graphe est de 0,6211146.

Graphique n°7.



Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : 37 000 liens ont un poids égal à un : il existe 37 000 paires uniques de locuteurs qui parlent lors d'une seule même séance.

L'étude de la distribution du nombre de liens selon le poids montre en effet que la majorité des liens sont rares : ils ont lieu entre 0 et 10 fois. Or, nous cherchons à observer des régularités, en ce sens il ne serait pas pertinent de conserver, par exemple, les quelque 37 000 liens qui n'ont eu lieu qu'une seule fois. Nous choisissons de fixer le seuil de poids minimal à 20 (représentée par la ligne verticale discontinue dans le graphique n°7), ce qui permet de conserver 2214 liens uniques, chacun d'entre eux pouvant se répéter si une même configuration d'interaction a eu lieu lors de plusieurs séances. Ainsi, nous n'étudions que les liens entre deux personnes qui se sont exprimées au cours d'une même séance au moins 20 fois (c'est-à-dire pour 20 séances), et ce sur un échantillon de 415 séances. Cette sélection, bien qu'elle puisse sembler stricte, n'exige en réalité qu'une expression de deux personnes dans la même séance sur moins de 5% de l'échantillon, ce qui permet de contrebalancer la définition large que nous avons de la colocation.

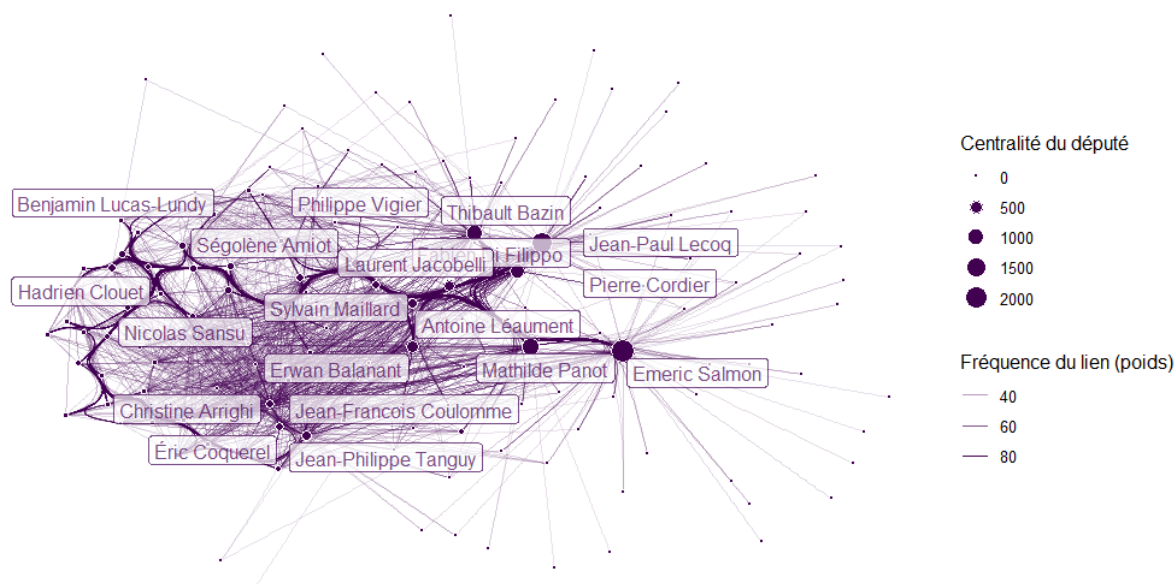
Cette opération de sélection modifie la structure de notre graphe : de 593 sommets il n'en reste plus que 154, sa densité passe de 0.61 à 0.19 (ce qui signifie que 19% des liens possibles existent effectivement), son diamètre de 3 à 92⁸⁵. Nous n'étudions donc plus les personnes qui parlent lors des mêmes séances, mais précisément les personnes qui parlent

⁸⁵ Il s'agit du plus long des plus courts chemins, c'est-à-dire de la longueur du plus court chemin reliant les deux sommets les plus éloignés.

régulièrement lors des mêmes séances, c'est-à-dire les unes devant les autres, voire, les unes aux autres.

Graphique n°8.

Réseau des colocations régulières entre députés



Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : Chaque nœud représente un député, et chaque lien une colocation répétée en séance publique. Plus le lien est opaque, plus la colocation est fréquente. Plus la taille du nœud est importante, plus le député est central.

Cette représentation du réseau (graphique n°8)⁸⁶ montre l'existence de nœuds très connectés où chaque sommet est relié à de nombreux autres sommets et où le poids des liens est particulièrement important (plus de 60 co-locutions), et à l'inverse de nœuds périphériques, plus faiblement connectés entre eux et avec le reste du réseau. Dans le même temps, aucune communauté n'a pu être détectée de façon fiable⁸⁷, ce qui semble indiquer qu'il n'existe pas de "groupes" de députés qui se parlent entre eux ou parlent aux mêmes moments de façon prédominante, mais davantage des députés qui parlent beaucoup dans des séances "prisées" où le nombre de locuteurs est important, et dont l'auditoire est, par

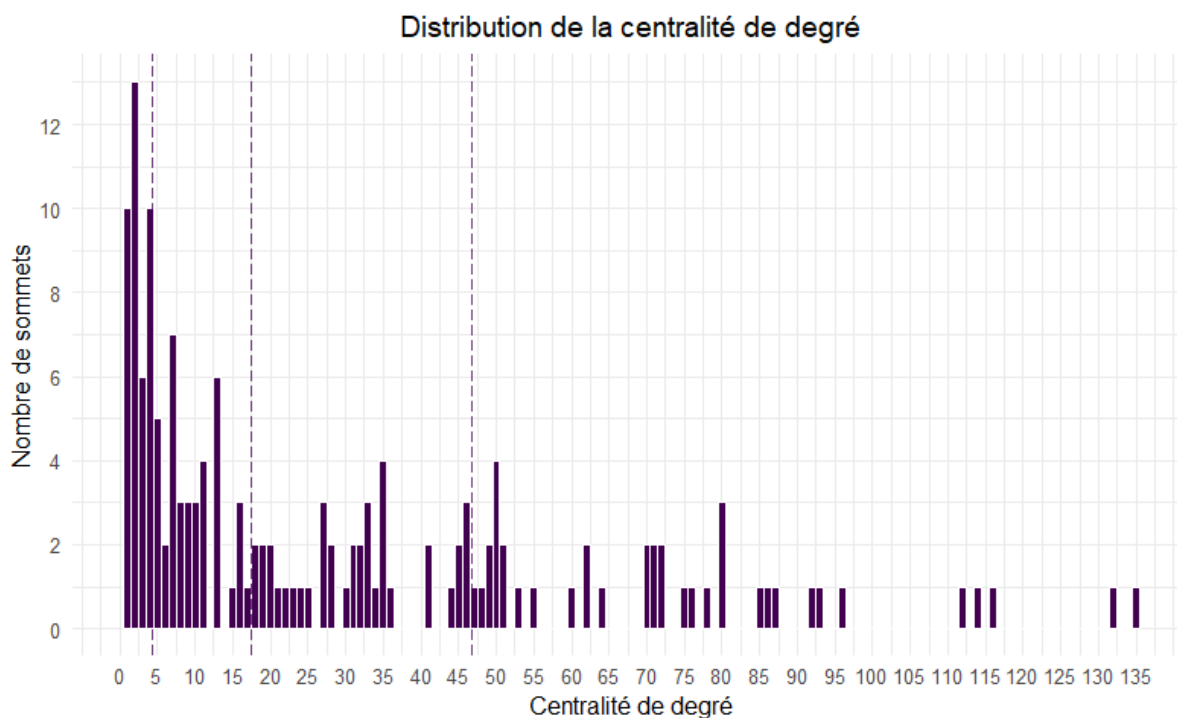
⁸⁶ Pour des questions de lisibilité, nous ne représentons que les labels des 20 locuteurs les plus centraux : la centralité, ou plus précisément dans ce cas-ci la centralité de degré, est un indicateur calculé sur la base du nombre de liens qu'a chaque sommet. Plus un locuteur a de sommets, plus il est central. Il en existe des variantes fondées sur l'intermédialité des sommets, mais cela nous paraissait moins adapté au cas présent puisque les liens sont des colocations et non des formes de relation, le passage d'un sommet à un autre par l'intermédiaire d'un troisième sommet n'aurait donc aucune portée analytique.

⁸⁷ Avec l'algorithme Louvain, nous obtenons 5 clusters et une modularité de 0.047 ; avec une méthode gloutonne (fast greedy), 4 clusters et une modularité de 0.043 ; avec un algorithme basé sur la diffusion entre sommets voisins (label prop), 1 cluster ; avec un algorithme basé sur les vecteurs propres (leading eigen), 4 groupes et une modularité de 0.04. La forte variation du nombre de clusters d'une méthode à l'autre ainsi que la modularité constamment faible laissent à penser que ces communautés seraient difficilement analysables.

extension, probablement plus important.

Cette hiérarchisation s’observe dans la répartition de la centralité de degré parmi les locuteurs (graphique n°9).

Graphique n°9.



En effet, certains locuteurs, très centraux dans notre représentations, concentrent un nombre important de co-locuteurs : le groupe délimité par le dernier quartile, c’est-à-dire les 25% de locuteurs qui ont le plus de co-locuteurs⁸⁸, concentre à lui seul 60% des liens du réseau⁸⁹. En deçà du premier quartile sont regroupés les locuteurs les moins connectés (avec un nombre moyen de liens égal à 2), tandis que le groupe délimité par le dernier quartile rassemble les individus les plus centraux (avec une médiane égale à 71). Cette progression s’accompagne d’une augmentation de la proximité⁹⁰. Autrement dit, seuls 25% des locuteurs de notre réseau s’expriment suffisamment souvent et dans des séances suffisamment prisées ou mouvementées pour se retrouver dans des configurations de colocation régulière (au

⁸⁸ Plus de 47, en l'occurrence. V. annexe, tableau n°1.

⁸⁹ 2894 sur 4542 liens (incluant les doublons) soit 63.7%.

⁹⁰ Elle est calculée en effectuant l’inverse de la somme des distances des chemins les plus courts qui relient un sommet donné à tous autres sommets du réseau. Plus elle se rapproche de 0, plus le sommet est éloigné des autres. Inversement, plus elle se rapproche de 1, plus l’individu étudié est proche des autres membres du réseau.

moins 20 fois) avec plus de 47 personnes⁹¹. Rapporté à la totalité de l'hémicycle, cela ne représente en réalité que 6% des locuteurs⁹². Ces députés centraux sont donc reliés à davantage de personnes, et l'intensité des liens qu'ils entretiennent est également plus élevée, ainsi que le montre le tableau n°1.

Tableau n°1.

Intensité des liens selon la centralité				
	G1 ¹	G2 ¹	G3 ¹	G4 ¹
Intensité des liens	22.3 (21.5, 23.0)	24.1 (23.7, 24.8)	26.3 (25.2, 27.2)	30.3 (28.4, 32.3)

¹ Median (Q1, Q3)

Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Députés du réseau de colocation.

Les groupes de G1 à G4 représentent les groupes délimités par les quartiles de centralité de degré. G1 est le groupe composé des 25% de locuteurs du réseau dont la centralité de degré est la plus faible (de 0 à Q1), et G4 celui dont la centralité est la plus élevée (de Q3 au maximum).

Note de lecture : Les 25% de locuteurs du réseau qui ont le moins de liens uniques ont une intensité médiane de liens égale à 22.3. L'intensité est le nombre de fois qu'un lien se répète, il augmente d'une unité pour chaque séance où la colocation se produit.

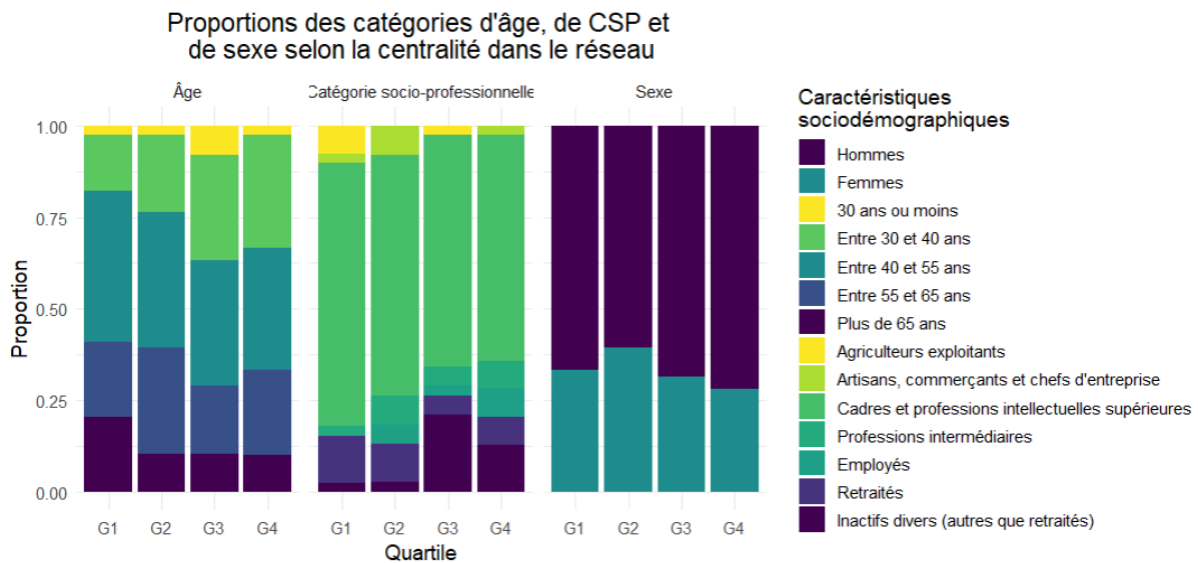
Il s'agit donc de députés qui s'expriment non seulement dans des séances de premier plan où le nombre de locuteurs est élevé, mais qui le font en outre de façon particulièrement fréquente, de sorte que les liens de colocations que nous observons sont de plus en plus réguliers à mesure que le degré de connexion du locuteur augmente. Les positions centrales ne correspondent donc pas seulement à une forte activité oratoire. Elles semblent plutôt associées à la capacité à s'exprimer de façon répétée dans les séances où se concentrent les interactions les plus nombreuses, où l'auditoire est vraisemblablement le plus important et où l'enjeu (symbolique, législatif, politique) est le plus fort. La composition de chacun de ces groupes de centralité n'est pas homogène ; il n'est donc pas possible de déceler un profil caractéristique de chaque position dans cet espace, d'autant que nos variables sont à ce stade limitées. Nonobstant ces limitations, des régularités émergent⁹³ : les femmes se concentrent dans les tranches intermédiaires, et se font de plus en plus rares à mesure que l'on se rapproche du centre (elles représentent 26% du quatrième groupe contre 36% du premier dans le graphique n°10).

⁹¹ V. annexe, tableau n°2.

⁹² Puisque notre réseau comporte 154 sommets, les quartiles délimitent des groupes comprenant 38 ou 39 personnes. Or il existe 577 députés : il en résulte que $39/577 = 0.06$.

⁹³ V. annexe, tableau n°3.

Graphique n°10.



La proportion de personnes âgées de plus de 65 ans diminue à mesure que l'on se rapproche du centre du réseau, passant ainsi de 21% à 11% entre le premier et le dernier groupe ; alors que les députés ayant moins de 30 ans se concentrent dans le troisième groupe et peinent à investir le quatrième, c'est-à-dire le plus central. Ceux ayant entre 30 et 40 ans sont quant à eux plus nombreux dans les deux groupes les plus centraux que dans les deux les plus périphériques. Similairement, le troisième groupe est celui où la proportion d'inactifs (étudiants) non retraités est la plus importante. Cela suggère que l'accès aux positions centrales repose sur des ressources inégalement distribuées entre les acteurs, et que les efforts pour s'en rapprocher peuvent-être plus ou moins bien accueillis et plus ou moins fructueux selon les ressources dont disposent les locuteurs.

Nous savons par exemple que certains groupes politiques ont pour objectif de faire usage de l'Assemblée comme d'une tribune, ainsi que le soulignait E. Ollion⁹⁴, et il est manifeste que cette stratégie politique passe par l'expression orale. Pour autant, cette dimension politique peut conduire à expliquer les différences de position en les renvoyant à tort aux catégories politiques indigènes, or, celles-ci ne sont pas le seul déterminant de la centralité des acteurs, car leurs volontés tribunicienne peuvent se heurter d'une part à leurs

⁹⁴ É. Ollion, *Les candidats*, op. cit., p. 146.

propres dispositions, et d'autre part à la structure du champ qui n'est pas organisé de sorte à favoriser mécaniquement ceux qui disent vouloir être entendus. Les députés insoumis, figures archétypales de « *ce type d'engagement parlementaire* »⁹⁵, sont certes particulièrement présents dans le groupe le plus central si l'on rapporte cela à la taille du groupe⁹⁶, mais ils sont présents, dans les mêmes proportions, dans le groupe le plus périphérique. A l'inverse, les députés écologistes et de la Droite Républicaine voient leur proportion augmenter à mesure que l'on se rapproche du centre, sans pour autant que ces groupes ne soient particulièrement connus pour « *refuser les règles du jeu* »⁹⁷ ou « *chercher à se distinguer* »⁹⁸, pour citer les propos que tenait un haut responsable de l'administration de l'Assemblée au sujet des insoumis au cours d'un entretien avec E. Ollion. Alors, cette notion de l'Assemblée comme tribune semble ne pas être un descripteur des positions effectives qu'occupent les individus dans l'espace social, mais ressemble davantage à une façon de mettre en cohérence la quête de centralité dans un espace avec l'inadaptation (réelle ou perçue) aux normes et attendus de celui-ci. Les gardiens de ces normes (de par leur fonction, à l'instar du responsable, ou de par leur position dominante comme les députés opposés aux insoumis), décrivent alors cela comme un « *refus des règles* », tandis que les locuteurs susceptibles de se voir opposer cette illégitimité, l'investissent et la théorisent comme une stratégie. Cela se confirme lorsque nous rapportons le nombre de mots aux effectifs des groupes politiques. Les députés insoumis sont certes ceux qui ont le plus grand nombre de locutions, pour autant, il sont de ceux qui disent le moins de mots par personne (13 000), nombre qu'excèdent largement 8 groupes parlementaires sur 11, allant de 16 000 à plus de 28 000 mots par personne, sans que cela ne puisse nécessairement être justifié par le statut d'opposition, ainsi qu'en témoignent la position des groupes écologiste ou de la droite républicaine (DR)⁹⁹.

Ainsi, en filigrane derrière le reproche qui peut être fait à certains locuteurs de « trop parler », l'on entrevoit le fait de parler de façon inappropriée, soit que l'on n'y soit pas autorisé (formellement ou socialement), soit qu'on le fasse par des moyens qui déplaisent.

⁹⁵ *Ibid.*

⁹⁶ Voir annexe, graphique n°1.

⁹⁷ *É. Ollion, Les candidats, op. cit., p. 146.*

⁹⁸ *Ibid.*

⁹⁹ Les statistiques se voient légèrement modifiées lorsque l'on considère l'ensemble des députés et non seulement ceux que l'on a conservés dans le réseau, mais le constat demeure : LFI y est le 7ème en termes de nombre de mots par personne. A ce titre, se référer aux tableaux n°4 et 5 en annexe, le premier pour les députés du réseau de colocation et le second pour l'ensemble des députés.

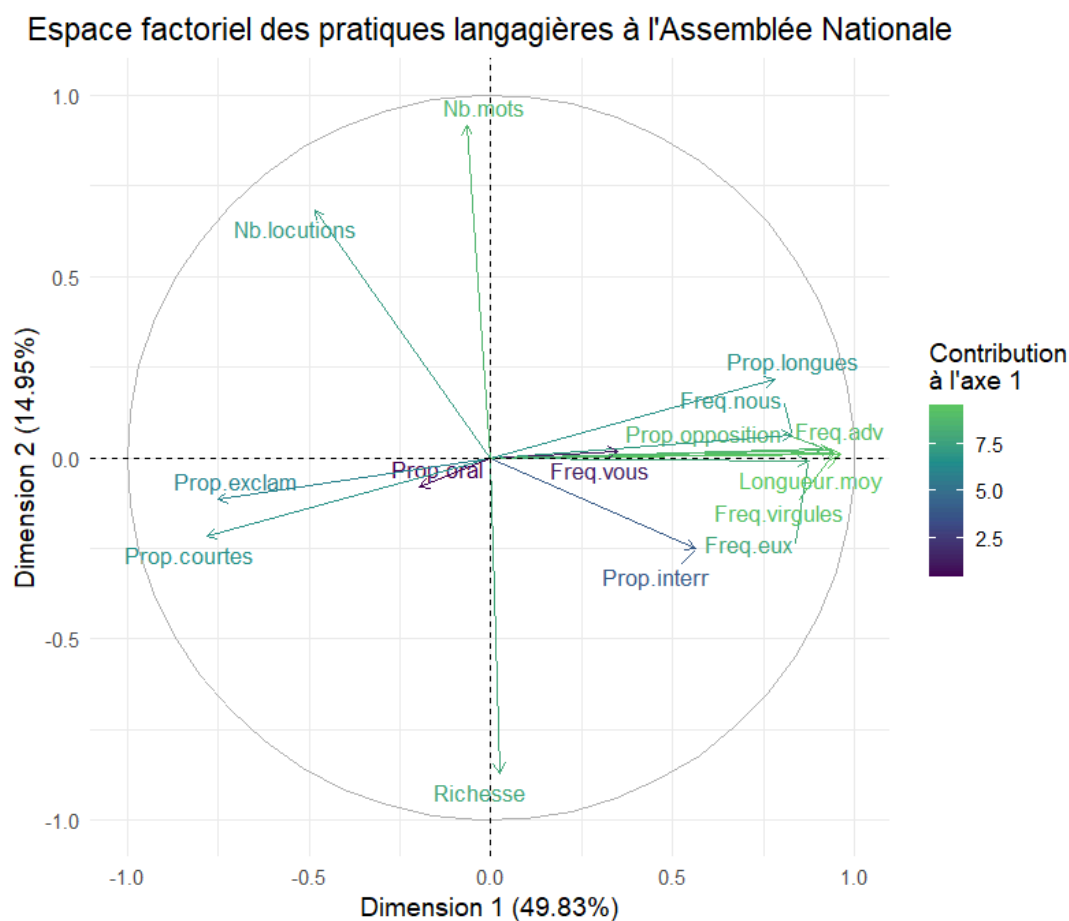
B. Dire : des pratiques langagières et de leur légitimité

Afin de mettre au jour le lien entre la position des acteurs et les caractéristiques de leur expression par le langage, nous nous proposons de construire un espace factoriel à partir de certains indicateurs que nous avons construits sur le langage, puis d'y projeter des variables socio-démographiques d'une part, et des variables relatives à la position dans le réseau d'autre part. Pour ce faire, nous sélectionnons les variables actives suivantes :

- Le nombre de locutions et le nombre total de mots, pour décrire la fréquence des interventions.
- La longueur moyenne des phrases et la proportion de phrases longues (rapportées au total des locutions), autre aspect de la densité des prises de parole.
- La proportion d'exclamations et d'interrogations, afin de capter tant l'expressivité que la finalité des locutions : sont-elles de simples affirmations, des réactions, ou des questions (potentiellement adressées à un interlocuteur) ?
- La fréquence de recours aux pronoms personnels relatifs au "nous", au "vous" et à "eux", ces trois variables visant à déceler le positionnement discursif du locuteur.
- La fréquence des adverbes, des virgules, des connecteurs logiques d'opposition (mais, pourtant, alors que...), des marqueurs d'oralité (ah, oh, euh...), et la richesse du vocabulaire (proportion de mots uniques sur l'ensemble des mots); ce qui peut entre autres déceler le discours formel de l'expression spontanée.

L'espace factoriel alors obtenu est le suivant (graphique n°11).

Graphique n°11.



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

L'on observe d'abord que le premier axe représente à lui seul plus de 49.83% de l'inertie totale, tandis que le second en représente 14.95%. Nous limiterons donc notre analyse à ces deux axes¹⁰⁰, tout en tenant compte du fait que le premier domine la représentation en résumant une grande partie de ce qui distingue nos locuteurs. À droite de notre espace factoriel se trouvent les individus qui prononcent régulièrement des phrases longues, interrogatives, saccadées par des virgules, avec un recours fréquent au “nous”, au “eux”, au “vous”, aux adverbes et aux conjonctions de coordination marquant des oppositions. À gauche, les individus sont caractérisés par une plus faible propension à recourir aux éléments que nous venons de citer, mais ils se distinguent par un nombre important de locutions, avec une forte proportion de phrases courtes, exclamatives, de marqueurs d'oralité, ainsi que par un discours “riche”, entendu ici comme non répétitif. La

¹⁰⁰ V. histogramme des valeurs propres en annexe, graphique n°2.

contribution des variables à la construction de cet axe est très équilibrée (la plus contributive n'y contribue qu'à hauteur de 12%¹⁰¹). Toutefois, certaines variables y sont très bien représentées : il s'agit essentiellement de variables ayant trait à la structure des locutions plutôt qu'à leur contenu. Ce premier axe semble donc être celui de l'intensité et de la densité des prises de parole¹⁰². Le second axe, quant à lui, oppose un discours riche, expressif, fait d'interrogations, d'exclamations et de phrases courtes (en sa partie supérieure), à des formes de locutions plus nombreuses et plus longues¹⁰³, avec davantage d'interpellation de l'interlocuteur par le "vous" (en sa partie inférieure). Cet axe peut donc s'interpréter comme la tonalité de l'interaction, selon qu'elle soit spontanée ou contrôlée.

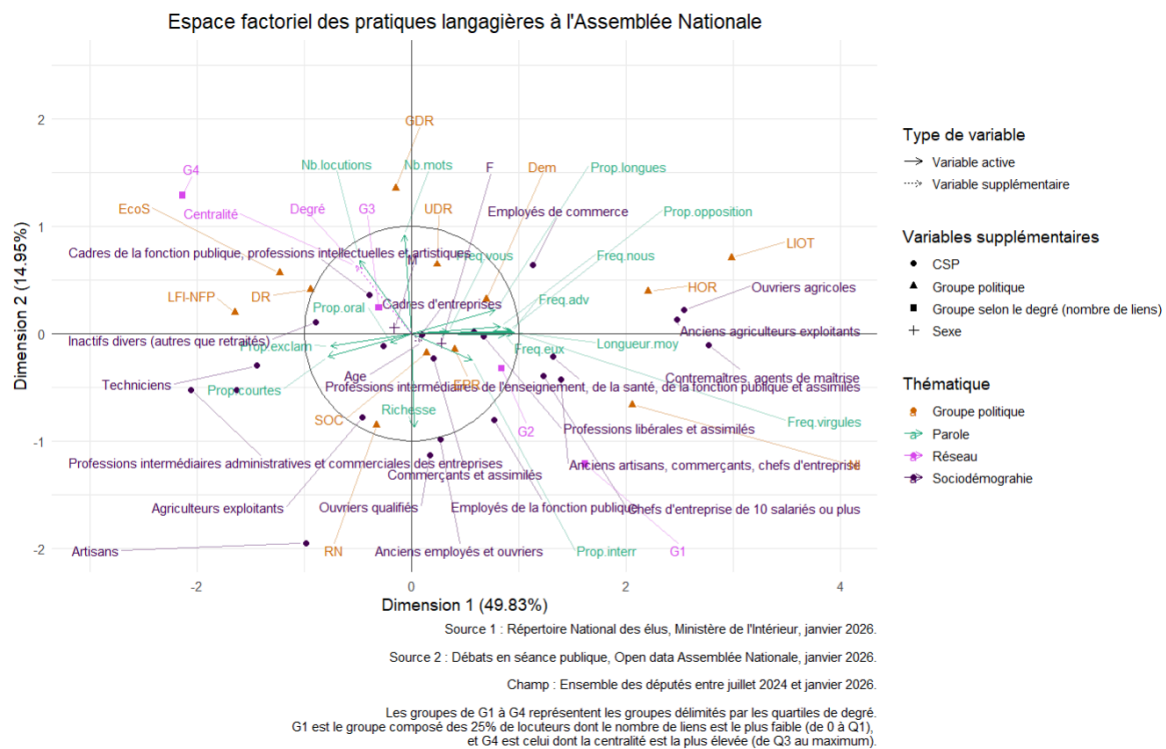
En ajoutant les variables supplémentaires, il ressort que la centralité et le nombre de liens (degré) des locuteurs est certes corrélé positivement au nombre de locutions, mais également que les individus qui concentrent ces caractéristiques se situent dans le quadrant en haut à gauche de l'espace factoriel : ils sont les plus centraux. A l'inverse, les moins centraux se situent en bas à droite de notre espace (graphique n°12).

¹⁰¹ V. annexe, tableau n°6.

¹⁰² Il convient de noter que la structure du graphique demeure similaire lorsque nous passons les variables "Nombre de mots" et "Nombre de locutions" en variables illustratives. La seule modification que cela induit est une meilleure représentation des variables relatives à l'oralité et du recours au "vous". V annexe, graphique n°3.

¹⁰³ Le vecteur représentant le nombre de mots se trouve dans une situation de quasi-superposition avec l'axe, et est dirigé vers le bas. Il est inverse à celui portant sur la richesse, c'est-à-dire sur le nombre de mots uniques, qui croît mécaniquement à mesure que le nombre de mots décroît. Cette opposition résulte en grande partie de la construction des indicateurs mais elle demeure interprétable, en ce qu'un discours contrôlé,

Graphique n°12.

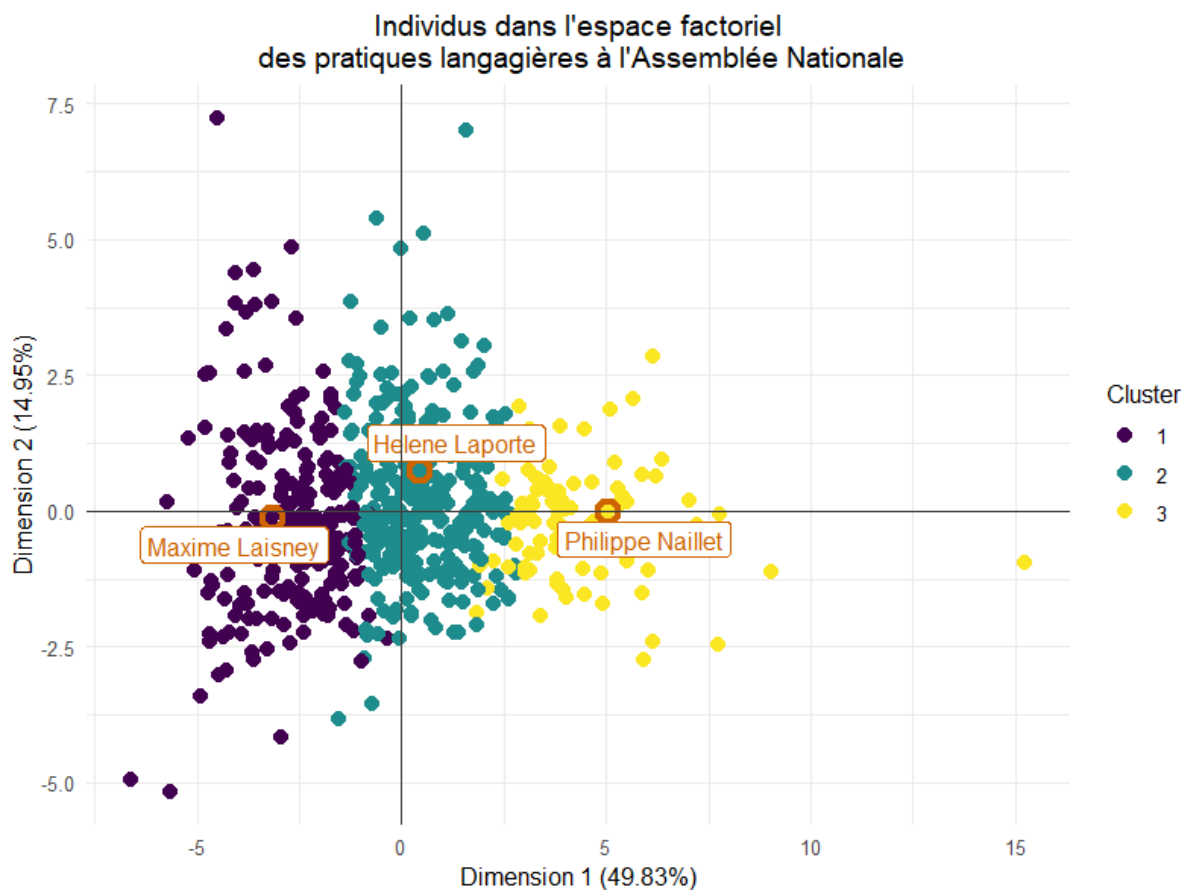


Ces derniers parlent longtemps, utilisent des phrases structurées et rythmées par des adverbes, des virgules et des conjonctions de coordination. Ils adressent des questions à leurs auditeurs, et se désignent souvent. Ces caractéristiques sont souvent associées aux agriculteurs, aux employés de commerces et aux professions libérales, mais également aux personnes plus âgées. Il s'agit là d'une façon de parler dense et contrôlée, sans doute lue ou, *a minima*, préparée en avance. Les groupes parlementaires qui s'en rapprochent sont les députés non-inscrits, rattachés à Horizons, Ensemble pour la République, LIOT, et dans une moindre mesure les Démocrates ainsi que l'Union des Droites pour la République. Dans le quadrant inverse se trouvent les députés les plus centraux, qui parlent plus souvent mais font des phrases plus courtes, plus souvent exclamatives, avec davantage de marqueurs d'oralité. Il s'agit de députés plus jeunes, souvent cadres de la fonction publique, techniciens, étudiants (inactifs non retraités), anciens employés et ouvriers, employés administratifs d'entreprise ou chefs de petites entreprises. Ce style plus spontané est entre autres le propre des Écologistes, de la France insoumise et de la Droite Républicaine. Les cadres d'entreprises et les professions intermédiaires se concentrent quant à eux dans une position intermédiaire, proche des deuxième et troisième groupes : ils parlent moins souvent que ceux du groupe le plus central, mais ont en commun avec eux la spontanéité du style et les interventions courtes et expressives. Cette distribution laisse entrevoir une double opposition : d'une part entre ceux

qui disposent des ressources légitimes (G2, G3 G4) et ceux qui n'en disposent pas (G1), et d'autre part entre ceux qui oeuvrent sciemment à investir l'espace (G4), ceux dont la position y est assise (G2, G3) et ceux qui se conforment aux attendus institutionnels qu'ils en perçoivent (G1), par exemple par un auto-contrôle strict du discours.

Afin de vérifier comment cette opposition se traduit dans l'espace factoriel, nous subdivisons nos locuteurs selon leur proximité par le biais d'une classification ascendante hiérarchique (CAH), c'est-à-dire en rapprochant les individus deux à deux jusqu'à constituer des groupes dont la composition est aussi homogène que possible en termes de distribution des variables, et qui sont aussi différents que possibles les uns des autres (graphique n°13).

Graphique n°13.



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Nous obtenons alors trois groupes qui se répartissent dans l'espace suivant leur position sur le premier axe. Le premier cluster, dont l'expression est fréquente et peu contrôlée, est le plus jeune de tous, avec un âge médian de 52 ans, contre 53 pour le second et

56 pour le troisième¹⁰⁴. Il est également le plus masculin, la proportion d'hommes y étant supérieure à celle du troisième cluster de 8 points de pourcentage. Enfin, s'y concentrent les cadres et professions intellectuelles supérieures, dont les capitaux, en particulier culturel, confèrent vraisemblablement une aisance pour s'exprimer dans cet espace, ainsi que les "inactifs non retraités", qui en représentent 11% contre 5 à 6% dans les autres clusters : il s'agit majoritairement d'anciens étudiants n'ayant jamais eu d'activité professionnelle autre que politique. Nous avons fait ressortir 3 figures archétypales¹⁰⁵ de leur classe respective (graphique n°14). Dans le cadran où l'expression est la plus formelle et les positions les plus périphériques, l'individu le plus proche du centre est Maxime Laisney, professeur des écoles âgé de 45 ans, élu pour la première fois en 2022 après avoir été investi par la France insoumise, dont il est membre depuis sa création (voire avant, puisqu'il avait rejoint le Parti de Gauche dès sa fondation¹⁰⁶). Cette expérience politique et ce rattachement partisan stable se doublent d'un court passage sur la « file d'attente », puisqu'il est élu dès les premières élections auxquelles il se présente, ce qui semble le rapprocher des « jeunes gens pressés » que décrivait E. Ollion, c'est-à-dire ceux dont les aspirations n'ont pas été révisées à la baisse par une attente trop longue avant de pouvoir vivre de la politique. Il a à son actif 146 locutions, dont un tiers n'excède pas les 5 mots. Celles-ci sont fréquemment exclamatives (pour 77% d'entre elles), plus rarement ponctuées de virgules ou d'adverbes (respectivement dans seulement 2% et 1% des cas). Il se distingue à cet égard de Philippe Naillet, figure illustrant le troisième cluster, le moins central. Né en 1960 à la Réunion, il en a été le député de la première circonscription entre 2016 et 2017 puis à partir de 2020. Lorsqu'il se présente aux élections, il se déclare "chef d'entreprise de 10 salariés ou plus", et indique sur le site de l'Assemblée être "inspecteur assurance-vie". Depuis 2008, il est conseiller municipal de la ville qui l'a vu naître, Saint-Denis. Sa première accession à la députation se fait par la suppléance de sa prédécesseuse Ericka Bareigts, issue du parti socialiste, qui est d'ailleurs retournée sur les bancs de l'hémicycle entre 2017 et 2020. Philippe Naillet se rattache ensuite, à compter de 2017, à une structure partisane plus petite fondée en 2017 et dont il est le seul député, nommée "Association Ambition Réunion". En séance publique, il n'est à l'origine que de 19 locutions, dont la longueur moyenne est de 2200 caractères, soit 1700 de plus que Maxime Laisney. Ses prises de parole sont souvent interrogatives et structurées. Quoique moins loquace, avec 75 locutions, Hélène Laporte occupe une position

¹⁰⁴ V. annexe, tableau n°7.

¹⁰⁵ Elles sont en réalité celles qui se rapprochent le plus du centre des groupes (ou clusters) formés par la CAH.

¹⁰⁶ Ces deux structures ayant un fondateur commun, elles demeurent très proches, une multitude d'acteurs circulant entre les deux.

intermédiaire. Sa position dans le champ est stabilisée : elle en maîtrise les codes et dispose d'assises locale et partisane importantes. Cadre d'entreprise, elle rejoint le Rassemblement National¹⁰⁷ en 2014, puis est élue conseillère régionale et députée européenne dans les deux années qui suivent, et accède finalement à la députation en 2022. Elle occupe, depuis 2020, la qualité de vice-présidente du parti. La conjonction de ces éléments lui confère une aisance qui lui permet de prendre la parole pour s'exclamer de façon impromptue : « Soyez honnête, madame Genevard ! », sans être enserrée par le respect des règles ou l'habitus que l'on imagine être celui d'un.e député.e, mais ces interventions ne visent pas non plus à se démarquer. Cela s'observe dans les postes qu'elle occupe à l'Assemblée Nationale : sixième vice-présidente, ses chances de présider la séance sont faibles, mais cela lui confère d'autres avantages : elle participe à l'élaboration de l'ordre du jour, préside une délégation¹⁰⁸, ainsi qu'un groupe d'études et un groupe d'amitié¹⁰⁹.

Ces trois figures donnent à voir une première ébauche de ce que signifie parler selon sa position dans le champ. À la périphérie, la parole est dense, préparée, structurée : elle porte en elle sa propre justification, comme si le locuteur devait convaincre de sa légitimité à chaque intervention. Au centre, elle est fréquente, courte, exclamative — non par impréparation, mais parce que la position recherchée exige de s'imposer et dispense de se justifier longuement. Entre les deux, la maîtrise des codes du champ permet de s'en affranchir ponctuellement, sans que cet affranchissement ne soit perçu comme une transgression. La domination, ici, se lit moins dans ce qu'on dit que dans l'économie des moyens avec lesquels on le dit.

Conclusion

En somme, nous avons cherché, dans le cadre de ce mémoire exploratoire, à traiter les députés non seulement comme des acteurs politiques, mais comme des agents et locuteurs socialement situés, en essayant de saisir dans leurs façons de parler l'expression incorporée de leur position dans le champ parlementaire.

Cela supposait d'abord de prendre la mesure de ce que les données disponibles permettent et ne permettent pas de dire. Les comptes-rendus de séance, produits

¹⁰⁷ Sous les couleurs duquel son grand-père s'était présenté à des élections en 1997, et sa mère est élue municipale.

¹⁰⁸ Celle en charge du patrimoine artistique et culturel : cela implique de rendre des rapports d'information sur ces thématiques, en contrepartie d'un supplément d'indemnité équivalent à 1099 €.

¹⁰⁹ L'ensemble des informations présentées dans cette section sont issues de sources secondaires (articles de presse, sites ou réseaux sociaux des partis, etc).

institutionnels remis en forme *a posteriori*, donnent accès à une version partielle et stabilisée de l'activité langagière des députés : l'oralité, la corporéité, les interactions non publiques leur échappent, ce qui appelle à une certaine prudence dans l'interprétation. C'est pourquoi nous avons fait le choix de travailler sur des indicateurs aussi bruts que possible, construits par expressions régulières, en refusant les opérations de nettoyage qui font abstraction de la forme au profit du fond.

Par la suite, situer socialement ces locuteurs a constitué le second enjeu du mémoire. Le Répertoire national des élus, principale source mobilisée, renseigne moins sur ce que les députés sont que sur ce qu'ils disent être au moment de se présenter à l'élection : en particulier, la CSP déclarée est une mise en récit biographique autant qu'une donnée sociologique, et doit être interprétée comme telle. Cette source de données pourra par la suite être mise en relation avec d'autres indicateurs plus riches, notamment d'ancienneté dans le champ politique.

Enfin, la mise en relation de ces indicateurs langagiers et sociodémographiques a permis de dégager une première ébauche de ce que signifie parler selon sa position dans le champ : à la périphérie, une parole dense et contrôlée qui porte en sa forme sa propre justification ; au centre, une expression plus fréquente, plus courte, plus exclamative, plus « orale », dont on peut faire l'hypothèse qu'elle cherche à s'imposer en faisant montre d'une plus grande familiarité avec le champ. Entre ces deux positions, une parole ponctuelle qui maîtrise les codes et sait s'en dispenser quand son émetteur estime qu'il le faut.

Ces résultats exploratoires, s'ils confirment que les pratiques langagières ne sont jamais tout à fait indifférentes à la place occupée dans l'espace social, appellent toutefois à être approfondis, notamment par l'intégration de variables relatives à la trajectoire biographique et au degré de professionnalisation politique. Une autre voie à explorer consisterait à gagner en précision dans l'étude des pratiques langagières : d'abord, celle-ci peut être enrichie par l'ajout de variables, en outre elle se résume pour l'instant à une étude des variations interpersonnelles. Elle pourrait à cet égard être approfondie en se concentrant sur les variations intrapersonnelles, c'est-à-dire sur les différentes façons qu'a une même personne de parler (ou de ne pas parler) dans différents contextes, c'est-à-dire selon la thématique abordée, la taille de l'auditoire, etc, ce qui pourrait donner à voir les rapports qu'entretiennent les locuteurs à la parole.

Bibliographie

ACHARD Pierre, *La sociologie du langage*, Paris, Presses universitaires de France (coll. « Que sais-je ? »), 1993, 127 p.

ANGERMULLER Johannes et GLADY Marc, « La sociologie du langage. Perspectives d'un champ émergent », *Langage et société*, 24 mai 2017, vol. 160161, n° 2, p. 163-178.

BAUDELLOT Christian et ESTABLET Roger, « 2. Que valent les statistiques du suicide ? » dans *Durkheim et le suicide*, Paris cedex 14, Presses Universitaires de France (coll. « Philosophies »), 2011, vol.8e éd., p. 46-75.

BENOIT Kenneth, WATANABE Kohei, WANG Haiyan, NULTY Paul, OBENG Adam, MÜLLER Stefan et MATSUO Akitaka, « Quanteda: An R package for the quantitative analysis of textual data », *Journal of Open Source Software*, 6 octobre 2018, vol. 3, n° 30, p. 774.

BERNSTEIN Basil, *Langage et classes sociales: codes socio-linguistiques et contrôle social*, Paris, Éd. de Minuit (coll. « Le sens commun »), 1975, 347 p.

BOLTANSKI Luc, DARRÉ Yann et SCHILTZ Marie-Ange, « La dénonciation », *Actes de la recherche en sciences sociales*, 1984, p. 3-40.

BOLTANSKI Luc et THÉVENOT Laurent, « Comment s'orienter dans le monde social », *Sociologie*, 31 mars 2015, Vol. 6, n° 1, p. 5-30.

BOURDIEU Pierre, *Langage et pouvoir symbolique*, Paris, Éditions Fayard (coll. « Points Essais »), 2001, 423 p.

BOURDIEU Pierre, *Choses dites*, Paris, Éd. de Minuit (coll. « Le sens commun »), 1987, 229 p.

BOURDIEU Pierre, « La force du droit », *Actes de la recherche en sciences sociales*, 1986, n° 64, p. 3-19.

BOURDIEU Pierre, « L'illusion biographique », *Actes de la recherche en sciences sociales*, 1986.

BOURDIEU Pierre, *Ce que parler veut dire: l'économie des échanges linguistiques*, Paris, Fayard, 1982, 244 p.

BOURDIEU Pierre, « Décrire et prescrire », *Actes de la recherche en sciences sociales*, 1981.

BOURDIEU Pierre, CHAMBOREDON Jean-Claude et PASSERON Jean-Claude, *Le métier de sociologue: préalables épistémologiques*, Nouvelle éd., Paris, Éditions EHESS (coll. « EHESS poche »), 2021, 576 p.

DESROSIÈRES Alain, GOY Alain et THÉVENOT Laurent, « L'identité sociale dans le travail statistique : la nouvelle nomenclature des professions et catégories socioprofessionnelles », *Economie et statistique*, 1983, vol. 152, n° 1, p. 55-81.

FARGE Arlette, *Le goût de l'archive*, Paris, Éd. du Seuil (coll. « Points Histoire »), 1998, 152 p.

GAÏTI Brigitte et COLLOVALD Annie, « Discours sous surveillances : Le social à l'Assemblée » dans *Le « social » transfiguré. Sur la représentation politique des préoccupations « sociales »*, Presses universitaires de France., s.l., 1990, p. 9-54.

GALEMBERT Claire de, ROZENBERG Olivier et VIGOUR Cécile, « Chapitre 7. Réflexivité, rôles parlementaires et discours de légitimation. Les apports du logiciel Alceste à l'analyse lexicale d'un débat parlementaire relatif aux rapports entre justice et politique » dans *Faire parler le parlement: méthodes et enjeux de l'analyse des débats parlementaires pour les sciences sociales*, Paris, LGDJ-Lextenso (coll. « Droit et société »), 2013, p. 127-149.

GOFFMAN Erving, *Façons de parler*, Paris, Ed. de Minuit (coll. « Le sens commun »), 1987, 285 p.

HEURTIN Jean-Philippe, « Architectures morales de l'Assemblée nationale », *Politix, Revue des sciences sociales du politique*, 1994, p. 109-140.

JAVEAU Claude, « Sociologies du langage », *Raisons politiques*, 2001, vol. 2, n° 2, p. 79-87.

LABOV William, *The social stratification of English in New York City*, 2nd ed., s.l., Cambridge University Press, 1966, 498 p.

LAHIRE Bernard, « Achard Pierre, La sociologie du langage. », *Revue française de sociologie*, 1994, n° 35-1, p. 144-146.

LAVOINNE Yves, « Publicité des débats et espace public », *Études de communication. langages, information, médiations*, 1999, n° 22, p. 115-132.

MAGRI Véronique, « La linguistique et le nombre », *Le Français moderne - Revue de linguistique française*, 2020, p. 9.

OLLION Étienne, *Les candidats: novices et professionnels en politique*, Paris, Presses universitaires de France, 2021, 304 p.

PINTO Louis, « La communication orale en milieu académique – Note programmatique », *Regards sociologiques*, 2024, n° 63, p. 11-25.

Sources

ANONYME, « De la publicité des débats parlementaires », 1846.

ASSEMBLÉE NATIONALE, « Liste des députés en exercice », 2025 et 2026.

INSEE, « Portraits des circonscriptions législatives », 2022.

MINISTÈRE DE L'INTÉRIEUR, « Répertoire national des élus », septembre 2025 et janvier 2026.

Règlement de l'Assemblée Nationale, 2025, 336 p.

« Décret n°2001-777 du 30 août 2001 portant création au ministère de l'intérieur d'un fichier des élus et des candidats aux élections au suffrage universel ».

« Constitution du 4 octobre 1958 ».

Fiche n°30 - Les comptes rendus,

<https://www.assemblee-nationale.fr/dyn/synthese/organisation-assemblee-nationale/les-comptes-rendus>, consulté le 25 mai 2026.

Annexes

Tableau n°1.

Concentration des liens selon le nombre de colocuteurs		
Groupe de degré (nombre de colocuteurs)	Poids total des liens	Part du total (%)
G1	38771	4.418126
G2	106984	12.191298
G3	204960	23.356094
G4	526829	60.034483

Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.
Champ : Ensemble des députés entre juillet 2024 et janvier 2026.
 Les groupes de G1 à G4 représentent les groupes délimités par les quartiles de degré. G1 est le groupe composé des 25% de locuteurs du réseau dont le degré (nombre de liens) est le plus faible (de 0 à Q1), et G4 celui dont le degré est le plus élevé (de Q3 au maximum).
Note de lecture : Le quart des députés ayant le moins de colocuteurs concentre 4.42% des liens totaux.

Tableau n°2.

Distribution du degré (nombre de liens uniques) des locuteurs		
1er quartile	Médiane	3e quartile
4.25	17.5	46.75

Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.
Champ : Ensemble des députés entre juillet 2024 et janvier 2026.
Note de lecture : Un quart des députés ont interagi 20 fois dans la même séance que 5 personnes ou moins.

Tableau n°3.

Distributions des caractéristiques socio-démographiques selon le groupe de centralité				
Variables	G1 ¹	G2 ¹	G3 ¹	G4 ¹
Sexe				
Femme	33%	39%	32%	28%
Homme	67%	61%	68%	72%
Age				
30 ans ou moins	3%	3%	8%	3%
Entre 30 et 40 ans	15%	21%	29%	31%
Entre 40 et 55 ans	41%	37%	34%	33%
Entre 55 et 65 ans	21%	29%	18%	23%
Plus de 65 ans	21%	11%	11%	10%

Catégorie socio-professionnelle

Agriculteurs exploitants	8%	0%	3%	0%
Artisans, commerçants et chefs d'entreprise	3%	8%	0%	3%
Cadres et professions intellectuelles supérieures	72%	66%	63%	62%
Employés	0%	5%	3%	8%
Inactifs divers (autres que retraités)	3%	3%	21%	13%
Professions intermédiaires	3%	8%	5%	8%
Retraités	13%	11%	5%	8%

Groupe politique (abrégé)

Dem	8%	8%	5%	5%
DR	13%	11%	18%	18%
EcoS	8%	11%	16%	18%
EPR	28%	3%	8%	5%
GDR	3%	3%	5%	8%
HOR	5%	3%	3%	0%
LFI-NFP	21%	32%	18%	26%
LIOT	5%	3%	0%	0%
RN	5%	18%	11%	13%
SOC	3%	3%	11%	8%
UDR	3%	8%	5%	0%

¹ %

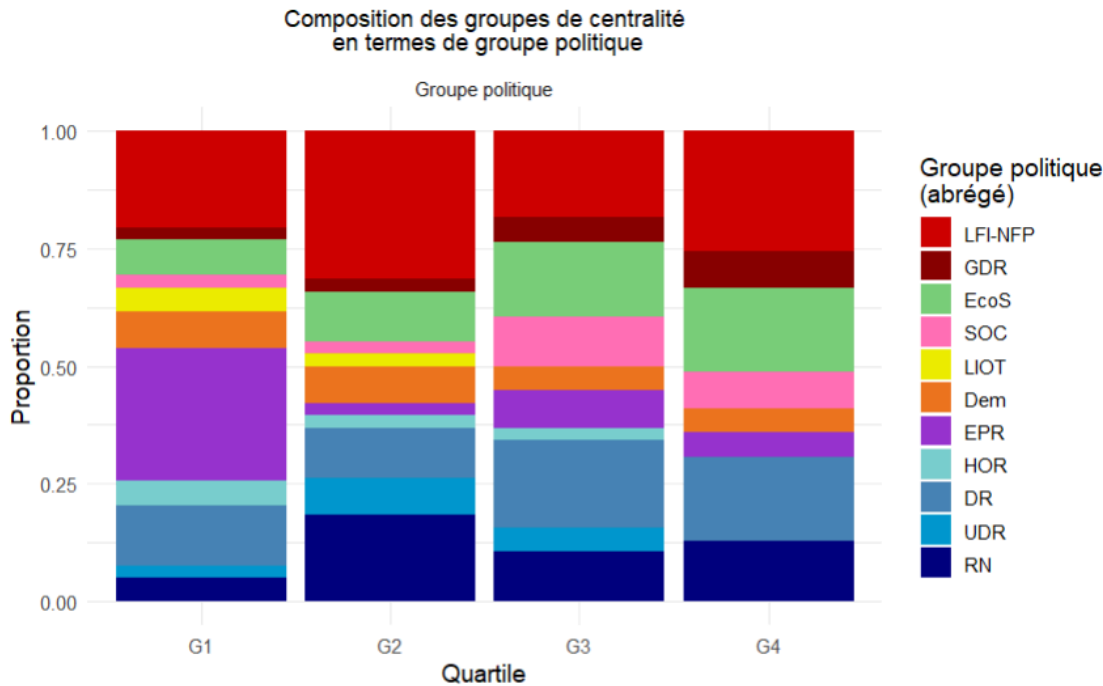
Source 1 : Répertoire National des élus, Ministère de l'Intérieur, janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Députés appartenant au réseau de colocation

Note de lecture : Les groupes de G1 à G4 représentent les groupes délimités par les quartiles de centralité G1 est le groupe composé des 25% de locuteurs du réseau dont la centralité de degré est la plus faible (de 0 à Q1), et G4 celui dont la centralité est la plus élevée (de Q3 au maximum). Les députés âgés de plus de 65 ans représentent 21% du groupe le moins central du réseau de colocation.

Graphique n°1.



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Députés du réseau de colocation.

Note de lecture : Les groupes de G1 à G4 représentent les groupes délimités par les quartiles de centralité de degré. G1 est le groupe composé des 25% de locuteurs du réseau dont la centralité de degré est la plus faible (de 0 à Q1), et G4 celui dont la centralité est la plus élevée (de Q3 au maximum). Les députés LFI représentent 21% du groupe le moins central du réseau de colocation.

Tableau n°4.

Statistiques de parole par groupe politique				
Groupe politique (abrégé)	Nombre de membres	Nombre de locutions	Nombre de mots	Nombre de mots par personne
LIOT	3	325	85664	28555
GDR	7	2142	192474	27496
UDR	6	1097	140035	23339
HOR	4	555	89963	22491
EPR	17	3668	311935	18349
Dem	10	1868	166049	16605
DR	23	6123	361244	15706
EcoS	20	4874	306861	15343
LFI-NFP	37	8792	497046	13434
SOC	9	1970	115857	12873
RN	18	4342	175867	9770

Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Députés du réseau de colocation.

Note de lecture : Les 3 députés du groupes Libertés, Indépendants, Outre-mer et Territoires figurant dans notre réseau de colocations régulières sont à l'origine de 325 locutions.

Tableau n°5.

Statistiques de parole par groupe politique				
Groupe politique (abrégé)	Nombre de membres	Nombre de locutions	Nombre de mots	Nombre de mots par personne
GDR	17	2722	316921	18642
LIOT	25	1301	326808	13072
UDR	18	1641	223747	12430
EcoS	38	5826	434639	11438
DR	56	7780	581319	10381
Dem	37	2905	364997	9865
LFI-NFP	71	10226	670758	9447
HOR	37	1950	348018	9406
EPR	104	6636	712366	6850
SOC	66	4181	441262	6686
NI	12	276	60434	5036
RN	123	7028	496312	4035

Source : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : Les 17 députés du groupes de la Gauche Démocrate et Républicaine sont à l'origine de 2722 locutions.

Graphique n°2.

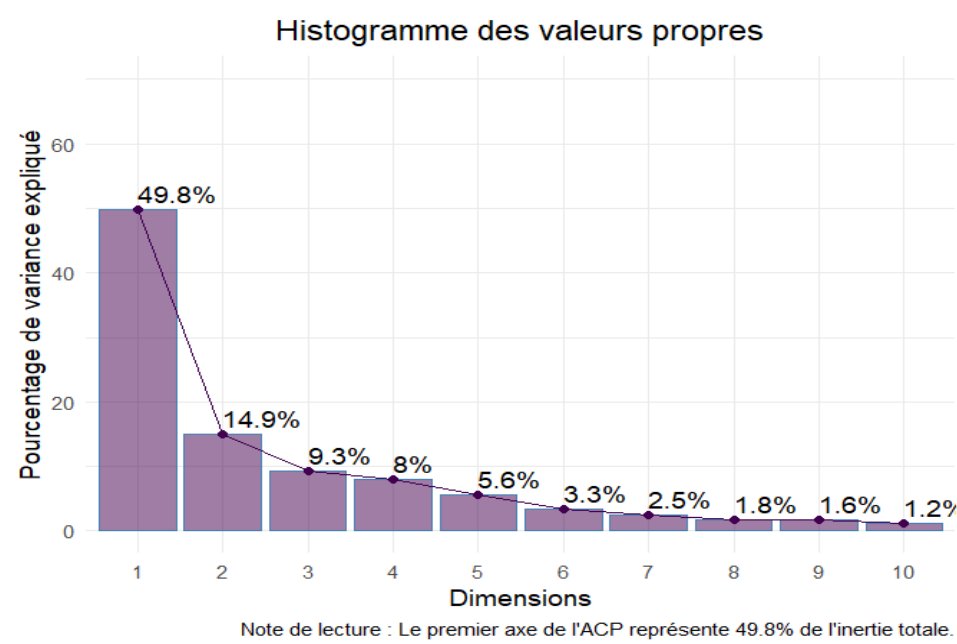
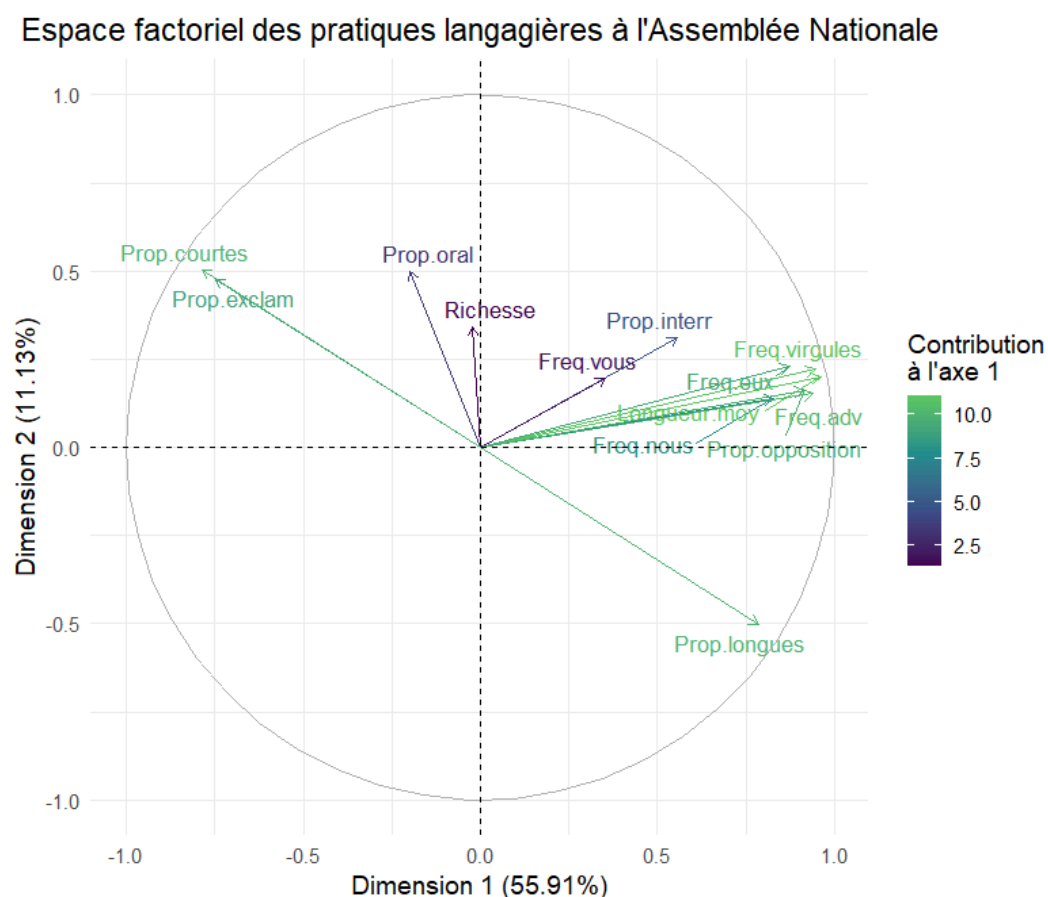


Tableau n°6. Coordonnées, contribution, représentation et corrélation des 10 variables les plus contributives à l'axe 1.

Variable	Coord	Contrib	Cos2	Cor
Longueur.moy	0.961	12.35	0.923	0.961
Freq.virgules	0.944	11.93	0.891	0.944
Freq.adv	0.938	11.76	0.879	0.938
Prop.opposition	0.914	11.17	0.835	0.914
Freq.eux	0.873	10.20	0.762	0.873
Freq.nous	0.827	9.16	0.685	0.827
Prop.courtes	-0.781	8.16	0.610	-0.781
Prop.longues	0.781	8.16	0.610	0.781
Prop.exclam	-0.751	7.55	0.564	-0.751
Prop.interr	0.562	4.23	0.316	0.562

Graphique n°3. Analyse en composantes principales alternative à celle présentée, excluant le nombre de locutions et le nombre de mots : la position des variables reste relativement stable en comparaison avec l'ACP principale.



Source 1 : Répertoire National des élus, Ministère de l'Intérieur, septembre 2025 et janvier 2026.

Source 2 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Tableau n°7.

Distribution des caractéristiques socio-démographiques et partisans selon le cluster			
Caractéristiques	1 N = 215 ^J	2 N = 278 ^J	3 N = 100 ^J
Age	52 (40, 60)	53 (41, 61)	56 (50, 62)
Sexe			
F	33% (72)	37% (103)	41% (41)
M	67% (143)	63% (175)	59% (59)
CSP			
Agriculteurs exploitants	1.9% (4)	2.9% (8)	2.0% (2)
Artisans, commerçants et chefs d'entreprise	2.8% (6)	6.8% (19)	9.0% (9)
Cadres et professions intellectuelles supérieures	61% (131)	62% (172)	53% (53)
Employés	4.7% (10)	4.3% (12)	5.0% (5)
Inactifs divers (autres que retraités)	11% (23)	5.4% (15)	6.0% (6)
Ouvriers	0.5% (1)	1.1% (3)	0% (0)
Professions intermédiaires	8.4% (18)	6.8% (19)	13% (13)
Retraités	10% (22)	11% (30)	12% (12)
Groupe politique (abrégé)			
Dem	4.2% (9)	6.5% (18)	10% (10)
DR	12% (25)	9.7% (27)	4.0% (4)
EcoS	10% (22)	5.0% (14)	2.0% (2)
EPR	13% (28)	19% (52)	19% (19)
GDR	3.3% (7)	2.5% (7)	3.0% (3)
HOR	1.9% (4)	7.2% (20)	13% (13)
LFI-NFP	22% (48)	7.9% (22)	1.0% (1)
LIOT	0.5% (1)	3.6% (10)	14% (14)
NI	0.5% (1)	2.2% (6)	4.0% (4)
RN	20% (42)	23% (63)	14% (14)
SOC	12% (26)	9.4% (26)	14% (14)
UDR	0.9% (2)	4.7% (13)	2.0% (2)

^J Median (Q1, Q3); % (n)

Source 1 : Débats en séance publique, Open data Assemblée Nationale, janvier 2026.

Source 2 : Répertoire national des élus, Ministère de l'intérieur, septembre 2025 et janvier 2026

Champ : Ensemble des députés entre juillet 2024 et janvier 2026.

Note de lecture : L'âge médian du premier cluster est de 52 ans.