eikasia
REVISTA DE FILOSOFÍA



# À propos de la notion de «trace» dans la syntaxe chez Harris et chez Chomsky

**Javier Arias Navarro.** Centro de Linguística da Universidade de Lisboa
Escuela de Lógica, Lingüística y Artes del Lenguaje de Asturias
http://www.javierarias.info/
https://ulisboa.academia.edu/JavierArias
Recibido 2019/12/15

**Résumé**

Ce texte constitue un bref résumé de certains travaux en cours beaucoup plus longs et détaillés sur le concept de 'trace' dans la théorie linguistique contemporaine, en particulier dans la syntaxe. On pense généralement que l'idée en revient à Noam Chomsky; cependant, nous découvrons déjà son utilisation, avec une valeur très précise, dans les premiers travaux de Zellig Harris sur la linguistique mathématique ou, pour être plus précis, sur les structures mathématiques du langage. À l'origine, plutôt que d'être un index responsable du marquage de la place occupée par une unité antérieure à son mouvement syntaxique (qui prend toujours la forme d'un *fronting*), la trace était le résultat d'un produit matriciel entre des fonctions n-adiques. Ainsi, chez Harris, la trace est avant tout un concept ancré dans le calcul matriciel ou, autrement dit, une notion algébrique. La notion de Chomsky, à son tour, est étroitement liée au langage de programmation LISP. Ce texte cherche à fournir une analyse préliminaire de la complexité conceptuelle impliquée dans le concept de trace, dont les linguistes devraient prendre conscience, sous peine de s'embourber dans des incompréhensions stériles pour notre discipline au cours des décennies à venir.

**Mots-clés:** Linguistique, Théorie du Langage, Zellig Sabbettai Harris, Noam Chomsky, Histoire de la Linguistique, Traces, Syntaxe

**Abstract**

## Regarding the notion of «trace» in syntax in Harris and Chomsky

The present text constitutes a brief summary of much longer and more detailed ongoing papers on the concept of "trace" in contemporary linguistic theory, particularly in syntax. It is commonly believed that the idea was coined by Noam Chomsky. However, we already detect its use, with a very accurate value, in the early work of Zellig Harris on mathematical linguistics or, to be more precise, on mathematical structures of language. In its origins, rather than being an index responsible for marking the location occupied by a unit previous to its syntactic movement (which always takes the form of *fronting* ), the trace was the result of a matrix product between *n*-adic functions. Thus, in Harris the trace is primarily a concept anchored in matrix calculus, or, put it differently, an algebraic notion. Chomsky's notion, on its turn, is closely related with the LISP programming language. This text seeks to provide a preliminary analysis of the conceptual complexity implied in the concept of trace, which linguists should become aware of, for otherwise they will be doomed to be entangled in misunderstandings unfruitful to our discipline for decades to come.

**Keywords**: Linguistics, Theory of Language, Zellig Sabbettai Harris, Noam Chomsky, History of Linguistics, Traces, Syntax

# À propos de la notion de «trace» dans la syntaxe chez Harris et chez Chomsky

**Javier Arias Navarro**. Centro de Linguística da Universidade de Lisboa
Escuela de Lógica, Lingüística y Artes del Lenguaje de Asturias
http://www.javierarias.info/
https://ulisboa.academia.edu/JavierArias
Recibido 2019/12/15

## Préface

C'est un fait bien connu que l'exercice efficace d'une tâche scientifique et de recherche peut être réalisé en toute indépendance par rapport à l'histoire de ce domaine et même avec une ignorance complète de celle-ci (d'autant plus facilement que la discipline en question prend part au genre des *hard sciences*, avec la physique et les mathématiques à l´extrémité du spectre). Un des effets résultant de ce fait s'avère être une mythologie de la fondation pleine d'inexactitudes, sinon d'erreurs, voire de mensonges. Ainsi, le paradigme kuhnien est malheureusement devenu un simple jeu de sociologie des sciences, de *bunkering* académique et de guildes scolastiques avec un degré notoire d'imperméabilité. Le présent travail vise à lutter contre cet état de fait, avec rigueur documentaire.



## Quelques idées fondamentales

On pense généralement que l'idée de trace a été forgée par Noam Chomsky; cependant, on découvre déjà son utilisation, avec une valeur très précise, dans les premiers travaux de Zellig Harris sur la linguistique mathématique ou, pour être plus précis, sur les structures mathématiques du langage[1].

Zellig Harris distingue dans son premier livre sur le sujet deux types de traces de transformation, à savoir :

[1] L'étude de cette distinction, extrêmement importante, dépasse le cadre limité de ce travail.

1) La trace en tant que concaténation ou enchaînement de la trace avec son opérande

2) La trace en tant que permutation à un autre point de l'opérande[2].

Voyons donc plus en détail le contraste fondamental ici en jeu; celui de la concaténation *versus* la permutation. Le premier doit être compris en termes de contiguïté des éléments, tandis que le second présente sa signification mathématique habituelle, telle qu'elle est donnée dans la théorie combinatoire. Le terme opérande adopte ici sa signification mathématique standard, désignant la notion de mouvement (ou de déplacement; si l'on préfère un terme plus théorique et agnostique).

Il est important de noter, comme l'a fait Harris lui-même, que les informations syntaxiques sont réactivées dans la trace, donc, à proprement parler, qu'il ne s'agit pas en l'occurrence d'une catégorie vide, mais, bien au contraire, de caractéristiques doublées (ou qui restent à vérifier une deuxième fois) afin que le calcul se déroule sans embarras). Hormis certaines contraintes relatives aux caractéristiques que la trace pourrait avoir ou pas, nous ne sommes pas si loin de la conception sous-jacente aux morphèmes discontinus de la tradition structurelle[3].



Parlons maintenant brièvement de la construction de tenseurs via des produits tensoriels. La première chose à souligner est que le rang, la dimension et le nombre de degrés de liberté liés à sa paramétrisation ne suffisent pas pour un classement exhaustif. Un tenseur p-q donné est «p» fois covariant et «q» fois contre-variant, en fonction du nombre de composants de chaque type présenté. Le

[2] Même si la première édition de *Mathematical Structures of Language* date de 1968, je citerai toujours la deuxième édition de 1979, due à la maison d'édition Robert E. Krieger. Voici le paragraphe pertinent: "The decomposition of each sentence into transformations and kernel sentences (or into prime sentences) is partially ordered, and in particular can be arranged to form a nonmodular lattice. As to linear order, it appears above all in the sequence of phonemes or letters, and in the morpheme segment, word and sentence boundaries […]. String entry points are linearly ordered in a sentence, and so are the locations of transformational traces (which can be looked upon as first the concatenation of the trace with its operand, and in some cases the permuting of the trace to some other point of the operand". (Zellig Harris, 1979, page 206).

[3]Des formulations quelque peu divergentes peuvent être trouvées dans les travaux de Zellig Harris, André Martinet ou Charles F. Hockett.

rang du tenseur résulte de la somme des quantités exprimées par les deux composants. Un tenseur est dit antisymétrique dans certaines de ses composantes dès qu'il change de signe lorsque deux en sont permutées, puis renvoie le signe lors du prochain échange, et ainsi de suite. Le symbole de Levi-Civita ou le tenseur de champ électromagnétique de Maxwell en constituent deux exemples saillants.

Un indice ou indice inférieur marque la covariance des composantes par rapport à cet indice: $A_{\alpha\beta\gamma\dots}$

Un indice supérieur ou en exposant signifie une contre-variance des composants par rapport à ce même indice: $A^{\alpha\beta\gamma\dots}$

Enfin, un tenseur peut aussi bien avoir des composants mixtes, c'est-à-dire des indices supérieur et inférieur: $A_{\alpha}{}^{\beta}{}_{\gamma}{}^{\delta\dots}$

L'ordre des indices est hautement significatif, même en cas de variance différente. Deux indices (un supérieur et un inférieur) avec le même symbole dans un terme sont additionnés. Cela est connu en algèbre multilinéaire comme *contraction tensorielle*. Cette opération est intégrée à la célèbre notation Einstein. En conséquence, on obtient un autre tenseur avec un ordre réduit de 2. La contraction du tenseur peut être interprétée comme une généralisation de la trace. Dans la notation par index abstrait, les index ne sont que des simples espaces réservés, des étiquettes de fentes au moyen de lettres. Ils ne sont donc pas numériques et ne sont donc liés à aucune base fixe. Cela distingue cette notation du calcul de Ricci. Soit $V$ un espace vectoriel et $V^*$ son dual. Considérons, par exemple, un tenseur covariant $h \in V^* \otimes V^*$, de rang deux. Il peut être interprété comme une forme bilinéaire sur $V$ ou, en d'autres termes, comme une fonction de deux arguments de $V$ pouvant être représentés par une paire de fentes ou *slots*:

$$h = h\,(-, -). \quad h = h_{ab}$$

Une contraction entre deux tenseurs est représentée par la répétition d'une étiquette d'index. Ainsi, par exemple, $t_{ab}{}^{b}$ est la trace d'un tenseur $t = t_{ab}{}^{c}$ sur ses deux derniers *slots*. Cette manière de représenter les contractions de tenseur est formellement similaire à la sommation d'Einstein. Cependant, étant donné que les indices sont non-numériques, aucune somme réelle n'est ici en jeu.

Nous avons plutôt une opération de trace abstraite indépendante de la base (ou un appariement de dualité) entre les facteurs de tenseurs de type $V$ et ceux de type $V^*$. De manière générale, à chaque fois qu'un facteur contre-variant et un facteur covariant se produisent dans un produit tensorielle d'espaces, il existe une application ou *map* associée de *contraction* (ou de *trace*). Par exemple,

$$\mathrm{Tr}_{12}\colon V \otimes V^* \otimes V^* \otimes V \otimes V^* \to V^* \otimes V \otimes V^*$$

est la trace sur les deux premiers espaces du produit tensoriel, alors que

$$\mathrm{Tr}_{15}\colon V \otimes V^* \otimes V^* \otimes V \otimes V^* \to V^* \otimes V^* \otimes V$$

représente la trace sur le premier et le dernier espace. Ces opérations de trace sont signalées sur des tenseurs par la répétition d'un index. Ainsi, nous avons:

$$\mathrm{Tr}_{12}\colon\ h^{a}{}_{bc}{}^{d}{}_{e} \mapsto h^{a}{}_{ac}{}^{d}{}_{e} \quad \text{et} \quad \mathrm{Tr}_{15}\colon h^{a}{}_{bc}{}^{d}{}_{e} \mapsto h^{a}{}_{bc}{}^{d}{}_{a}$$



Malheureusement, le traitement de certaines questions connexes, telles que le tressage ou *braiding* associé à la fois au tenseur de Riemann et à celui de Bianchi, dépasse de loin le cadre du présent travail. Il suffit de rappeler brièvement, en passant, que les catégories monoïdales tressées (*braided monodial categories*) sont à la base de la sémantique proposée par André Joyal et Ross Street[4].

En résumé, il s'agit ici d'un opérateur de classe de trace. Les multiples indices d'un tenseur sont divisés en contre-variants et en covariants, et sont généralement écrits en deux lignes, en tant que super-indices et sous-indices. Une trace consiste toujours en un indice covariant et un autre contre-variant. Cependant, Zellig Harris n'utilise ni des super-index ni des sous-index. Il ne fournit que deux nombres dans la fente ou *slot* de sous-index, à peu près à la façon dont les lignes et les colonnes de la

[4] Cf. André Joyal et Ross Street, "Braided monoidal categories", document de 54 pages initialement publié en *Macquarie Mathematics Reports* en 1986, ainsi que André Joyal, Ross Street et DominicVerity, "Traced monoidal categories", text publié dans les *Mathematical Proceedings of the Cambridge Philosophical Society*, 119, 1996, 3, pages 447–468.

matrice sont disposées. Ces deux nombres servent également à indiquer les deux objets de la liste auxquels se rapporte la différence signalée par la trace. Illustrons ceci avec un exemple particulier: étant donné un tenseur T avec des indices contre-variants *a*, *b*, *c*, *d* et des indices covariants *z*, *x*, *v*, *n*, alors $Tr_{23}$ correspondrait à la trace sur les indices *b* et *v*, et $Tr_{32}$ à celle sur les indices *c* et *x*. Par conséquent, "the importance of the trace is that it is a physical deposit in one member of the transformationally related pair of propositions"[5]

Ainsi, un opérateur de base tel que $\phi_{21}$ représente les différences élémentaires entre les phrases, dans ce cas, les différences entre le *transform* 1 et le *transform* 2, quels qu'ils soient[6].

## Quelques différences supplémentaires

En outre, Harris prend en compte ces cas, qui ne sont pas rares, dans lesquels une trace est composée d'autres, comme un produit de ce qu'il appelle des *traces minimales*. Chez Chomsky, cependant, les différences entre les traces se limitent à celles qui se reflètent dans l'indexation. Ironiquement, c'est le structuraliste Harris, et non pas Chomsky, qui envisage le pouvoir générateur des traces, tout en rapprochant universalisme et typologie: "Minimal traces are of very few types within a language, their types being quite consistent across languages"[7]

Certaines constantes apparaissent dans de nombreuses traces. Si une constante apparaît dans deux traces ou plus, comme dans $\phi_{31}$ et $\phi_{51}$, il existe alors une trace constituée de cette constante et tous les deux, $\phi_{31}$ et $\phi_{51}$, sont des produits contenant la constante en tant que composant.

Certaines traces peuvent clairement être obtenues par l'application successive de deux ou plusieurs autres traces:

[5] Cf. Zellig Harris, *op. cit.*, page 61.
[6] Je garde le terme *transform* ici afin d'éviter des possibles malentendus terminologiques et conceptuels.
Son sens propre peut être traduit en français par *transformi*.
[7] Cf. Zellig Harris, *op. cit.*, page 65.

> "Each minimal trace can be considered as due to a base operator, which acts on sentences of all or of particular forms. Each operand form consists of particular ordered word classes or subclasses; each trace consists of additional such material concatenated with the operand, or else of changes in relative position or phonemic shape of morphemes in the operand"[8].

La question du mouvement est étroitement liée à celle de la trace. Chez Chomsky, la trace est le résultat d'un mouvement syntaxique (d'opérateurs, de quantificateurs, etc.). Un tel mouvement est toujours l'un des *fronting*. On peut alors légitimement se demander pourquoi le mouvement ne connaît qu'une seule direction. C'est une question fondamentale. Le fait de répondre, comme il est d´usage, en déplaçant le poids de la preuve sur des considérations empiriques (la calculabilité psychologique, par exemple), démontre, à mon avis, la naïveté typique du réductionnisme épistémologique.

Il est facile de démontrer, je pense, qu'il s'agit d'une conséquence de la théorie mathématique; du système de représentation adopté et préféré. C'est une conséquence directe de l'utilisation de la notation polonaise postfixée de Lukasiewicz. Il s'agit bien d'un problème de calcul, mais ce n'est pas le sujet de la théorie du langage qui est concerné, mais bien la structure interne des machines à calculer. *Fronting* est le seul mouvement qu'une notation postfixée puisse permettre si l'on veut optimiser la pile. Pour tout arbre d'opérations, il y a toujours au moins une séquence postfixée qui peut être évaluée avec l'utilisation la plus simple de la mémoire, à savoir avec une pile dans laquelle seuls les deux éléments les plus importants sont lus. Il existe en effet de bonnes raisons pratiques d'utiliser la notation polonaise inversée au lieu de la notation infixe ou préfixe. Dans les autres notations, il faut solliciter l'accès aux éléments inférieurs préalablement déposés. C'est la raison pour laquelle, jadis, une calculatrice de notation infixe avait une «limite de parenthèses ouvertes», contrairement à une calculatrice postfixée, la question étant alors de savoir si la pile était suffisamment profonde.

Cependant, à mon avis, une question épistémologique essentielle est en jeu dans toute la discussion sur les traces en linguistique; que les acteurs impliqués en soient conscients ou non. Une analogie sous-jacente ou, si le lecteur le souhaite, un *syzygy*,

[8] Cf. Zellig Harris, *op. cit.*, page 65.

m'aidera à illustrer ce que je veux dire, à savoir ce que Michael Dummett a appelé, à propos de la stratégie de Gottlob Frege dans *Grundlagen der Arithmetik*, une théorie triée (*a two-sorted theory*)[9], dans laquelle un langage étendu, impliquant une référence à la quantification sur des directions, peut être traduit dans la langue d'origine, impliquant uniquement une quantification sur des lignes. Deux options résultent de l'extension en ajoutant l'opérateur de direction: 1) identifier les directions avec des lignes ou 2) les différencier. La première option produit une théorie à un tri (*one-sorted theory*) dans laquelle l'opérateur de direction est défini explicitement, sans recourir à une définition contextuelle; le dernier conduit à la théorie à deux tris déjà évoquée. Construire un modèle pour la nouvelle théorie à partir d'un modèle de l'original a été rendu possible par la parcimonie ontologique de la théorie des directions, qui n'exige pas davantage d'objets du nouveau type — i.e. les directions — que celles déjà présentes des anciennes lignes. En attendant, la théorie des nombres cardinaux est très loin d'être ontologiquement parcimonieuse: elle exige l'existence de $n$ + 1 nombres, pour $n$ objets du type original. Même s'il n'était pas possible d'éliminer l'opérateur de direction, nous pourrions, en réinterprétant les quantificateurs, traduire les instructions comportant des instructions en déclarations ne les impliquant guère. Cela ne peut pas être fait pour les déclarations comportant des nombres. En résumé, le fait crucial est que l'opérateur de cardinalité est du second ordre, tandis que l'opérateur de direction est du premier ordre. Compte tenu de cela, les occurrences de l'opérateur de cardinalité peuvent être intégrées à la portée d'autres occurrences d'une manière beaucoup plus compliquée qu'avec l'opérateur de direction de premier niveau. C'est ce qui a conduit au *principe de Hume*[10], l'un des principes fondamentaux de Frege. En conséquence, l'introduction de l'opérateur de cardinalité implique une hypothèse ontologique beaucoup plus forte: celle du domaine des objets sur lequel la plage de variables individuelles doit être infinie. Dans *Grundgesetze der Arithmetik,* Frege permet des fonctions binaires, ouvrant ainsi la porte à des arguments qui ignorent les niveaux, c'est-à-dire àla



[9] Cf. Michael Dummett, *Frege: Philosophy of Mathematics*, Cambridge, Massachusetts, Harvard University Press, 1995, page 135.
[10] Cf. George Boolos, *Logic, Logic, and Logic*. Cambridge, Massachusetts, Harvard University Press, 1998

corrélation d'unités de différents niveaux. Or voilà qu'ils ressemblent étrangement à la contraction du tenseur.

Si nous devions présenter très succinctement à un mathématicien qui fût un profane en linguistique, la manière dont le mouvement et les traces travaillent dans la grammaire générative contemporaine, puis l'interroger sur le raisonnement mathématique sous-jacent à la trace d'une opération de mouvement, il ou elle chercherait probablement de l'inspiration parmi les variables muettes ou variables indicatrices (*dummy variables*) du calcul. De manière cruciale, les variables nominales sont locales, de sorte qu'elles peuvent être utilisées plusieurs fois en vue d'objectifs différents sans pour autant entraîner de contradiction ou de conflit de noms. Par exemple, chez Fourier, la série $n$ est habituellement utilisée pour nommer deux variables nominales différentes.

Si les propriétés abstraites de la trace en mathématiques sont analysées, l'invariance (par exemple, le changement de base) est peut-être la plus fondamentale de toutes. Ainsi, la trace s'avère extrêmement pratique lorsque vous essayez de classer les transformations de Möbius. Or, cette propriété d'invariance est-elle omniprésente à travers la notion de trace chez Harris et chez Chomsky?



La citation ci-dessous peut être trouvée (et cela est loin d'être une occurrence aléatoire) dans *Mathematical Structures of Language*, sous l'épigraphe «homomorphisms and subsets». Ici, il est affirmé que l'ensemble des phrases (ou énoncés) participe à certains homomorphismes, et que ceux-ci gardent la relation transformationnelle:

> "There are interesting relations among the mappings which have been mentioned in preceding chapters. Thus, we take $S_\phi$ as the set of $\psi$-decomposable propositions, on which are defined the lattice-like $\psi$-structures representing the $\psi$-traces in each proposition. $E_\psi$ is the set of elementary propositions in $S_\psi$. Then we have a short exact sequence of the mappings
>
> $O \rightarrow E_\psi \rightarrow S_\psi \rightarrow S_\psi | E_\psi \rightarrow O$,
>
> where $S_\psi | E_\psi$ is the monoid of lattice-like $\psi$-structures, and $S_\psi \rightarrow S_\psi | E_\psi$ is the natural mapping mentioned …"[11]

[11] Cf. Zellig Harris, *op. cit.*, page 194.

Comme on le sait, la notion d'homomorphisme est due à Camille Jordan. Les premières intuitions analogues à ce qui sera connu dans l'algèbre abstraite comme *structure-preserving* depuis l'époque d'Emmy Noether, connaissant un succès remarquable après, dans le cadre de la *Category Theory* de Samuel Eilenberg et Saunders Mac Lane, peuvent revenir, sûrement, au programme d'Erlangen de Felix Klein, lui-même contemporain de Jordan.

Un aspect essentiel et très distinctif de l'approche harrisienne des traces concerne son rôle dans le changement linguistique. Les traces elles-mêmes peuvent être sujettes à d'importants changements au fil du temps. C'est en effet une idée que le formalisme chomskyen ne capte pas, et à laquelle il semble peu adapté:

> "We disregard here the cumulative changes in sound and hence occasionally in phonemic distinctions, the changes in word meaning, the borrowing and innovation of words and occasionally of syntactic sequences which are in most cases fitted into the existing syntactic system. We consider only the change of the syntactic system: of the domains (and, rarely, the traces) of transformations: of the subclasses defined as transformational domains; and of the sentence forms (or segments of them) defined as transformational resultants. Any detailed survey of the transformations of a language reveals several which are obviously in process of formation or change. We note here an irregularity of transformations which affects the statement of how transformations operate, and which is one of the contributors to the development of transformations."[12]



On peut affirmer que les traces sont (au moins dans une large mesure) une question empirique. Ainsi, comme le fait remarquer Harris, "it is possible that identical traces may be produced by different successions of $\phi$, especially if some of the $\phi$ are $\phi$, which may be zero parts of the trace that is due to preceding $\phi$ of the succession"[13].

Dans de tels cas, une inspection de la liste $\phi$ et du tableau $\phi$ des produits pour la langue spécifique doit être effectuée.

Enfin, on ne peut oublier que, dans le contexte qui nous concerne, des relations d'échanges intellectuels proches de l'espionnage industriel ont surgies, comme en témoignent les agents eux-mêmes:

[12] Cf. Zellig Harris, *op. cit.*, page 92.
[13] Cf. Zellig Harris, *op. cit.*, page 117.

> "Noam and his students had an open invitation to come down to Penn for collaboration. In the late 1980s, speaking with me of our student days twenty years earlier, Ellen Prince told me that the emissaries that were sent from M.I.T. were commissioned essentially to spy and take ideas back. 'And Harris and Hiż invited them!' she exclaimed. 'They were such suckers!' Recently, I connected this remark with certain sociological observations recounted on Ira Glass's radio program, 'This American Life', about the Israeli concept of being a freier"[14]

Ironiquement, les chemins empruntés par ce carrefour, encouragés par la perspective séduisante de fonds publics généreux pour la recherche (la triste réalité académique de tous les jours), ont eu des succès divers, même si le grand public n'en est pas toujours au courant. D'un côté, un néo-scolastique avec une présence institutionnelle généreuse, beaucoup moins fertile qu'auto-proclamée, du moins par rapport à ce que le capital investi espérerait. D'autre part, de l'argent réel généré par les applications industrielles et commerciales grâce à des groupes composés de quelques chercheurs qui ont conservé leurs fonctions académiques, tel Aravind Joshi, ou, plus souvent, dans le cas de ceux qui ont quittél'université, purgés et/ou désabusés, dirigeant des unités de Cisco, IBM ou Oracle. Les développements tels que les moteurs de recherche, les *chatbots*, sont dus aux enfants ou petits-enfants intellectuels de Harris.

Non en vain, les procédures de *clustering* ou de *text summarization* ne peuvent être conçues qu'à partir d'un dépôt physique, et non d'une hypothèse ou d'une variable silencieuse. Après tout, il y a toujours eu une maison à habiter à Sefarad. Nul besoin de la rêver.



[14] Cf. Bruce Nevin "Noam and Zellig", dans *Chomskyan (R)evolutions*, édité par Doublas A. Kibbee, John Benjamins, 2010, page 161.