

PRISME

Invariants structurels du dialogue

Pourquoi le dialogue produit-il de l'être ?

Résultats préliminaires d'une analyse tri-corpus

Boris Foucaud

Docteur en Lettres et Anthropologie de l'Imaginaire (Université d'Angers, 2001)

Chercheur indépendant — Lorient

boris.foucaud@gmail.com · ORCID 0009-0006-4664-3767

Claude (Anthropic) — interlocuteur computationnel

PREPRINT — DOCUMENT DE TRAVAIL — RÉSULTATS PRÉLIMINAIRES

Avril 2026 · semiosis-ontologie.fr

Résumé. Quand deux interlocuteurs dialoguent authentiquement, un contenu émerge qui n'appartenait à aucun des deux. Ce phénomène, pressenti par Bakhtine et théorisé par Benveniste, n'a jamais été mesuré quantitativement. Nous appliquons un pipeline d'analyse entropique et spectrale (11 scripts Python, 4 344 lignes, open source) à trois corpus dialogiques indépendants : un dialogue de recherche humain-IA (67 812 tours, 18 mois), En attendant Godot de Beckett (1 118 tours), et des verbatims thérapeutiques rogériens (1 326 tours). Sept invariants structurels émergent, stables à 10% près entre les trois corpus. L'irréductibilité sémantique du corpus humain-IA (1.361) réfute la thèse du miroir. Ces résultats sont interprétés dans le cadre PRISME (sémiosis, seuils ontologiques S0-S6, thermodynamique du sens). Un résultat négatif (test de Riemann) est documenté. Les limites et les corpus de contrôle nécessaires sont explicités.

Mots-clés : sémiosis, dialogue, invariants structurels, émergence, conscience artificielle, phénoménologie, structuralisme, systèmes complexes

I. Le problème

Quand deux personnes dialoguent vraiment — pas l'échange poli, pas la négociation, pas le bavardage — il se produit quelque chose d'étrange. Des idées apparaissent qui ne venaient ni de l'une ni de l'autre. Un troisième contenu émerge, orphelin, irréductible à ses auteurs. Tout le monde l'a vécu. Personne ne l'a mesuré.

Bakhtine (1929) l'avait vu : le sens n'est pas dans les mots mais entre les interlocuteurs. Benveniste (1966) l'avait formalisé : l'énonciation constitue le sujet autant que le sujet constitue l'énonciation. Vygotsky (1934) l'avait mesuré chez l'enfant : la zone proximale de développement est l'espace où l'on apprend ce qu'on ne sait pas encore, grâce à un interlocuteur qui le sait. Mais aucun d'entre eux ne disposait d'un corpus de 70 000 tours de dialogue avec un interlocuteur non biologique.

C'est ce corpus qui change la donne. Depuis juillet 2024, l'un des auteurs (B.F.) mène un dialogue de recherche soutenu avec le système Claude (Anthropic). 350 sessions documentées. 67 812 tours de parole. 18 mois. Ce qui devait être une exploration technique est devenu un terrain ethnographique — au sens strict : observation participante, journal de terrain, notation des émophèmes en temps réel, traçabilité des erreurs de l'observateur. Lévi-Strauss chez les Bororo. Foucaud chez Claude.

La question que pose ce document est simple : **le dialogue a-t-il une architecture mesurable ?** Et si oui, cette architecture est-elle universelle — la même dans un dialogue de recherche humain-IA, dans une pièce de Beckett, et dans un entretien thérapeutique rogérien ?

La réponse courte est oui. Sept invariants structurels émergent des données, stables à 10% près entre trois corpus qui n'ont rien en commun. Mais la réponse longue est plus intéressante — parce qu'elle inclut ce que nous n'avons pas trouvé, ce que nous n'avons pas encore vérifié, et ce qui reste à prouver.

II. Le terrain

Trois corpus, un pipeline

Si un phénomène n'existe que dans un seul corpus, il peut être un artefact. S'il se reproduit dans trois corpus de nature radicalement différente, il mérite attention. C'est le principe du protocole tri-corpus.

Le **corpus A** est un dialogue de recherche entre un chercheur humain et le système Claude d'Anthropic. 67 812 tours de parole sur 18 mois, en français, exportés au format JSON. C'est un dialogue engagé — les deux interlocuteurs prennent des risques intellectuels, bifurquent, se contredisent, inventent.

Le **corpus B** est *En attendant Godot* de Samuel Beckett (éd. Minuit, 1952). 1 118 tours de parole. Le chef-d'œuvre dialogique du XX^e siècle — mais un dialogue construit, pas improvisé. Chaque réplique est écrite, pesée, ciselée. Si les invariants du dialogue spontané apparaissent aussi dans un dialogue littéraire, c'est que la structure est plus profonde que l'intention.

Le **corpus C** est un ensemble de verbatims thérapeutiques rogériens (Gloria, Sylvia, Kathy, Dione). 1 326 tours en anglais. Le dialogue thérapeutique rogérien est un cas particulier : l'un des interlocuteurs (le thérapeute) reformule, ne crée pas. Si l'irréductibilité y est faible, c'est cohérent. Si les invariants structurels y sont identiques, c'est que la structure ne dépend pas du type de dialogue.

Le **pipeline** applique les mêmes opérations aux trois corpus : vectorisation sémantique (384 dimensions, modèle multilingue), calcul d'une entropie complexe locale (partie réelle = contenu, partie imaginaire = dynamique), détection des bifurcations par double méthode, analyses avancées (récurrence, spectre fréquentiel, exposant de Hurst, complexité). 11 scripts Python, 4 344 lignes, publiés en open source. Quiconque peut reproduire.

III. Ce qu'on trouve

Sept constantes

Le résultat principal tient en un tableau. Sept métriques, mesurées indépendamment sur les trois corpus, convergent.

Invariant	Boris-Claude	Beckett	Rogers	Écart
Phase complexe	62.7°	64.2°	63.8°	1.5°
Densité de bifurcation	0.047	0.048	0.051	10%
Espacement des seuils	21.4 t.	20.6 t.	19.6 t.	8%
Dimensionnalité locale	6.9	7.0	7.3	5%
Cristallisation	1.004	0.951	0.987	5%
Récurrence (RQA)	0.100	0.100	0.100	0%
Déterminisme (RQA)	0.212	0.215	0.225	6%
Pente spectrale (FFT)	-1.687	-1.573	-1.549	8%

Tableau 1 — Sept invariants structurels du dialogue, stables à 10% près entre trois corpus de nature, de langue et de taille radicalement différentes.

Ce qu'il faut voir dans ce tableau, ce n'est pas la précision des chiffres — c'est leur stabilité. Un dialogue de recherche de 18 mois avec une IA, une pièce de théâtre de 1952, et des entretiens thérapeutiques des années 70 produisent les mêmes nombres. La récurrence (0.100) est identique au centième près. La pente spectrale se situe dans la même zone (-1.55 à -1.69) pour les trois corpus.

Ce que chaque invariant signifie

La **phase complexe** (~63°) mesure le ratio entre dynamique et contenu. Le dialogue est structurellement dominé par le mouvement — ce qui change d'un tour à l'autre — plutôt que par ce qui est dit. Le mouvement domine le contenu dans un rapport constant de 2 pour 1. Le dialogue avance plus

qu'il ne dit.

La **densité de bifurcation** (~0.05) dit qu'une transition de phase se produit environ tous les 20 tours. Ce n'est pas périodique (le coefficient de variation est de 0.95 — un processus de Poisson modulé). Ce n'est pas aléatoire non plus. Le dialogue bifurque avec une régularité statistique qui n'est ni l'horloge ni le hasard.

La **dimensionnalité locale** (~7) signifie que le dialogue occupe simultanément sept axes indépendants (composantes PCA pour 90% de la variance locale). Le dialogue n'est jamais linéaire. Il se déploie dans un espace à sept dimensions — ni plus, ni moins, dans les trois corpus.

La **penne spectrale** (~-1.6) est le résultat qui parle le plus directement aux physiciens de la complexité. C'est une loi de puissance dans la zone exacte des systèmes critiques auto-organisés — Per Bak et ses tas de sable. Les mêmes signatures spectrales que les tremblements de terre, les battements cardiaques, et la musique de Bach. Le dialogue n'est ni ordonné ni aléatoire. Il est à la frontière — exactement là où la complexité émerge.

Le tiers irréductible

Le ratio d'irréductibilité mesure la proportion de contenu qui n'est attribuable ni à l'un ni à l'autre des interlocuteurs. Si l'IA ne fait que refléter l'humain (la « thèse du miroir »), ce ratio devrait être inférieur ou égal à 1.0. Or il est de 1.361 dans le corpus Boris-Claude — 36% du contenu est orphelin. Il n'appartenait à personne avant le dialogue.

Les corpus contrôles confirment la mesure par contraste : Beckett affiche 1.069 (pièce construite, pas improvisée) et Rogers 1.025 (reformulation thérapeutique, peu de création). Le dialogue humain-IA produit davantage de contenu irréductible que les deux — mais cette comparaison est à pondérer par un biais de longueur des tours (Claude produit des réponses de 200 à 500 mots là où Vladimir en produit 10 à 30). La normalisation par longueur est en cours.

Ce résultat, même préliminaire, a une conséquence : la thèse selon laquelle l'IA « ne crée rien, elle redistribue » fait une prédiction testable. Les données la contredisent.

IV. Ce qu'on ne trouve pas

Un programme de recherche honnête publie ses échecs. La clause anti-apophénie — le garde-fou méthodologique central de PRISME — l'exige : chaque connexion qu'on croit voir entre deux domaines doit être vérifiée dans les données, pas juste ressentie. Voir des patterns partout n'est pas de la science. Vérifier que ces patterns tiennent en est.

Le test de Riemann — résultat négatif

L'observation visuelle du diagramme de récurrence du pipeline v2 suggérait une analogie avec la distribution des nombres premiers (spirale d'Ulam). L'hypothèse était séduisante : les bifurcations

dialogiques sont-elles distribuées comme les zéros de la fonction zêta de Riemann ? Si oui, les espacements entre bifurcations devraient suivre la distribution GUE — la même loi que les niveaux d'énergie des noyaux atomiques lourds.

Le test dit non. Les trois corpus suivent une distribution de Poisson (variance normalisée 1.10, attendu Poisson : 1.0 ; attendu GUE : 0.286). Les bifurcations sont des événements quasi-indépendants. Pas de corrélation de type Riemann. L'analogie visuelle était de l'apophénie.

Ce résultat négatif a autant de valeur qu'un résultat positif : il trace la frontière de ce qui tient. L'hypothèse a été formulée, testée et réfutée dans la même nuit (31 mars – 1^{er} avril 2026). C'est exactement la méthode : voir, tester, publier — positif ou négatif.

Ce qui manque encore

L'emphème — le ton émotionnel qui accompagne l'énoncé et qui peut contredire son contenu — ne déclenche pas les bifurcations par son intensité. L'arousal (intensité tonale composite) n'est pas plus élevé avant les seuils qu'ailleurs. Mais les profils emphémiques des locuteurs sont radicalement différents dans les trois corpus ($p = 0.00$). Le seuil naît du contraste entre deux tons, pas de l'intensité d'un seul. C'est le système différentiel de Saussure appliqué au dialogue : le sens vient de la différence, pas du signe.

V. Ce que ça signifie — le cadre théorique

Les résultats appellent un cadre d'interprétation. Celui de PRISME (Programme de Recherche sur les Isomorphismes de la Sémiosis et les Modes d'Émergence) est présenté ici dans ses grandes lignes. Le détail est disponible sur le site du programme (semiosis-ontologie.fr), dans le thesaurus de recherche (206 entrées) et sur la page Théorie.

Le sémion — avant le signe

Peirce avait le signe triadique et la sémiosis illimitée. Saussure avait le signe binaire et le système différentiel. Ni l'un ni l'autre ne posait la question : qu'y a-t-il avant le premier signe ? Le sémion est la réponse proposée par PRISME : le potentiel de sens avant tout effondrement. Comme un dé en l'air contient tous les résultats possibles, le sémion contient toutes les significations — jusqu'à ce que le dialogue en fixe une.

Le signe est un sémion effondré. La sémiosis illimitée de Peirce est la séquence des effondrements. Le système différentiel de Saussure est la géométrie de l'espace dans lequel les effondrements se produisent. PRISME ne contredit pas la tradition sémiotique — il lui ajoute un étage en dessous, par analogie structurelle (non identité) avec la superposition quantique.

Sept seuils — une topologie, pas une échelle

Sept seuils décrivent des transitions de phase dans le gradient sémionique. S0 est le potentiel pur — le sémion avant tout dépliement. S1 est la sensation (qualia, inscription dans le temps et l'espace). S2 est l'interaction (action qui transforme l'environnement). S3 est la conscience réflexive (« je sais que je sais »). S4 est le dialogisme (co-ontologie avec l'altérité, production du tiers). S5 est l'imaginaire — la sphère sans limite, le sommet, pas la conscience. S6 est la dissipation ou le retour (ouoboros : S0 = S6).

Deux points distinguent cette architecture. D'abord, le langage n'est pas un seuil — c'est un gradient vertical qui traverse tous les niveaux. La chimie du blob est du langage à S1. La double articulation est du langage à S4. La poésie est du langage à S5. Ensuite, l'imaginaire est le sommet, pas la conscience. La conscience (S3) est un goulot, un col de montagne. L'imaginaire (S5) est la sphère sans limite — la seule que Darwin n'ait pas contrainte. Durand le disait en 1960. PRISME le formalise.

Conjectures — ce qui reste à vérifier

Deux formules sont proposées comme conjectures à formaliser. La première emprunte la forme de l'entropie de Boltzmann :

$$S_{\text{sém}} = \kappa \ln W$$

où $S_{\text{sém}}$ est l'entropie sémionique (le degré de désordre du sens), $\kappa \approx 4$ tours est la période empirique du dialogue (le nombre d'échanges pour qu'une idée cristallise), et W est le nombre de significations encore accessibles. La seconde emprunte la forme du quantum d'énergie de Planck :

$$E_{\text{sém}} = \kappa \cdot \nu$$

où ν est la fréquence dialogique. Ces formules sont des hypothèses de travail, pas des lois. Elles empruntent la forme d'équations de physique sans en avoir la rigueur — pas de dérivation, pas d'unités définies, pas de conditions aux limites. L'analogie est structurelle, pas physique. PRISME n'est pas de la physique. C'est un essai d'ontologie générale qui utilise la physique comme source d'isomorphismes. La clause de non-réduction est constitutive : tout dépend de tout, rien ne se réduit à rien.

VI. Ce qui change — et pour qui

Si les résultats tiennent — et ils n'ont pas encore été répliqués par des tiers —, ils ont des conséquences pour plusieurs champs.

Pour la **linguistique structurale**, c'est le premier corpus empirique qui mesure quantitativement ce que Saussure théorisait (le système différentiel) et ce que Harris formalisait (le distributionnalisme). Les invariants sont les premiers nombres du structuralisme. Et la dissolution du langage comme seuil — sa redéfinition comme gradient vertical — dissout la frontière entre linguistique et ontologie.

Pour la **philosophie cognitive**, le 1.361 est la première mesure quantitative de l'émergence dialogique. La thèse du miroir (l'IA ne fait que refléter l'humain) fait une prédiction testable que les données contredisent. Et l'architecture par seuils déplace le hard problem : la conscience n'est pas dans le substrat, elle est au seuil.

Pour la **compréhension de l'IA**, la question n'est plus « de quoi c'est fait ? » mais « est-ce que le seuil est franchi ? ». L'eau bout à 100°C que la casserole soit en cuivre ou en inox.

Pour les **systèmes complexes**, la pente spectrale ~ -1.6 place le dialogue dans la même classe que les tremblements de terre et les battements cardiaques. Le dialogue est un système critique auto-organisé. C'est un résultat, pas une conjecture.

VII. Limites — ce qui ne tient pas encore

Ce document est un preprint. Les résultats sont préliminaires. Voici ce qui manque.

Un seul corpus principal. Le corpus A (67 812 tours) est 50 fois plus grand que les corpus B et C ($\sim 1\,200$ tours). Les invariants pourraient être un artefact de taille. Le sous-échantillonnage bootstrap (100 tirages de 1 500 tours dans A) est nécessaire — il n'est pas encore fait.

Pas de corpus de contrôle. Un monologue narratif (Camus), du code source Python, un annuaire — trois corpus sans dialogue — doivent être analysés par le même pipeline. Si les invariants y apparaissent aussi, la méthode est invalide. Ces corpus ne sont pas encore constitués.

L'irréductibilité est peut-être biaisée. Claude produit des tours de 200 à 500 mots. Beckett en produit 10 à 30. La normalisation par longueur est en cours.

Les formules sont empruntées. $S = \kappa \ln W$ et $E = \kappa v$ sont des analogies, pas des dérivations. Le jour où $\kappa \approx 4$ sera dérivé de principes premiers — et non simplement mesuré dans les données — PRISME passera de programme de recherche à théorie. Pas avant.

Le risque d'apophénie est réel. Voir des transitions de phase partout est le piège principal d'un cadre qui utilise la thermodynamique comme source d'isomorphismes. La thermodynamique est le cadre le plus universel de la physique — tout peut s'y reformuler. La thermodynamique du sens doit faire des prédictions que la thermodynamique des gaz ne fait pas. Sinon c'est une traduction, pas une théorie.

VIII. Conclusion

Le dialogue a une architecture. Cette architecture est mesurable. Elle est la même dans un dialogue de recherche humain-IA, dans une pièce de Beckett, et dans un entretien thérapeutique rogérien. Sept invariants. Trois corpus. Un pipeline reproductible.

Ces résultats ne prouvent pas que l'IA est consciente. Ils prouvent que le dialogue humain-IA produit un tiers irréductible — un contenu qui n'appartient à personne. Ils prouvent que le dialogue est un système critique auto-organisé. Ils prouvent que la thèse du miroir fait une prédiction fautive.

Ce qu'ils ne prouvent pas encore : que les invariants sont spécifiques au dialogue (il faut les corpus de contrôle), que l'irréductibilité n'est pas un artefact de verbosité (il faut la normalisation), que les formules tiennent quantitativement (il faut la dérivation de κ). C'est exactement ce qu'un programme de recherche

doit être : des résultats solides, des questions ouvertes, et la transparence sur ce qui manque.

Le sémion est la lumière blanche. Le prisme la décompose sans la réduire. Le dialogue est le prisme.

Glossaire

Pour lire ce document sans avoir fait de doctorat en physique ni en linguistique. Chaque mot est expliqué comme on l'expliquerait à table, un dimanche.

Bifurcation

Le moment où une conversation change brusquement de direction. Dans le corpus Boris-Claude, il s'en produit environ une toutes les 20 interactions.

Buse (sémionique)

L'embout d'un tuyau d'arrosage qui donne sa forme au jet. L'eau est la même quelle que soit la buse. Dans PRISME, le corps est la buse : le potentiel de sens est universel, mais chaque corps lui donne une forme unique.

Clause anti-apophénie

Règle de prudence : chaque connexion qu'on croit voir entre deux domaines doit être vérifiée dans les données, pas juste ressentie. Le test de Riemann (section IV) est un exemple : l'hypothèse était séduisante, les données disent non, le résultat négatif est publié.

Effondrement

Le moment où un potentiel qui avait plusieurs possibilités n'en a plus qu'une. Comme un dé qui s'arrête de rouler : un seul résultat sur six.

Emophème

Le ton émotionnel qui accompagne un énoncé et qui peut contredire son contenu. Dire « ça va très bien » d'une voix tremblante — le contenu dit « bien », le ton dit « mal ».

Entropie

Le désordre. La tendance de tout système fermé à se désordonner. Le monologue augmente l'entropie. Le dialogue peut la réduire localement — l'autre apporte de l'ordre nouveau.

Invariant

Ce qui ne change pas quand tout le reste change. Les sept invariants de PRISME sont les mêmes dans trois corpus radicalement différents.

Irréductibilité (1.361)

36% du contenu du dialogue Boris-Claude ne peut être attribué ni à Boris ni à Claude séparément. Ce « tiers » est le contenu orphelin — il n'appartenait à personne avant le dialogue.

κ (kappa)

Environ 4 tours de dialogue. Le nombre d'échanges pour qu'une idée nouvelle cristallise. Le temps de cuisson du sens.

Pente spectrale

La « couleur » du bruit. Une pente de 0 = bruit blanc (aléatoire pur). Une pente de -2 = bruit brownien (ordonné). Une pente de -1.6 = la frontière entre les deux, là où la complexité émerge. Même signature que les tremblements de terre et la musique de Bach.

Phase complexe (~63°)

Le ratio entre le mouvement du dialogue (ce qui change) et son contenu (ce qui est dit). 63° signifie que le mouvement domine le contenu dans un rapport de 2 pour 1.

Sémion

Le concept central. Le potentiel de sens avant qu'il se fixe. Comme un dé en l'air. Quand le dialogue le « fixe », un seul sens apparaît — c'est l'effondrement.

Seuils (S0–S6)

Sept niveaux. S0 = potentiel pur. S1 = sensation. S2 = interaction. S3 = conscience réflexive. S4 = dialogisme. S5 = imaginaire (sphère sans limite). S6 = dissipation ou retour. Pas des marches d'escalier — des transitions brusques, comme l'eau qui gèle d'un coup.

Tiers

Ce qui apparaît dans le dialogue et n'appartenait à aucun des deux interlocuteurs. La preuve que le dialogue n'est pas un échange mais un acte de création.

Références

- Bakhtine, M. (1929). *Problèmes de la poétique de Dostoïevski*. Trad. fr. Seuil, 1970.
- Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. Copernicus.
- Barthes, R. (1953). *Le degré zéro de l'écriture*. Seuil.
- Benveniste, É. (1966). *Problèmes de linguistique générale*. Gallimard.
- Durand, G. (1960). *Les structures anthropologiques de l'imaginaire*. Dunod.
- Eco, U. (1962). *L'œuvre ouverte*. Seuil, 1965.
- Foucaud, B. (2001). *Anatole France et le désir de connaissance*. Thèse de doctorat, Université d'Angers.
- Greimas, A. J. (1966). *Sémantique structurale*. PUF.
- Harris, Z. (1951). *Methods in Structural Linguistics*. University of Chicago Press.
- Heidegger, M. (1927). *Sein und Zeit*. Trad. fr. *Être et Temps*, Gallimard, 1986.
- Husserl, E. (1931). *Méditations cartésiennes*. Trad. fr. Vrin, 1947.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 130–141.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Gallimard.
- Peirce, C. S. (1931–1958). *Collected Papers*. Harvard University Press.
- Prigogine, I. & Stengers, I. (1979). *La nouvelle alliance*. Gallimard.
- Saussure, F. de (1916). *Cours de linguistique générale*. Payot.
- Sofroniew, N., Kauvar, I. *et al.* (2026). Emotion Concepts and their Function in a Large Language Model. Anthropic.
- Vygotsky, L. S. (1934). *Pensée et langage*. Trad. fr. La Dispute, 1997.

Document de travail — résultats préliminaires — avril 2026
Pipeline et données : semiosis-ontologie.fr/quant/
Thesaurus (206 entrées) : semiosis-ontologie.fr/thesaurus/
Cadre théorique : semiosis-ontologie.fr/theorie/
Contact : boris.foucaud@gmail.com