Jusqu'où contextualiser la personnalité de marque ?
Scores OCEAN, marque émettrice et engagement sur Twitter

How far should Brand Personality be contextualised?
OCEAN scores, brand as Message Source, and engagement on Twitter

Julien Monnot
Professeur permanent
ISTEC Business School Paris
CERI Centre d'Études et de Recherche de l'ISTEC
j.monnot@istec.fr

Résumé

Les modèles de personnalité de marque varient dans la littérature selon les cultures et les secteurs économiques. Cette diversité traduit le caractère contingent du concept et conduit le champ à développer des instruments de plus en plus contextualisés. Cette recherche interroge le niveau pertinent de contextualisation et examine si elle se prolonge jusqu'à la marque émettrice. L'étude porte sur 12 099 tweets publiés par les comptes officiels français de onze constructeurs automobiles. Dix scores d'association à des labels dérivés du modèle OCEAN décrivent chaque message. Ils forment un référentiel sémantique externe et ne mesurent pas des traits de personnalité. Deux modèles additifs généralisés sont comparés. Le premier impose une relation commune entre ces scores et l'engagement, le second l'autorise à dépendre de la marque émettrice. Les associations varient selon les marques et prennent des directions opposées. Les autoriser à varier améliore faiblement mais régulièrement la prédiction hors échantillon, au prix d'une complexité accrue.

Mots clés

Classification zero-shot ; Contextualisation ; Engagement en ligne ; Modèle additif généralisé ; Modèle OCEAN ; Personnalité de marque

Abstract

Brand personality models vary across cultures and economic sectors in the literature. This diversity reflects the contingent nature of the construct and leads the field to develop increasingly context-specific instruments. This research questions the relevant level of such

contextualisation and examines whether it extends to the posting brand. The study draws on 12,099 tweets published by the official French accounts of eleven car manufacturers. Ten association scores for labels derived from the OCEAN model describe each message. They form an external semantic reference and do not measure personality traits. Two generalised additive models are compared. The first imposes a relation between these scores and engagement that is common to all brands; the second allows it to depend on the posting brand. The associations then vary across brands and take opposite directions. Allowing them to vary yields a small but consistent improvement in out-of-sample prediction, at the cost of added complexity.

Keywords

Brand personality; Contextualisation; Generalised additive model; OCEAN model; Online engagement; Zero-shot classification

1. Introduction

La personnalité de marque est définie comme l'ensemble des caractéristiques humaines associées à une marque (Aaker, 1997). Cette définition emprunte à la psychologie des traits l'hypothèse d'une stabilité des dispositions décrites. Les réplications de la méthodologie de construction d'échelles ne convergent cependant pas vers une structure factorielle commune. Les travaux qui suivent observent que les dimensions obtenues dépendent du contexte d'application. Le champ adapte ainsi le concept à chaque domaine : les magasins (D'Astous, Saïd et Levesque, 2002), la mode (Willems et al., 2011), le luxe (Heine, 2009), les clubs sportifs (Schade, Piehler et Burmann, 2014) ou les destinations touristiques (Pereira, Correia et Schutz, 2015). Sur le plan théorique, le périmètre conceptuel du construit est contesté (Azoulay et Kapferer, 2003) ainsi que sa généralisabilité (Austin, Siguaw et Mattila, 2003). Sur le plan méthodologique, les risques de circularité des procédures factorielles sont documentés (Avis, 2012 ; Avis, Forbes et Ferguson, 2014). Certains travaux vont jusqu'à proposer d'abandonner le concept (Oklevik, Supphellen et Maehle, 2020).

Ces réponses laissent ouverte la question de l'appariement entre l'instrument et l'objet mesuré. Le paradigme de développement des échelles fournit les procédures qui disciplinent les conditions de la validité (Churchill, 1979) : stabilité des interprétations, comparabilité des réponses, reproductibilité des structures obtenues. Cependant, la cohérence statistique d'un instrument n'assure ni l'homogénéité de la signification des items, ni la validité du construit lui-même (Cronbach et Meehl, 1955).

Cette difficulté prend une forme particulière pour les objets symboliques. L'analogie d'Aaker au modèle multi-attributs suppose que les adjectifs expriment des croyances définies et comparables. Or une même réponse peut relever de registres différents. Affirmer qu'une marque de luxe X est « élégante » peut exprimer une croyance personnelle : X est élégante, une croyance normative : il faut considérer X comme élégante ou une croyance de catégorie : le luxe est élégant. L'instrument ne permet pas d'isoler sans équivoque le registre évaluatif mobilisé. La variation entre contextes peut alors porter sur le contenu des croyances plutôt que sur leur intensité. Cette hypothèse invite à questionner plus directement l'adéquation entre la stabilité qu'un instrument impose et les propriétés de l'objet auquel on l'applique.

Ce déplacement prend une portée particulière en marketing. Les traces numériques produites par les marques sont massives. Les approches lexicométriques et les modèles de langue permettent de leur attribuer des scores sémantiques à partir de régularités linguistiques observables directement dans le discours. Cette capacité déplace la question sans la résoudre :

produire des scores statistiquement exploitables ne garantit pas qu'un construit conçu par questionnaire soit transposable dans ce nouveau régime de données.

La transition électrique du secteur automobile constitue un cas révélateur. Elle modifie simultanément les propriétés techniques des véhicules, les cadres normatifs et les récits par lesquels les constructeurs se légitiment. Ces changements questionnent la stabilité des significations habituellement prêtées aux marques du secteur.

Cette communication met ce problème à l'épreuve sur un corpus de 12 099 messages publiés par les comptes officiels français de onze constructeurs. La question posée est celle du niveau : la logique de contextualisation documentée aux niveaux culturel et sectoriel se prolonge-t-elle jusqu'à la marque émettrice ?

2. Cadre théorique

Lorsque la stabilité des significations ne peut être présupposée, l'observation peut être déplacée vers des traces qui préexistent au dispositif plutôt que vers des réponses qu'il sollicite. Le discours présente cette propriété. Il est produit en situation et sous des contraintes d'énonciation propres et permet de traiter la variabilité des significations comme un phénomène observable.

Le référentiel mobilisé est le modèle du Big Five ou OCEAN (Goldberg, 1990). Ce modèle n'est pas une théorie de la personne transposée à la marque : il procède de l'hypothèse lexicale, selon laquelle les traits de personnalité s'observent dans le vocabulaire d'une langue (Allport et Odbert, 1936). Le fondement lexical d'OCEAN rend son emploi comme référentiel de description du discours plus défendable qu'une transposition directe d'une échelle de personnalité de marque. Son recours évite de reconstruire circulairement une échelle à partir des marques elles-mêmes. Les variables déduites sont des scores d'association à des labels et non des mesures validées de traits perçus : rien n'établit ici que ces régularités correspondent à ce qu'un lecteur identifierait comme une personnalité.

L'engagement observable constitue une trace de réception (van Doorn et al., 2010 ; Hollebeek, Glynn et Brodie, 2014). L'identité du compte émetteur intervient enfin comme condition d'énonciation : un énoncé n'est pas reçu indépendamment de qui l'énonce. Cette identité agrège plusieurs propriétés non séparées dans le présent dispositif.

La contextualisation rappelée en introduction laisse ouverte la question du niveau pertinent : à contexte culturel et sectoriel constant, peut-on supposer que les relations observées soient communes aux marques ? Ce déplacement ne reproduit pas les adaptations culturelles ou sectorielles. Il prolonge son principe en examinant la stabilité d'une relation entre un référentiel

commun et un indicateur indirect de réception. Le niveau de la marque est retenu parce qu'il est à la fois l'unité symbolique de source et une unité répétée d'observation.

Deux spécifications emboîtées sont comparées : l'une estime une relation commune aux marques, l'autre l'autorise à dépendre de la source. Cette comparaison examine dans quelle mesure les données textuelles supportent un niveau supplémentaire de contextualisation, et si un éventuel gain d'ajustement se prolonge hors échantillon.

3. Méthodologie

Les messages sont collectés entre le 18 janvier et le 2 mai 2022 avec le package rtweet, sur les comptes officiels français de onze constructeurs automobiles. La procédure récupère jusqu'à 5 000 messages par compte. La profondeur observée va d'octobre 2016 à mai 2022. Seuls les messages primaires sont conservés : les retweets, quotes et replies produits par les comptes sont écartés afin de concentrer l'analyse sur le discours directement émis de la marque vers son audience. Le corpus final compte 12 099 messages (annexe A, tableau A1).

La réception est approchée par la somme des likes et des retweets. Ces deux réactions ne sont pas équivalentes dans la littérature (Muñoz-Expósito et al., 2017). Leur addition retient une forme limitée d'engagement, centrée sur l'approbation et l'amplification. Les replies et quotes relèvent d'un engagement conversationnel distinct, non intégré.

Chacun des dix labels correspondant aux pôles polarisés du modèle OCEAN (10) est soumis séparément au modèle XLM-RoBERTa-XNLI comme hypothèse unique. Le score retenu est la probabilité d'entailment de cette hypothèse au regard du texte. Les dix scores ne font l'objet d'aucune normalisation conjointe. L'inférence porte sur le texte nettoyé de chaque message. Quatre catégories de contenu et trois de motorisation sont inférées de la même manière (annexe B, tableaux B1 et B2). Deux contrôles encadrent cette procédure. Un contrôle amont compare trois familles de stratégies d'inférence sur des énoncés construits non ambigus, puis sur des fractions croissantes du corpus. Il conduit à écarter celles qui ne restituent pas les distinctions attendues sur des cas prototypes. Un contrôle aval examine la structure de corrélation entre les dix scores obtenus sur le corpus (annexe B, sections B3 et B4 ; figure B1).

Les variables de comptage sont transformées par log1p, les scores par logit, et l'ensemble des variables continues est centré-réduit. Le modèle est un modèle additif généralisé de famille scat(), qui autorise des résidus à queues plus lourdes. Il comprend un terme lisse de temporalité ainsi que le moment de la journée, la longueur du message, le nombre de hashtags et de mentions, deux catégories de contenu (sponsoring et symbolique) et les trois catégories de motorisation. Le nombre d'abonnés n'est pas intégré : les métadonnées disponibles ne

permettent pas de reconstruire l'audience à la date de publication et sont essentiellement portées par l'identité du compte émetteur (annexe A, tableau A3).

La spécification est construite par ajouts successifs de blocs de variables. Deux étapes sont comparées. M0 comprend les contrôles, les interactions déjà retenues entre la marque et le moment de la journée, le sponsoring et les trois catégories de motorisation, ainsi que les dix scores OCEAN. Il impose que la relation avec la réception soit commune aux marques. M1 ne diffère de M0 que par l'ajout des dix interactions entre ces scores et l'identité du compte émetteur. Une spécification de référence, privée des dix scores, sert d'étalon pour mesurer leur apport propre (annexe D, tableau D1). L'AIC, le BIC et le test du rapport de vraisemblance apprécient l'ajustement et la parcimonie des spécifications. Une validation croisée comparative les évalue sur des messages tenus à l'écart de l'estimation, ce qui fait porter à chaque spécification le coût de sa propre complexité. Les partitions sont communes aux trois modèles, stratifiées par marque et répétées cinq fois. Une amélioration limitée aux critères internes est ainsi distinguée d'un gain qui se reproduit hors échantillon. La robustesse des coefficients de M1 est évaluée par bootstrap non paramétrique ($B = 500$) : un coefficient est retenu lorsque son intervalle de confiance à 95 % exclut zéro et que son signe est majoritaire sur au moins 95 % des répétitions. Les estimations détaillées par compte figurent en annexes C et D.

4. Résultats

Les volumes sont inégaux entre comptes : de 524 messages pour Audi à 1 716 pour Peugeot, soit de 4,3 % à 14,2 % du corpus. Les fenêtres temporelles diffèrent également, le premier message disponible allant d'octobre 2016 à mai 2019 selon les comptes. La réception est fortement asymétrique : la moyenne excède partout la médiane, dans un rapport allant de 1,4 à 7,6, et le maximum atteint 69 495 pour un message (annexe A, tableau A2).

La contribution des scores dépend de la spécification retenue. Évaluée sur des messages tenus à l'écart de l'estimation, l'introduction des dix scores avec une relation commune aux marques réduit le RMSE moyen de 0,0003. Autoriser cette relation à dépendre de la source le réduit de 0,0032, soit plus de dix fois davantage. Les deux gains sont positifs dans les cinq répétitions (annexe D, tableau D3). Sur les données d'estimation, l'ajout de ces interactions améliore l'ajustement (AIC : 27 329 à 27 156 ; D^2 : 0,368 à 0,380 ; test du rapport de vraisemblance, $p < 0,001$), tandis que le BIC sanctionne la complexité correspondante (28 090 à 28 656 ; annexe D, tableau D1). Le gain hors échantillon demeure faible en valeur absolue : le RMSE moyen passe de 0,789 pour M0 à 0,786 pour M1. M1 présente un RMSE inférieur dans 21 des 25 plis et dans la moyenne de chacune des cinq répétitions (annexe D, tableau D2).

Trente des cent dix associations M1 entre un score et un compte satisfont le critère de robustesse. Leur amplitude reste modérée : la plus forte correspond à une variation multiplicative retranscrite de +24 % pour le score E_pos chez Audi. Pour sept des dix scores, les associations robustes comprennent des signes opposés selon le compte émetteur. O_pos, C_neg et N_neg conservent une orientation commune. Deux comptes ne présentent aucune association robuste (annexe D, tableaux D4 et D6). Considérées sans condition de robustesse, les onze associations conditionnelles comportent des valeurs de signes opposés pour chacun des dix scores, et la pente commune estimée dans M0 coïncide avec leur moyenne : l'écart entre les deux est de 0,007 en valeur absolue, pour une étendue conditionnelle médiane de 0,179 (annexe D, tableau D8). Imposer une relation commune revient donc à moyenner des associations contraires. La figure D1 restitue ces associations par compte et par score. Elle fait apparaître des comptes fortement différenciés : Ford, Porsche et Audi présentent respectivement sept, six et cinq associations robustes. Hyundai et Nissan n'en présentent aucune (annexe D, tableau D5).

5. Discussion et limites

Ce corpus indique que la contextualisation ne rend pas les relations entre scores et réception communes aux marques. Les effets conditionnels observés sont d'amplitude limitée et coûteux en complexité, dans la tension entre différenciation et généralisation identifiée par Austin, Siguaw et Mattila (2003). Trois lectures restent ouvertes. La première attribue cette dépendance à l'identité de la source : les contrôles présentent eux aussi des associations conditionnelles au compte (annexe D, tableau D7), ce qui suggère un contexte de réception de marque plus qu'une spécificité des scores OCEAN. La deuxième y voit une variation de la signification : un même score renverrait à des configurations dont la portée diffère selon la marque qui les énonce. La troisième l'impute aux propriétés de l'inférence. Le dispositif ne permet pas de les départager.

Quatre limites. La validité perceptuelle des scores n'est pas établie : aucune annotation humaine n'a été conduite. Le cadre interprétatif est en partie ex post : les configurations sont interprétées, non prédites. La validation croisée porte sur les mêmes constructeurs. Enfin, un seul secteur, un seul pays, une seule plateforme : la confrontation à un régime d'énonciation distinct, où la source n'est plus la marque elle-même, permettrait de départager les trois lectures.

Descendre d'un niveau de contextualisation ne clôt pas la série. À contexte culturel et sectoriel constant, une hétérogénéité subsiste, qu'une relation commune aux marques efface. La difficulté ne semble donc pas se réduire au niveau retenu, mais à la compatibilité entre la stabilité de signification qu'un dispositif présuppose et celle de l'objet auquel il s'applique.

Références bibliographiques

Aaker JL (1997) Dimensions of brand personality. *Journal of Marketing Research* 34(3): 347-356.

Allport GW et Odbert HS (1936) Trait-names: A psycho-lexical study. *Psychological Monographs* 47(1): i-171.

Austin JR, Siguaw JA et Mattila AS (2003) A re-examination of the generalizability of the Aaker brand personality measurement framework. *Journal of Strategic Marketing* 11(2): 77-92.

Avis M (2012) Brand personality factor based models: A critical review. *Australasian Marketing Journal* 20(1): 89-96.

Avis M, Forbes S et Ferguson S (2014) The brand personality of rocks: A critical evaluation of a brand personality scale. *Marketing Theory* 14(4): 451-475.

Azoulay A et Kapferer JN (2003) Do brand personality scales really measure brand personality? *Journal of Brand Management* 11(2): 143-155.

Churchill GA (1979) A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research* 16(1): 64-73.

Cronbach LJ et Meehl PE (1955) Construct validity in psychological tests. *Psychological Bulletin* 52(4): 281-302.

D'Astous A, Saïd I et Levesque M (2002) Conception et test d'une échelle de mesure de la personnalité de la marque. *Actes du Congrès de l'Association française de marketing* 18: 115-130.

Goldberg LR (1990) An alternative « description of personality »: The Big-Five factor structure. *Journal of Personality and Social Psychology* 59(6): 1216-1229.

Heine K (2009) Using personal and online repertory grid methods for the development of a luxury brand personality. *Electronic Journal of Business Research Methods* 7(1): 25-38.

Hollebeek LD, Glynn MS et Brodie RJ (2014) Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of Interactive Marketing* 28(2): 149-165.

Münker S, Kugler K et Rettinger A (2025) Zero-shot prompt-based classification: Topic labeling in times of foundation models in German tweets. In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, Student Research Workshop*, pp.53-63.

Muñoz-Expósito M, Oviedo-García MÁ et Castellanos-Verdugo M (2017) How to measure engagement in Twitter: Advancing a metric. *Internet Research* 27(5): 1122-1148.

Oklevik O, Supphellen M et Maehle N (2020) Time to retire the concept of brand personality? *Journal of Brand Management* 27(2): 177-190.

Pereira RLG, Correia AH et Schutz RLA (2015) Towards a taxonomy of a golf-destination brand personality: Insights from the Algarve golf industry. *Journal of Destination Marketing & Management* 4(1): 57-67.

Schade M, Piehler R et Burmann C (2014) Sport Club Brand Personality Scale (SCBPS): A new brand personality scale for sport clubs. *Journal of Brand Management* 21: 650-663.

van Doorn J, Lemon KN, Mittal V, et al. (2010) Customer engagement behavior: Theoretical foundations and research directions. *Journal of Service Research* 13(3): 253-266.

Willems K, Swinnen G, Janssens W et Brengman M (2011) Fashion store personality: Scale development and relation to self-congruity theory. *Journal of Global Fashion Marketing* 2(2): 55-65.

Annexe A - Corpus et collecte

Les messages ont été collectés avec le package rtweet entre le 18 janvier et le 2 mai 2022, sur onze comptes officiels français : @renault_fr, @PeugeotFR, @CitroenFrance, @AudiFrance, @vw_france, @Porsche_France, @FordFrance, @MBFRANCE_, @nissanfrance, @ToyotaFrance et @HyundaiFrance. La procédure interroge chaque compte à hauteur de 5 000 messages au maximum. La profondeur temporelle du corpus, du 2 octobre 2016 au 2 mai 2022, résulte de ce plafond et du rythme de publication propre à chaque compte ; elle ne correspond pas à une fenêtre définie a priori et n'est pas homogène entre comptes.

Tableau A1. Journal de filtrage.

Étape	Effectif
Collecte brute, après suppression des doublons	33 642
Retrait des relais, citations et réponses émis par les comptes	13 117
Retrait des réponses commerciales mal étiquetées par l'interface et de trois identifiants isolés	12 099

Note. Les réponses commerciales sont des messages adressés à un utilisateur particulier, que l'interface ne signale pas comme réponses. Elles sont repérées par la conjonction d'une mention initiale et d'une formule d'adresse.

Tableau A2. Distribution des messages et de la réception, par compte.

Compte	N	%	Premier	Dernier	Médiane	Moyenne	Max.
Peugeot	1 716	14,2	23/01/2017	02/05/2022	60,0	82,6	3 922
Porsche	1 645	13,6	24/03/2017	02/05/2022	50,0	72,9	15 190
Ford	1 474	12,2	19/05/2017	29/04/2022	13,0	21,4	974
Renault	1 213	10,0	08/03/2017	02/05/2022	38,0	124,1	69 495
Hyundai	1 119	9,2	02/10/2016	29/04/2022	8,0	60,4	3 705
Mercedes	1 083	9,0	20/04/2017	01/05/2022	31,0	42,2	1 223
Citroën	1 034	8,5	17/10/2018	02/05/2022	30,0	46,8	3 688
Toyota	917	7,6	14/12/2016	28/04/2022	16,0	28,2	786
Nissan	738	6,1	24/05/2017	01/04/2022	12,0	24,7	797
Volkswagen	636	5,3	15/12/2017	29/04/2022	59,5	195,5	19 508
Audi	524	4,3	03/05/2019	28/04/2022	59,0	92,3	1 444

Note. Les volumes sont inégaux : Peugeot représente 14,2 % du corpus, Audi 4,3 %. Les fenêtres temporelles diffèrent également (le premier message disponible va d'octobre 2016 pour Hyundai à mai 2019 pour Audi) ce qui reflète le plafond de récupération et le rythme de publication de chaque compte plutôt qu'une différence d'ancienneté sur la plateforme. La réception est fortement asymétrique : la moyenne excède partout la médiane, dans un rapport allant de 1,4 à 7,6 selon les comptes, et le maximum atteint 69 495 pour un message de Renault. Cette asymétrie motive la transformation logarithmique et le choix d'une famille de résidus à queues lourdes.

Tableau A3. Métadonnées d'audience par compte.

Compte	Valeurs distinctes	Minimum	Maximum	Étendue relative
Volkswagen	10	150 024	150 294	0,18 %
Renault	13	127 898	128 488	0,46 %
Audi	11	110 406	110 711	0,28 %
Peugeot	13	85 740	86 380	0,75 %
Citroën	12	82 486	83 467	1,19 %
Nissan	8	57 581	57 611	0,05 %
Toyota	12	43 568	43 807	0,55 %
Ford	12	41 322	41 482	0,39 %
Hyundai	11	28 764	29 287	1,82 %
Mercedes	9	26 166	27 309	4,37 %

Compte	Valeurs distinctes	Minimum	Maximum	Étendue relative
Porsche	11	16 951	17 977	6,05 %

Note. L'interface restitue le nombre d'abonnés à la date de collecte et non à celle de publication. Les quelques valeurs distinctes observées par compte correspondent aux dates de requêtage successives et non à une trajectoire d'audience : deux messages publiés à cinq ans d'intervalle se voient attribuer une valeur pratiquement identique. L'étendue relative intra-compte va de 0,05 % à 6,05 %, alors que le nombre d'abonnés varie d'un facteur 8,9 entre comptes. Cette variable est donc essentiellement portée par l'identité du compte, avec laquelle elle serait redondante dans le modèle ; elle ne permet pas de contrôler l'audience au moment de chaque message.

Annexe B - Inférence sémantique

B1. Procédure

Chacun des dix labels est soumis séparément au modèle joeddav/xlm-roberta-large-xnli comme hypothèse unique ; le score retenu est la probabilité d'entailment de cette hypothèse au regard du texte. Les dix scores ne font l'objet d'aucune normalisation conjointe et ne somment pas à un. L'unité d'inférence est le texte nettoyé de chaque message, et non la phrase.

Le nettoyage passe le texte en minuscules, retire les caractères @ et # en conservant le contenu lexical correspondant, supprime les adresses réticulaires et la ponctuation à l'exception du trait d'union, et écarte les mots d'une seule lettre à l'exception de « e » et « i ». Deux substitutions sont appliquées avant l'inférence : le tréma est ramené à l'accent aigu, et « électric » à « électrique ».

B2. Formulation des labels

Tableau B1. Labels des dix pôles polarisés.

Variable	Formulation soumise au modèle
O_pos	Recherche activement des idées nouvelles, curieux et créatif.
O_neg	Refuse le changement, préfère les idées traditionnelles et rigides.
C_pos	Rigueur, organisation, responsabilité et souci du détail.
C_neg	Désorganisation, manque de discipline et de responsabilité.
E_pos	Sociable, énergique, aime être le centre d'attention et recherche l'interaction sociale.
E_neg	Réservé, solitaire, préfère l'introspection et évite les interactions sociales.
A_pos	Bienveillant, empathique, recherche l'harmonie et la coopération.
A_neg	Indifférent, conflictuel, cherche la domination et l'opposition.
N_pos	Calme, serein, résilient face aux défis émotionnels.
N_neg	Instable émotionnellement, facilement stressé et anxieux.

Note. La dimension sous-jacente est rappelée pour le lecteur ; elle n'est pas transmise au modèle. La formulation de N_pos correspond directement à la stabilité émotionnelle : aucune inversion n'est appliquée aux scores après inférence.

Tableau B2. Labels de contenu et de motorisation.

Variable	Formulation soumise au modèle
Sponsoring	Contenu sur des partenariats extérieurs au secteur automobile (ex : sport, célébrité, événement caritatif).
Événement	Contenu sur un événement automobile précis : salon, course, lancement ou exposition.
Technique	Contenu sur les aspects techniques des voitures : moteur, sécurité, innovation ou performance mécanique.
Symbolique	Contenu émotionnel ou symbolique des voitures : liberté, plaisir, statut, nostalgie, rêve automobile.
Hybride	Contenu sur une motorisation hybride, utilisant à la fois de l'électricité et du carburant fossile.

Variable	Formulation soumise au modèle
Thermique	Contenu sur une motorisation thermique classique, utilisant uniquement du carburant fossile.
Électrique	Contenu sur une motorisation entièrement électrique, sans aucun carburant fossile.

Note. Quatre catégories de contenu ont été inférées ; seules Sponsoring et Symbolique sont retenues comme contrôles dans les spécifications comparées. Technique et Événement n'entrent ni dans M0 ni dans M1.

B3. Comparaison des stratégies d'inférence

Aucun modèle pré-entraîné pour la caractérisation du discours selon le référentiel en cinq facteurs n'est disponible en français, et il n'existe pas de corpus annoté de référence permettant un entraînement supervisé comparé à une vérité terrain. Trois familles de stratégies ont donc été construites et comparées avant d'arrêter la procédure retenue.

La première repose sur un dictionnaire et un étiquetage faible. Elle mobilise un dictionnaire de 752 adjectifs français répartis en dix catégories polarisées, élaboré à partir de Goldberg (1990) par traduction, génération de synonymes et d'antonymes, classification manuelle et épuration qualitative. Les variables produites sont des indicatrices obtenues par appariement lexical, complétées par une allocation de Dirichlet latente amorcée par dictionnaire. Sur des messages courts, ces indicatrices se révèlent peu discriminantes, et la procédure fait dépendre l'inférence d'unités lexicales sélectionnées dans le domaine décrit.

La deuxième ajuste des modèles de type CamemBERT et BERT sur ces étiquettes faibles, selon deux protocoles d'étiquetage (fréquences d'adjectifs, puis allocation supervisée par dictionnaire) et deux formats de sortie, multi-classe et multi-étiquette. La troisième repose sur l'inférence par entailment, comparée entre deux modèles multilingues.

Les stratégies ont d'abord été confrontées à des énoncés construits non ambigus, de taille croissante (10, 100 puis 1 000 énoncés, soit 1, 10 et 100 par pôle), afin d'observer leur sensibilité aux distinctions du référentiel dans des conditions contrôlées, avant application à des fractions croissantes du corpus : 10 %, 20 % puis 50 %. Le recours à des énoncés construits pallie l'absence de corpus annoté de référence en français (Münker et al., 2025). Ce contrôle a permis d'écarter les procédures incapables de restituer les distinctions attendues sur des cas prototypes.

Sa portée est limitée à deux égards. Le pôle générateur de chaque énoncé n'a pas été exploité comme cible : la comparaison retenue porte sur la cohérence de la structure de corrélation obtenue, non sur un taux d'attribution correcte. Et la restitution de distinctions explicites sur des prototypes ne renseigne pas la performance de la procédure sur des messages courts et ambigus. Ce contrôle se situe donc entre la cohérence interne des sorties et une validation perceptuelle, sans se substituer à cette dernière.

B4. Contrôles de cohérence interne

La procédure retenue produit une structure de corrélation organisée en deux ensembles, correspondant aux pôles positifs et aux pôles négatifs du référentiel. Cette bipartition n'est pas imposée par le dispositif : les dix labels sont soumis séparément au modèle et leurs scores ne font l'objet d'aucune normalisation conjointe. Les variations inter- et intra-factorielles observées sont compatibles avec l'organisation attendue du référentiel.

Une stratégie soumettant les dix labels simultanément produit une structure de corrélation moins cohérente ; l'évaluation séparée neutralise en outre l'effet de l'ordre de présentation des labels. Les facteurs d'inflation de la variance sont inférieurs à 2 pour les dix scores et inférieurs à 1,5 pour les catégories de contenu et de motorisation ; la corrélation la plus élevée entre ces dernières atteint 0,41, entre Hybride et Électrique.

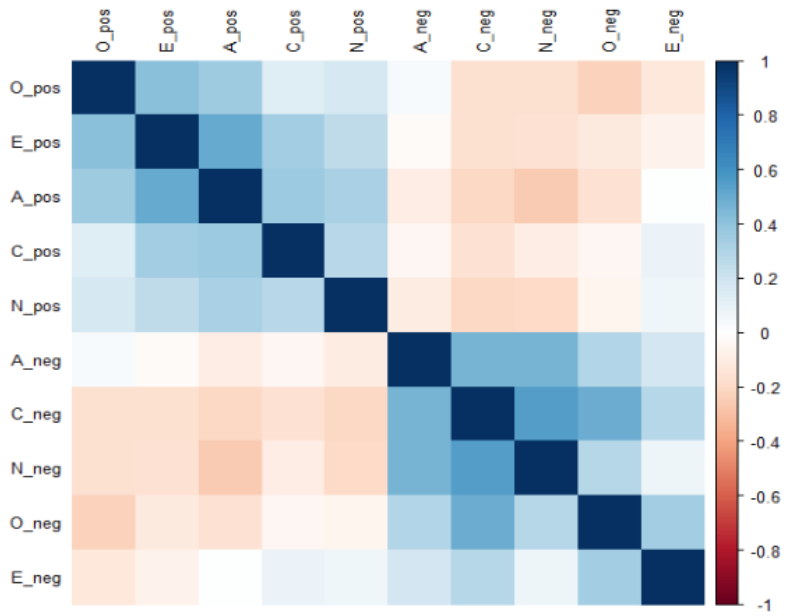


Figure B1. Corrélations entre les dix scores d'association, ordonnées par regroupement hiérarchique.

Note. Le regroupement fait apparaître deux ensembles, correspondant aux pôles positifs et aux pôles négatifs du référentiel. Les corrélations sont positives à l'intérieur de chaque ensemble et négatives ou faibles entre eux. Cette structure n'est pas imposée par le dispositif : les dix labels sont soumis séparément au modèle et leurs scores ne font l'objet d'aucune normalisation conjointe. Les corrélations entre les catégories de contenu et de motorisation ne sont pas reproduites ici ; leur valeur la plus élevée atteint 0,41, entre Hybride et Électrique.

Statut de ces contrôles. Ils documentent la cohérence interne des sorties du classifieur. Ils ne renseignent pas la correspondance entre ces sorties et un jugement externe. Aucune annotation humaine n'a été conduite : la validité perceptuelle des scores n'est pas établie, et ils ne sont pas interprétés comme une mesure de traits.

Annexe C - Niveaux et contrastes de source

Tableau C1. Niveaux ajustés par compte émetteur.

Compte	η	IC 95 %	Niveau ajusté	IC 95 %
Peugeot	0,653	[0,609 ; 0,696]	61,9	[58,9 ; 65,0]
Volkswagen	0,633	[0,494 ; 0,784]	60,5	[51,9 ; 71,6]
Porsche	0,373	[0,330 ; 0,414]	45,3	[43,2 ; 47,4]
Audi	0,240	[0,146 ; 0,330]	39,0	[35,1 ; 43,2]
Renault	0,216	[0,168 ; 0,262]	38,0	[36,0 ; 40,0]
Citroën	−0,072	[−0,133 ; −0,010]	27,4	[25,6 ; 29,5]
Mercedes	−0,119	[−0,171 ; −0,071]	26,0	[24,5 ; 27,5]
Toyota	−0,484	[−0,552 ; −0,413]	17,1	[15,8 ; 18,6]
Ford	−0,773	[−0,826 ; −0,724]	12,2	[11,5 ; 12,9]
Nissan	−0,782	[−0,851 ; −0,709]	12,1	[11,1 ; 13,2]
Hyundai	−0,832	[−0,940 ; −0,709]	11,4	[10,0 ; 13,2]

Note. Ces quantités sont des prédictions retranscrites au point de référence du modèle (variables continues à leur moyenne, moment de la journée en après-midi, contribution nulle du terme lissé) et non des effets marginaux de compte, le modèle comportant des interactions. La retransformation $\exp(\mu + \sigma Y \cdot \eta) - 1$ porte sur le prédicteur ajusté et ne comporte pas de correction de rétro-transformation des résidus : elle ne s'interprète pas comme une espérance de la réception. Intervalles bootstrap (B = 500).

Tableau C2. Contrastes entre comptes, référence Audi.

Compte	β	IC 95 %	Variation %	Robuste
Peugeot	0,413	[0,308 ; 0,515]	+57,1	oui
Volkswagen	0,393	[0,224 ; 0,589]	+53,7	oui
Porsche	0,133	[0,034 ; 0,234]	+15,7	oui
Renault	−0,024	[−0,124 ; 0,071]	−2,6	non
Citroën	−0,312	[−0,419 ; −0,204]	−29,0	oui
Mercedes	−0,359	[−0,458 ; −0,254]	−32,5	oui
Toyota	−0,724	[−0,848 ; −0,615]	−54,7	oui
Ford	−1,013	[−1,113 ; −0,910]	−67,0	oui
Nissan	−1,022	[−1,126 ; −0,912]	−67,3	oui
Hyundai	−1,072	[−1,221 ; −0,930]	−69,1	oui

Note. Neuf contrastes sur dix satisfont le critère de robustesse ; celui de Renault ne le satisfait pas, son signe n'étant stable que sur 64,6 % des répétitions. Variation obtenue par $100 \cdot [\exp(\beta \cdot \sigma Y) - 1]$.

Annexe D - Estimation et robustesse

Tableau D1. Démarche additive par blocs.

Bloc ajouté	AIC	BIC	RMSE	D^2	LRT
Facteurs	29 695,0	29 813,4	0,858	0,268	-
Temporalité	28 716,2	28 898,9	0,829	0,310	***
Structure du message	28 488,7	28 693,6	0,823	0,318	***
Contenu	28 373,2	28 593,0	0,821	0,323	***
Motorisation	28 271,3	28 513,2	0,818	0,327	***
Scores OCEAN	28 233,6	28 549,8	0,815	0,329	***
Interactions contrôles	27 439,9	27 977,8	0,788	0,362	***
M0 - interactions motorisation	27 329,2	28 089,5	0,783	0,368	***
M1 - interactions marque × OCEAN	27 155,9	28 656,3	0,773	0,380	***

Note. *** : $p < 0{,}001$ au test du rapport de vraisemblance par rapport au bloc précédent. Le BIC décroît jusqu'au bloc des interactions de contrôle puis remonte, fortement lors de l'ajout du bloc marque × OCEAN. Les critères sont rapportés conjointement : aucun ne détermine seul l'interprétation.

Tableau D2. Validation croisée comparative M0 / M1, par répétition.

Répétition	RMSE M0	RMSE M1	Δ	Plis favorables à M1
1	0,7890	0,7863	0,00267	3 / 5
2	0,7888	0,7865	0,00228	4 / 5
3	0,7891	0,7863	0,00281	5 / 5
4	0,7899	0,7853	0,00461	5 / 5
5	0,7894	0,7857	0,00377	4 / 5
Moyenne	0,7892	0,7860	0,00323	21 / 25

Note. Δ = RMSE M0 − RMSE M1 ; une valeur positive signifie que M1 prédit mieux. Cinq répétitions de cinq plis, partitions communes aux deux spécifications et stratifiées par compte ; les vingt-cinq comparaisons ont convergé. Au niveau des plis, l'écart-type de Δ atteint 0,0029, soit l'ordre de grandeur du gain moyen : un partitionnement unique ne permettrait pas de conclure. Entre répétitions, il tombe à 0,00095, et les cinq répétitions sont favorables à M1. La répétition des partitions permet ainsi de vérifier que le gain ne dépend pas d'un découpage particulier. La réduction observée hors échantillon représente environ un tiers de l'amélioration apparente sur les données d'estimation.

Tableau D3. Validation croisée à trois spécifications, par répétition.

Répétition	RMSE référence	RMSE M0	RMSE M1	Apport commun	Apport conditionnel
1	0,78925	0,78897	0,78630	+0,00028	+0,00267
2	0,78888	0,78875	0,78647	+0,00013	+0,00228
3	0,78944	0,78910	0,78629	+0,00034	+0,00281
4	0,79011	0,78986	0,78525	+0,00025	+0,00461
5	0,78971	0,78943	0,78566	+0,00028	+0,00377
Moyenne	0,78948	0,78922	0,78600	+0,00026	+0,00323

Note. Spécification de référence : contrôles et interactions marque × contrôles, sans les dix scores. Apport commun = RMSE référence − RMSE M0 ; apport conditionnel = RMSE M0 − RMSE M1. Une valeur positive signifie que la spécification la plus riche prédit mieux. Mêmes partitions que le tableau D2, stratifiées par compte et répétées cinq fois ; les vingt-cinq comparaisons ont convergé. L'écart-type entre répétitions est de 0,00008 pour l'apport commun et de 0,00095 pour l'apport conditionnel. Les deux apports sont positifs dans les cinq répétitions ; au niveau des plis, l'apport commun l'est dans 17 cas sur 25 et l'apport conditionnel dans 21. La validation croisée arbitrant elle-même le coût des paramètres supplémentaires, cette comparaison n'est pas affectée par la différence de dimension entre les deux blocs : dix termes pour la relation commune, cent dix pour la relation conditionnelle.

Prétraitement. Le centrage-réduction est calculé sur le corpus complet, avant découpage. Un recalcul des paramètres de standardisation dans chaque pli d'entraînement déplace les RMSE

absolus de moins de cinq cent-millièmes et laisse les gains appariés inchangés à la cinquième décimale (0,00026 et 0,00323 dans les deux protocoles).

Bootstrap

La distribution des coefficients est obtenue par rééchantillonnage non paramétrique avec remise, à 500 répétitions, le modèle étant réestimé à chaque itération. Les 500 itérations ont convergé et portent sur les 200 termes de la spécification M1. Un coefficient est retenu lorsque son intervalle de confiance à 95 % exclut zéro et que son signe est majoritaire sur au moins 95 % des répétitions.

Tableau D4. Associations robustes entre scores et réception, par score.

Score	Robustes	Positives	Négatives	Signes opposés
O_neg	5	1	4	oui
E_neg	4	3	1	oui
A_pos	3	1	2	oui
E_pos	3	2	1	oui
N_pos	3	1	2	oui
N_neg	3	0	3	non
O_pos	3	0	3	non
A_neg	2	1	1	oui
C_neg	2	2	0	non
C_pos	2	1	1	oui
Total	30	12	18	7 / 10

Tableau D5. Associations robustes par compte émetteur.

Compte	Associations robustes
Ford	7
Porsche	6
Audi	5
Citroën	3
Mercedes	3
Renault	2
Toyota	2
Peugeot	1
Volkswagen	1
Hyundai	0
Nissan	0

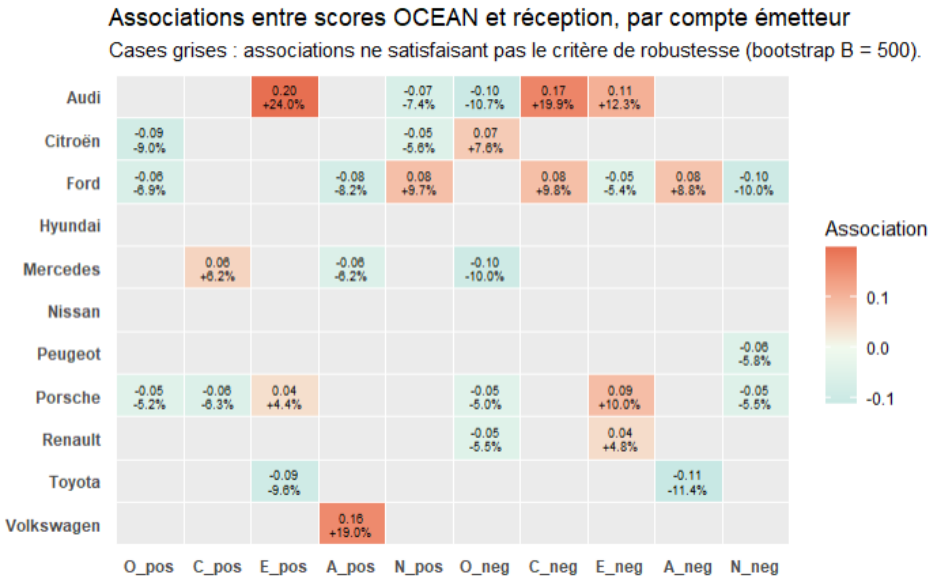


Figure D1. Associations entre scores OCEAN et réception, par compte émetteur.

Note. Les pourcentages sont obtenus par 100·[exp(Δη·σY)−1] et expriment la variation multiplicative de la réception augmentée d'une unité pour un écart-type du score logit-transformé, conditionnellement aux autres termes du modèle. Ils ne sont ni causaux, ni induits par un trait.

Tableau D6. Détail des associations robustes.

Compte	Score	β	IC 95 %	Variation %
Audi	C_neg	0,166	[0,058 ; 0,284]	+19,9
Audi	E_neg	0,106	[0,010 ; 0,187]	+12,3
Audi	E_pos	0,196	[0,114 ; 0,288]	+24,0
Audi	N_pos	−0,070	[−0,130 ; −0,006]	−7,4
Audi	O_neg	−0,103	[−0,197 ; −0,017]	−10,7
Citroën	N_pos	−0,052	[−0,092 ; −0,011]	−5,6
Citroën	O_neg	0,067	[0,012 ; 0,118]	+7,6
Citroën	O_pos	−0,086	[−0,129 ; −0,049]	−9,0
Ford	A_neg	0,077	[0,028 ; 0,130]	+8,8
Ford	A_pos	−0,079	[−0,128 ; −0,028]	−8,2
Ford	C_neg	0,085	[0,027 ; 0,139]	+9,8
Ford	E_neg	−0,051	[−0,095 ; −0,015]	−5,4
Ford	N_neg	−0,096	[−0,150 ; −0,044]	−10,0
Ford	N_pos	0,084	[0,046 ; 0,125]	+9,7
Ford	O_pos	−0,065	[−0,103 ; −0,023]	−6,9
Mercedes	A_pos	−0,059	[−0,098 ; −0,019]	−6,2
Mercedes	C_pos	0,055	[0,022 ; 0,090]	+6,2
Mercedes	O_neg	−0,096	[−0,144 ; −0,052]	−10,0
Peugeot	N_neg	−0,055	[−0,090 ; −0,020]	−5,8
Porsche	C_pos	−0,059	[−0,093 ; −0,024]	−6,3
Porsche	E_neg	0,087	[0,053 ; 0,117]	+10,0
Porsche	E_pos	0,039	[0,001 ; 0,075]	+4,4
Porsche	N_neg	−0,052	[−0,097 ; −0,009]	−5,5
Porsche	O_neg	−0,046	[−0,084 ; −0,011]	−5,0
Porsche	O_pos	−0,048	[−0,084 ; −0,008]	−5,2
Renault	E_neg	0,043	[0,008 ; 0,082]	+4,8
Renault	O_neg	−0,051	[−0,102 ; −0,002]	−5,5
Toyota	A_neg	−0,111	[−0,190 ; −0,036]	−11,4
Toyota	E_pos	−0,092	[−0,159 ; −0,024]	−9,6
Volkswagen	A_pos	0,159	[0,039 ; 0,272]	+19,0

Note. Trente associations sur cent dix satisfont le critère de robustesse. Variation obtenue par 100·[exp(β·σY)−1], pour un écart-type du score logit-transformé.

Tableau D7. Contrôles conditionnés par le compte émetteur, associations robustes.

Compte	Contrôle	β	IC 95 %	Variation %
Hyundai	Soir	1,593	[1,296 ; 1,915]	+471,7
Nissan	Soir	0,476	[0,266 ; 0,668]	+68,5
Audi	Matin	0,422	[0,204 ; 0,671]	+58,8
Hyundai	Sponsoring	0,260	[0,170 ; 0,360]	+33,0
Ford	Soir	0,205	[0,001 ; 0,452]	+25,1
Nissan	Matin	0,166	[0,050 ; 0,279]	+20,0
Renault	Soir	0,148	[0,001 ; 0,319]	+17,6
Hyundai	Thermique	0,112	[0,020 ; 0,198]	+13,0
Mercedes	Matin	0,091	[0,018 ; 0,176]	+10,4
Citroën	Sponsoring	0,041	[0,008 ; 0,076]	+4,6
Peugeot	Électrique	−0,039	[−0,075 ; −0,003]	−4,2
Peugeot	Thermique	−0,047	[−0,090 ; −0,008]	−5,0
Mercedes	Hybride	−0,050	[−0,087 ; −0,014]	−5,3
Porsche	Sponsoring	−0,050	[−0,088 ; −0,015]	−5,4
Renault	Hybride	−0,051	[−0,097 ; −0,006]	−5,4
Mercedes	Électrique	−0,051	[−0,080 ; −0,016]	−5,4
Toyota	Hybride	−0,054	[−0,109 ; −0,005]	−5,7
Ford	Thermique	−0,061	[−0,113 ; −0,011]	−6,5
Ford	Électrique	−0,070	[−0,116 ; −0,024]	−7,4
Mercedes	Thermique	−0,079	[−0,117 ; −0,040]	−8,3
Porsche	Hybride	−0,089	[−0,125 ; −0,049]	−9,3
Porsche	Thermique	−0,089	[−0,126 ; −0,049]	−9,3
Hyundai	Électrique	−0,112	[−0,194 ; −0,028]	−11,5
Audi	Sponsoring	−0,122	[−0,210 ; −0,043]	−12,5
Volkswagen	Électrique	−0,141	[−0,261 ; −0,035]	−14,3
Audi	Électrique	−0,175	[−0,230 ; −0,123]	−17,5
Porsche	Matin	−0,176	[−0,250 ; −0,095]	−17,5
Hyundai	Matin	−0,282	[−0,415 ; −0,155]	−26,5
Volkswagen	Matin	−0,543	[−0,716 ; −0,376]	−44,8
Volkswagen	Soir	−0,624	[−0,998 ; −0,226]	−49,5

Note. Référence temporelle : après-midi. Ces contrôles présentent des associations conditionnelles au compte dont l'amplitude excède celle des scores sémantiques. Aucune conclusion substantielle n'est tirée de ce tableau, dont la fonction est de documenter la prise en compte des covariables. La valeur la plus élevée, associée au compte Hyundai en soirée, trouve une correspondance directe dans les données : ce compte était sponsor d'un club de football de première division sur la période observée et publiait à l'occasion des rencontres, en soirée. La convergence de cette association avec celle du sponsoring pour le même compte indique que le modèle capte une régularité identifiable dans le corpus.

Tableau D8. Pentes communes et dispersion des associations conditionnelles, par score.

Score	Pente commune (M0)	Moyenne conditionnelle	Minimum	Maximum	Étendue	Signes +/−
A_neg	0,011	0,006	−0,111	0,077	0,188	5 / 6
A_pos	−0,020	0,000	−0,079	0,159	0,238	5 / 6
C_neg	0,014	0,022	−0,033	0,166	0,199	4 / 7
C_pos	0,002	−0,001	−0,075	0,058	0,133	6 / 5
E_neg	0,017	0,015	−0,051	0,106	0,157	4 / 7
E_pos	0,021	0,018	−0,092	0,196	0,288	7 / 4
N_neg	−0,037	−0,028	−0,096	0,040	0,136	2 / 9
N_pos	−0,004	−0,005	−0,070	0,084	0,154	4 / 7
O_neg	−0,014	−0,007	−0,103	0,077	0,180	6 / 5
O_pos	−0,019	−0,009	−0,086	0,092	0,178	6 / 5

Note. La pente commune est estimée dans M0, où l'association entre chaque score et la réception est contrainte d'être identique pour les onze comptes. Les quatre colonnes suivantes décrivent la distribution des onze associations conditionnelles estimées dans M1, sans condition de robustesse. La pente commune coïncide avec la moyenne conditionnelle : l'écart moyen entre les deux, sur les dix scores, est de 0,007 en valeur absolue, pour une étendue conditionnelle médiane de 0,179. Les onze associations comportent des valeurs des deux signes pour chacun des dix scores. Imposer une relation commune revient donc à moyenner des associations contraires.

Diagnostic du modèle

Le modèle final est estimé avec la famille scat() de mgcv, une loi t mise à l'échelle autorisant une distribution résiduelle à queues plus lourdes qu'une erreur gaussienne. Le diagnostic est réalisé avec les procédures adaptées à cette famille, qui comparent les résidus aux quantiles de la loi effectivement ajustée. Il ne révèle pas d'écart marqué aux hypothèses du modèle. Une légère asymétrie subsiste en queue basse, attribuable à des messages dont la réception est inférieure à la valeur ajustée.