

FICHE DE PROJET : “Parler de soi à travers la machine : l’auto comme miroir de l’identité”

1. Contexte et justification

Depuis ses débuts, la publicité automobile ne vend pas seulement un moyen de transport : elle vend une image de soi, une façon d'être au monde. La voiture est une **machine symbolique** qui incarne le désir de liberté, le statut social, mais aussi, de plus en plus, une forme de responsabilité morale face aux enjeux écologiques.

Une première observation exploratoire de notre corpus (voir *Figure 1*) permet de saisir cette nature profondément humaine du discours publicitaire.



Figure 1 : Nuage de mots des publicités automobiles (2014-2024)

Comme le révèle ce panorama lexical, le vocabulaire de la machine s'efface derrière celui du sujet. Les termes dominants (« **Homme** », « **Femme** », « **Voix** ») et les verbes d'action (« **Partir** », « **Profitez** ») montrent que la publicité est avant tout un récit d'incarnation et d'évasion, tiraillé par des réalités matérielles (« **Euros** », « **Mois** »).

Cependant, dans le contexte actuel de la transition énergétique, ce récit traditionnel est mis à l'épreuve. Les marques semblent modifier leur manière de s'adresser au public : elles ne parlent plus simplement de performances techniques, mais convoquent des valeurs, des styles de vie et une nouvelle subjectivité.

Ce projet propose d'analyser cette mutation.

Il s'agira de comprendre comment la publicité automobile construit la figure du conducteur ou de la conductrice : est-il défini comme un sujet émotionnel (le plaisir), rationnel (l'économie) ou moral (la responsabilité) ? En articulant la machine et l'identité, nous chercherons à voir si le miroir publicitaire est en train de changer de reflet.

2. Problématique

Comment la publicité automobile met-elle en scène le **conducteur ou la conductrice comme miroir identitaire**, et comment cette représentation évolue-t-elle entre 2014 et 2024 à travers les registres narratifs et moraux ?

Sous-questions :

- Comment les marques s'adressent-elles à leurs publics (je / tu / nous) ?
- Quels styles de discours traduisent des **formes de subjectivité** différentes (individuelle, collective, morale) ?
- Quelles marques ou quelles périodes privilégient une identité *libre, performante, responsable ou connectée* ?
- Comment les visuels et les scripts collaborent-ils pour représenter le rapport entre **l'humain et la machine** ?

3. Hypothèses

1. **H1 – La publicité automobile est un miroir de la subjectivité contemporaine.**
→ Elle met en scène l'individu à travers la machine, comme un prolongement de soi (identité technologique et affective).
2. **H2 – Le style d'adresse publicitaire traduit une évolution du "je" vers le "nous".**
→ Avant 2018 : individualisme et liberté ;
→ Après 2020 : responsabilité, collectif, écologie.

4. Objectifs

- **Identifier les formes de subjectivation** (rationnelle, émotionnelle, morale) à travers des scores lexicaux pour mesurer "qui" la publicité met en scène.
- **Mettre en évidence la rupture chronologique** de l'identité publicitaire, en datant précisément le basculement entre l'ascension de l'ego (jusqu'en 2021) et son effacement soudain.
- **Cartographier l'espace sémantique (BERT)** pour démontrer que le passage du "Je" au "Nous" n'est pas une opposition binaire, mais une hybridation où la performance tente de cohabiter avec l'éthique.
- **Analyser la recomposition thématique (LDA)** liée à l'électromobilité, pour voir comment le récit électrique remplace le "plaisir d'achat" par un "sens d'agir".
- **Caractériser la mutation du miroir publicitaire**, passant de la vente d'un statut social individuel à la promotion d'une responsabilité collective durable.

5. Corpus et variables

Corpus :

- 8 938 publicités du dataset automobile (2014–2024).
- Variables exploitées :
 - **year, Marque, Script, Visuel, MotsCles, Secteur, Electric, Hybrid, Support, Format, Incrustation.**

Variables dérivées :

- **Script_clean, Visuel_clean, MotsCles_clean** (textes nettoyés).
- **score_identite** : fréquence des lexiques de subjectivité.
- **score_moral** : termes liés à la responsabilité, à la temporalité future et à la préservation du bien commun. (styles d'adresse).
- **score_émotionnel** : champ lexical des sensations, des sentiments et de l'expérience vécue. (styles d'adresse).
- **score_rationnel** : vocabulaire de l'ingénierie, des caractéristiques physiques et des métriques chiffrées. (styles d'adresse).
- **Energie, periode** : variables de catégorisation.

6. Méthodologie

Prétraitement:

- Nettoyage du texte et normalisation des colonnes.
- Sélection du secteur automobile.
- **création des registres narratifs = styles d'adresse**
→ rationnel (démonstratif) / émotionnel (empathique) / moral (prescriptif).

Création des registres:

score moral: Ce score mesure l'intensité du discours axiologique (portant sur les valeurs). Il reflète la prise en compte des enjeux sociétaux et environnementaux par la marque.

- **Logique de calcul :** Il détecte les termes liés à la responsabilité, à la temporalité future et à la préservation du bien commun.
- **Mots-clés ciblés :** *Planète, durable, responsable, avenir, ensemble, préserver, éthique, transition, environnement, demain, collectif.*

score émotionnel: Ce score quantifie l'appel aux sens et à l'affect. Contrairement au score identitaire (qui définit *qui je suis*), le score émotionnel décrit *ce que je ressens*.

- **Logique de calcul :** Il identifie le champ lexical des sensations, des sentiments et de l'expérience vécue.
- **Mots-clés ciblés :** *Plaisir, frisson, amour, joie, rêve, silence, sensation, harmonie, aventure, douceur, souvenir.*

score rationnel: Ce score évalue la dimension technique, fonctionnelle et factuelle du discours. Il correspond à l'argumentaire de vente traditionnel, fondé sur la preuve et la performance tangible.

- **Logique de calcul :** Il recense le vocabulaire de l'ingénierie, des caractéristiques physiques et des métriques chiffrées.
- **Mots-clés ciblés :** *Technologie, moteur, performance, watts, sécurité, consommation, système, autonomie, précision, fiabilité, contrôle.*

7. Représentations graphiques

Ce score mesure la place accordée à l'individu et à l'affirmation de soi dans le message. Il ne s'agit pas de décrire la voiture, mais de définir celui ou celle qui la conduit.

- **Logique de calcul :** Il comptabilise les pronoms personnels (singularité et collectif) ainsi que les termes liés à la projection de soi.
- Création d'un **lexique identitaire** : ["je", "tu", "vous", "nous", "moi", "mon", "ma", "notre", "liberte", "plaisir", "vie", "famille", "style", "confort", "audace", "caractere", "envie", "quotidien", "partager", "ensemble"]

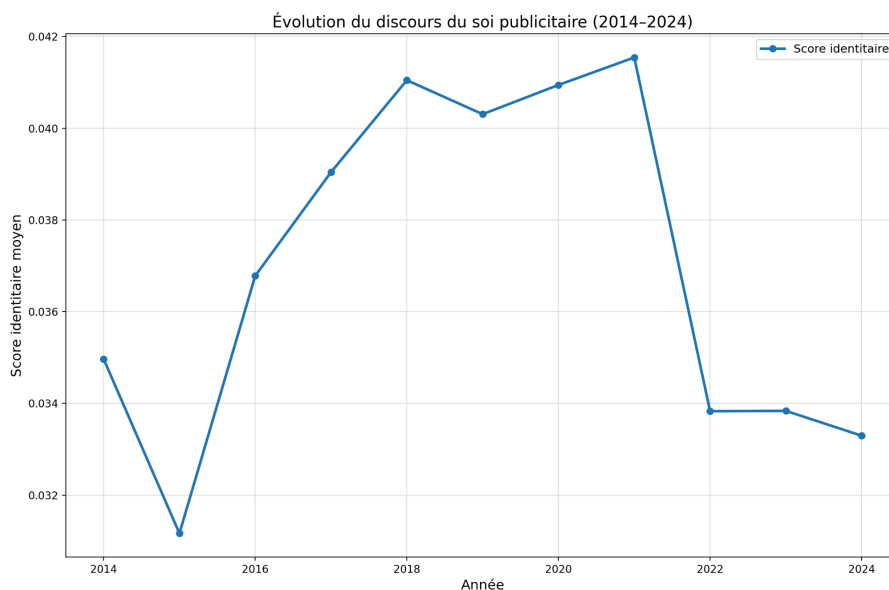


Figure 2 : Graphique de l'évolution du "Score identitaire moyen" (2014-2024)

Ce graphique illustre l'évolution du "Score identitaire moyen" dans le discours publicitaire entre 2014 et 2024.

Voici les faits observables :

1. **Une tendance à la hausse (2015-2021)** : Après un point bas en 2015, l'utilisation de ce discours augmente régulièrement pour atteindre son maximum en 2021.
2. **Une chute brutale (2021-2022)** : En l'espace d'un an, le score s'effondre, perdant tout le gain des années précédentes. Le vocabulaire de la responsabilité, du partage et de la planète augmente sensiblement
3. **Une stagnation à un niveau bas (2022-2024)** : Depuis 2022, la courbe reste plate et ne montre aucun signe de remontée, se situant à un niveau inférieur à celui de 2014. (signe que le nouveau régime discursif est durable)

En résumé, le graphique montre une montée en puissance de cette métrique pendant 6 ans, suivie d'un arrêt soudain et durable à partir de 2022.

Évolution des registres narratifs

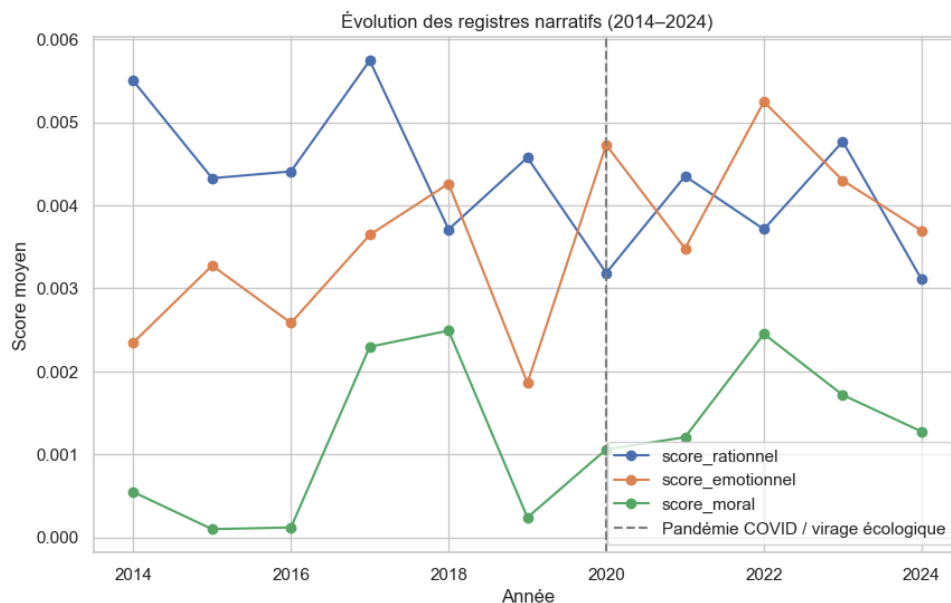


Figure 3 : Graphique de l'évolution des registres narratifs (2014-2024)

Cette courbe temporelle synthétise l'évolution moyenne des trois registres lexicaux entre 2014 et 2024. Trois dynamiques se distinguent clairement :

- **Le registre rationnel** reste dominant, mais tend à s'éroder légèrement.
- **Le registre émotionnel** s'intensifie après 2018, marquant une recherche de proximité et de sensibilité.
- **Le registre moral**, quasi inexistant avant 2017, connaît une ascension rapide à partir de 2019, avec un pic autour de 2022.

Le point de rupture de 2020, signalé par la ligne verticale (pandémie et accélération de la transition écologique), marque le véritable tournant discursif : l'automobile cesse d'être seulement un objet technique pour devenir un sujet social et éthique.

Carte sémantique BERT (UMAP) colorée par la probabilité d'ego

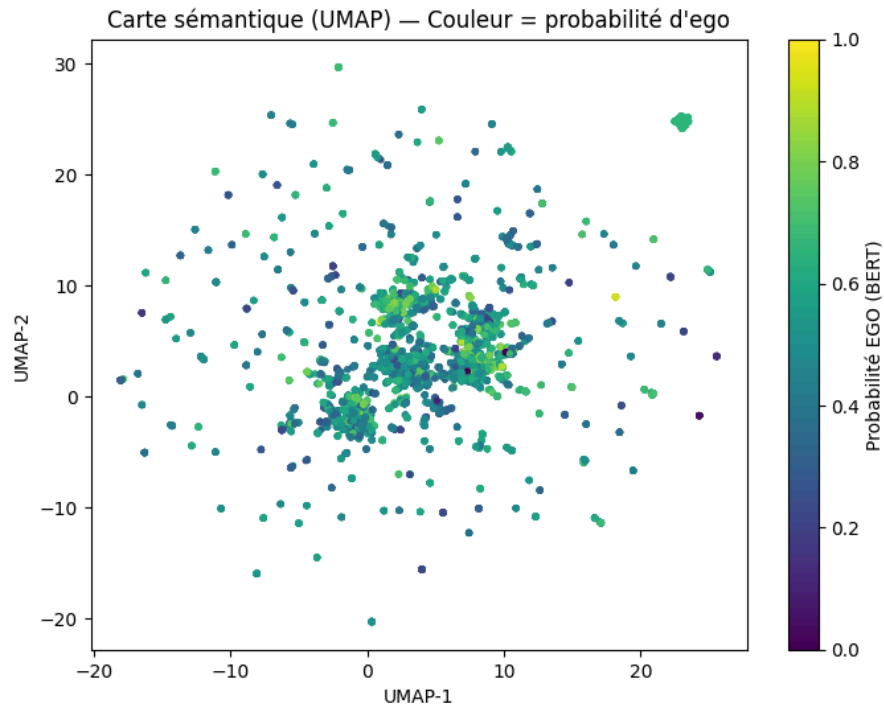


Figure 4 : Carte sémantique BERT (UMAP) colorée par la probabilité d'ego

Chaque point représente une publicité. La distance entre deux points indique leur proximité sémantique (plus ils sont proches, plus ils "racontent la même histoire" ou utilisent le même contexte). La couleur représente la **probabilité d'Ego** calculée par le modèle (Jaune = Discours fortement centré sur le "Je/Puissance" ; Violet = Discours centré sur le "Nous/Collectif").

Interprétation des résultats :

- **L'absence de rupture binaire :** Contrairement à ce que l'on pourrait attendre (deux clusters séparés : les "anciennes pubs égoïstes" vs les "nouvelles pubs morales"), le graphique montre un **nuage continu**. Il n'y a pas de frontière étanche. Cela signifie que la transition identitaire est progressive : les publicités "morales" conservent souvent des codes de l'individu, et les publicités "de performance" intègrent désormais des nuances.
- **Le Centre Teal/Vert :** La très grande densité de points au centre, de couleur intermédiaire (vert d'eau/bleu clair, correspondant à une probabilité de 0.4 à 0.6), indique que la majorité du discours publicitaire automobile est **hybride**. Les marques cherchent un compromis : elles vendent encore de la distinction individuelle tout en la saupoudrant de responsabilité collective.
- **Les pôles de résistance (Les Extrêmes) :** On distingue néanmoins des zones périphériques marquées :
 - Des points **jaunes vifs** (haute probabilité d'ego) isolés, incarnant probablement les publicités pour les sportives ou le luxe (le "Je" pur).
 - Des points **violet foncé** (faible ego), représentant les campagnes purement RSE ou citoyennes (le "Nous" pur).

L'analyse sémantique révèle que la métamorphose du "Je" au "Nous" ne se fait pas par remplacement brutal, mais par **hybridation**. L'identité automobile contemporaine est une nébuleuse complexe où la performance individuelle et la conscience morale tentent de cohabiter au sein d'un même message.

Thématiques dominantes du corpus : Topic Modeling (LDA)

Afin de compléter l'approche lexicale et factorielle, une analyse par **topic modeling** a été menée à l'aide de l'algorithme LDA (*Latent Dirichlet Allocation*). Cette méthode non supervisée vise à détecter les **thématiques latentes** dans les scripts publicitaires, c'est-à-dire les ensembles de mots qui co-apparaissent fréquemment et structurent le discours de manière implicite.

Le modèle, construit sur la base vectorisée des 8 938 scripts nettoyés, a identifié **cinq topics principaux** :

Topic	Principaux mots-clés	Interprétation thématique
0	voix, femme, électrique, mois, euros, partir, hybride, kia, homme, nouvelle	Électromobilité et consommation : discours orienté vers l'achat de véhicules hybrides/électriques et les offres commerciales.
1	voix, femme, citroën, homme, euros, renault, skoda, profitez	Publicité de marque et plaisir de conduite : lexique promotionnel et convivial, centré sur les marques généralistes.
2	voix, homme, euros, femme, partir, mois, renault, opel, dacia	Mobilité accessible et quotidienne : thématique du déplacement utilitaire, du confort et du prix.
3	voix, homme, euros, femme, mois, sans, fiat, toyota, plus	Discours transactionnel et rationnel : renvoie aux messages axés sur l'offre et la performance économique.
4	homme, voix, plus, femme, ça, voiture, tout, peugeot, volkswagen, nouveau	Récit universel de la conduite : vocabulaire transversal, incarnant la voiture comme expérience partagée.

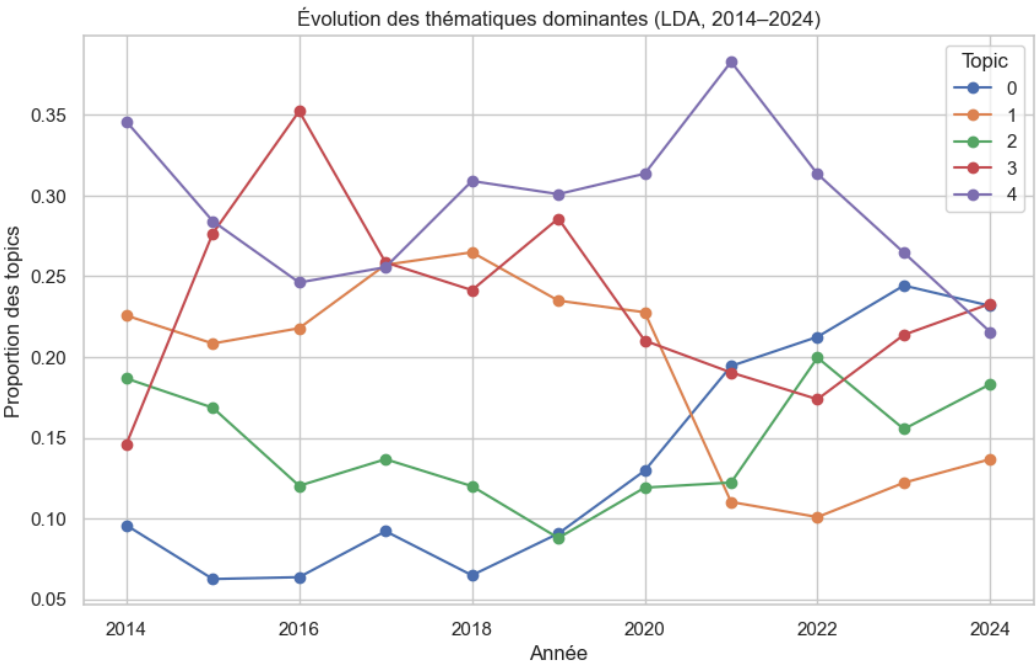


Figure 5 : Graphique de l'évolution des thématiques dominantes (LDA, 2014 - 2024)

Le graphique ci-dessus présente la proportion annuelle de chaque topic entre 2014 et 2024. On distingue trois phases majeures :

- 2014-2017 :**
Les topics 3 et 4 dominent, traduisant la prééminence d'un discours **promotionnel et universel** fondé sur la performance, la nouveauté et la promesse de plaisir.

2. **2018–2020 :**

L'équilibre entre les topics se modifie ; le **topic 1** (plaisir de conduite et marque) et le **topic 0** (hybride/électrique) progressent, signalant une période de **transition discursive** où la technologie devient valeur symbolique.

3. **2021–2024 :**

Le **topic 0** atteint sa plus forte progression, confirmant la **centralité de l'électromobilité** dans les récits publicitaires récents.

Simultanément, les topics 3 et 4 reculent, indiquant la **disparition du discours purement transactionnel** au profit d'un registre moral et collectif.

Le LDA révèle que la thématique "électrique" n'est pas simplement technique : elle fédère un ensemble narratif qui lie innovation, consommation responsable et reconfiguration des genres discursifs.

Le passage du "plaisir d'acheter" (topics 3–4) au "sens d'agir" (topic 0) incarne la mutation du capital symbolique de la voiture — d'objet de possession à signe de transformation.

8. Conclusion

L'ensemble des analyses confirme l'hypothèse d'une publicité automobile conçue comme miroir de la subjectivité contemporaine, où la voiture fonctionne d'abord comme prolongement de soi avant de devenir support d'une identité plus morale et collective (Figure 1, Figure 4).

Sur le plan quantitatif, la montée continue du score identitaire jusqu'en 2021, suivie de sa chute brutale et durable à partir de 2022, atteste du basculement d'un régime discursif centré sur l'ego vers un régime de responsabilité partagée, en cohérence avec l'hypothèse H2 (Figure 2).

L'évolution conjointe des registres rationnel, émotionnel et moral montre que, si l'argumentaire technique demeure dominant, il est progressivement complété par une intensification des appels émotionnels après 2018 et surtout par l'essor d'un registre moral à partir de 2019, indiquant que l'automobile cesse d'être seulement un objet de performance pour devenir un sujet social et éthique (Figure 3).

Enfin, la cartographie sémantique BERT et la recomposition thématique liée à l'électromobilité confirment que le passage du "je" au "nous" ne relève pas d'une simple opposition, mais d'une hybridation où les promesses de liberté et de performance tentent désormais de cohabiter avec l'exigence de durabilité et de bien commun (Figure 4).