

ARTÍCULO DE INVESTIGACIÓN

Some preliminaries to the study of traces in linguistics

Preliminares al estudio de la huella en lingüística

JAVIER ARIAS NAVARRO

Universidad de Lisboa, Portugal

ABSTRACT The present paper constitutes a brief advance of much longer and more detailed ongoing work 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 Sabbetai Harris; Noam Chomsky; History of Linguistics; Traces; Syntax.

RESUMEN El presente documento constituye un breve avance de una obra en curso mucho más larga y más detallada sobre el concepto de “huella” en la teoría lingüística contemporánea, particularmente en la sintaxis. Se cree, por lo común, que la idea fue acuñada por Noam Chomsky. Sin embargo, ya detectamos su uso, con un valor muy preciso, en los primeros trabajos de Zellig Harris sobre lingüística matemática o, para ser más exactos, sobre estructuras matemáticas del lenguaje. En sus orígenes, en lugar de ser un índice respon-

sable de marcar la ubicación de una unidad antes de su movimiento sintáctico (que siempre toma la forma de *fronting*), la traza o huella era el resultado de un producto matricial entre funciones n -ádicas. Por lo tanto, en Harris la huella es principalmente un concepto anclado en el cálculo matricial o, dicho de otro modo, una noción algebraica. La noción de Chomsky, por su parte, está estrechamente relacionada con el lenguaje de programación LISP. EL presente texto busca proporcionar un análisis preliminar de la complejidad conceptual implícita en el concepto de huella, del cual los lingüistas deben tomar conciencia, porque de lo contrario estarán condenados a enredarse en malentendidos infructuosos para nuestra disciplina durante las próximas décadas.

PALABRAS CLAVE Lingüística; Teoría del Lenguaje; Zellig Sabbettai Harris; Noam Chomsky; Historia de la Lingüística; Huella; Sintaxis.

1. Foreword

The present paper has been conceived as a first and somewhat informal presentation of forthcoming work about the concept of *trace* in linguistics, in which the origin and foundations of the term will be thoroughly discussed. Considering the wide linguistic audience this journal targets, the decision has been made to keep the text as non-technical and free from formalisms as possible.

2. Introduction

Despite conventional wisdom that the idea of *trace* was coined, as far as the theory of language is concerned, by Noam Chomsky, the reality is that its use in modern linguistics can be drawn back to Zellig Harris's first publications on mathematical linguistics or, to be exact, on mathematical structures of language¹. This would remain a mere curiosity for linguistics historians, were it not for the indisputable fact that both trace conceptions entail different mathematical (and philosophical) assumptions. For instance, it is by no means trivial that for Zellig Harris a trace will always be a physical deposit, irrespective of any kind of transformations it might be subjected to, and cannot be stripped of such condition under any circumstances, whereas traces in Chomsky's work have undergone an increasing process of abstraction, to the point of being devoid of some fundamental properties, such as a phonetic form. A comparative analysis regarding their relative power and adequacy is thus due. Such is the main goal of the present paper.

1. For the sake of brevity, I will not elaborate here on this crucial and yet often overlooked distinction.

This article focuses on the basic properties of traces in Zellig Harris's work, as compared to those present in trace theory within generative grammar, going back to the EST and GB periods (for its prehistory in the form of *transformation marker*, see 3.4), and still essentially unaltered, despite claims to the contrary, in the minimalist program.

Relevant examples are drawn from different (sub-) fields of mathematics, aiming at the clarification of the rationale behind the concept of trace and its usages.

3. A brief history of the trace in linguistics

This text seeks to provide a preliminary and yet well-thought and broad analysis of the conceptual complexity associated with the notion of trace in the theory of language and, more specifically, in syntax.

Does such an inquiry require justification? "Why bother?" some might ask, in a more or less cynical or even triumphant fashion. Other than claiming in an equally smirky manner our own right to get forever lost in the huge bulk of publications in the field of linguistics (or, for that matter, in any other), it might be worth bringing up the reflections by Geoffrey Pullum upon his revisiting *Syntactic Structures* (SS):

Why care about a retrospective evaluation of a monograph over 50 years old? Because myths about scientific breakthroughs and results can warp perceptions of the history of a field. Creation myths attributing everything to one individual are known in other fields too.

The truth about science is that discoveries and innovations develop over time and build on earlier developments in the field or in adjacent fields, and myths of monogenesis and individual glorification damage contemporary theorizing in at least two ways. First, they encourage scientists in the complacent maintenance of false assumptions: if almost every linguist is convinced that SS showed transformations to be necessary back in 1957, non-transformational research will be underdeveloped or ignored (and indeed I think in general it has been over the past fifty years). Second, they promote biased and lazy citation practices — the same old references passed from paper to paper without anyone checking the sources. Both consequences are worth guarding against (Geoffrey Pullum, 2010, pp. 251-252).

Guarding against a dramatic loss of naturality may have also become a desirable goal in itself. If I recall correctly, Karl Lachmann once blushed before his audience of students in the middle of a lecture because, in the midst of a momentary lapse, he could not recall the Latin word for coal. Albeit having words on the tip of your tongue (or on the brink of your mind, so to speak) is also a well-attested phenomenon in the native language of monolingual speakers, such ordeal arose from languages conflicting in his head: no matter how reputed a Latinist one might be, nothing can replace the natural lightness of having been brought up in an environment stemming

from *carbon/carbonis* instead of one pervaded by *Kohle*. The fate of all sophistication echoes in Nietzsche's words in *Menschliches, allzumenschliches: Wie? Ein großer Mann? Ich sehe immer nur den Schauspieler seines eignen Ideals*.

It should be noted on passing how symptomatic of the bad taste linguists show when giving names to their animals (objects and operations)² happens to be the fact that modularity hides the mathematical problem (not a minor one) of naturality, tightly linked to Hilbert's Theorem from 1890 regarding syzygies in algebra (Krieger, 2003, p. 22, footnote 2). Regrettably, the term "naturality" has been recklessly eroded after decades of speculation, including phonological theories, which, regardless of how empirically fruitful they may have panned out to be, are a far cry away from even the slightest echo of such concept.

The inability of many colleagues to understand this has certainly caused episodic stagnation in the field (or even regression, someone might argue). Linguists should become aware of what is at stake at every significant intellectual crossroad, for otherwise they will be doomed to be entangled in misunderstandings unfruitful to our discipline for decades to come³.

3.1 Harris, Chomsky and LISP

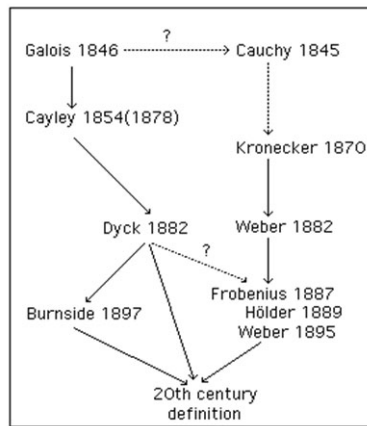
If one has a look at the definition of LISP in McCarthy's *History of Lisp* (McCarthy, 1981), one cannot help but notice how much closer it is to Chomsky than to Turing. In McCarthy's recollections from 1978, he explicitly states that one of his goals was to make a usable version of recursive definitions. It is worth remarking that LISP shares the same theoretical underpinnings as regular expressions and context-free grammars. That being said, there is no lack of connection (albeit a more indirect one) between Harris and the famous programming language. Thus, Oehrle (Oehrle,

2. Indeed a tragic irony when considering the background their knowledge relies on, which ought to make them extremely sensitive, Saussurean arbitrariness notwithstanding, to the "adequate name", to what Chinese people call "*zhengming*", 正名.

3. Truth to be told, I grow more and more convinced of that fact that linguistics attests a delay or *décalage* of 70 years, that has been occasionally reduced, in certain blossoming periods, to, at best, some 30 or 40, with regard to true leading research and epistemology. Perhaps one single example will suffice to illustrate my point. I beg the reader not to consider it an ad hominem argument (for I have no personal connection whatsoever with the author, whose texts I have sometimes found enriching), but rather see it all at the light of what Hegel called *Geistfigur*. In a lecture entitled "Traces Exist (Hypothetically)" given by Carl Pollard at Stanford University in 2013, the following words were pronounced: "I wish we had known about natural deduction 30 years ago!!". This is indeed startling. As a matter of fact, Gentzen published in the 1930s!! And, for those scholars who might eventually live entrenched within the limits of their own idiomatic tradition, an English translation of his *Collected Papers* is available since 1969.

2010, pp.46-47) has convincingly shown how the linguistic properties of intercalation which Harris considered in his early work on Semitic morphology may be formulated resorting to LISP factors *mapcar* and *mapshuffle*. Now, in order to fully account for such non concatenative phenomena, usage of more sophisticated tools, such as the shuffle product, as developed by Eilenberg and MacLane, is required.

It is, however, in algebra where the main keys for the development of the harrisian approach may be found: “Born in 1909. Harris was the exact contemporary of many great algebraists, his peers, such as G. Birkhoff and S. MacLane in the United States or A.I. Mal’tsev in the Soviet Union: it was in this sphere of influence that he found ‘his’ algebra” (Lentin, 2002, p. 4).



Some of the most prominent figures from the 20th century in the field trace back their genealogy to the above tree diagram. Thus, Emmy Noether was strongly influenced by Heinrich Weber, while Issai Schur, for example, directly relates to Frobenius. I leave it to the reader to explore the subtleties involved in the transfer of this powerful set of analytical tools to the realm of linguistics, an exercise in inquiry that should bring him or her joy and amusement. It should suffice to say that Zellig Harris inspiration from Noether should come as no surprise to future researchers (the reasons for that being too deep and sometimes subtle to be exhaustively listed here). As to Chomsky, it might be worth referring how the context-free grammars from his hierarchy of formal languages and the Backus-Naur form underlying rewriting rules can be smoothly handled by the Perron-Frobenius Theorem with its eigenvalue and nonnegative matrices, which at the same time accounts for the entropy, growth-sensitivity and ergodicity of such languages (Ceccherini-Silberstein and Woess, 2003). Furthermore, the problem of the Frobenius number has been recently construed in terms of symbolic regression and grammatical evolution (Adžaga, 2017, with reference to the concept of *codons* as discussed in the Epilogue of the present text). It has also been

shown (Andrews, 2003) how the issue of learning a symbolic dynamical system with an invariant probability distribution over its state-space (what is labelled as *the inverse Frobenius-Perron problem*) is to be reduced to the problem of learning a physical dynamical system with an invariant probability distribution⁴.

3.2 Traces in Zellig Harris vs. traces in Noam Chomsky

As Bruce Nevin (Nevin, 2010, p. 115) has pointed out in a work of the utmost relevance for anyone genuinely interested in intellectual history, divergence between Noam Chomsky and his early master Zellig Harris can be found as soon as 1951. A comparison of their respective views on Rudolf Carnap's contributions clearly unfolds a deep discrepancy as to the role of science, as well as regarding the concept of "transformation". Thus, Chomsky's stance on the latter can be construed along the lines of Rudolf Carnap's rules of transformation, as opposed to an algebraic reading of the term, which is what Harris always had in mind when employing it. Nevin's conclusions on the matter are sobering:

"There is no evidence that Noam ever understood Zellig's origination of the notion of grammatical transformation in algebra. Zellig's transformations are a property of language, Noam's are a formal device for representing that property by 'enriching' the rules of a phrase-structure grammar. Rules of grammar may be widely variant in form, as a matter of notation and system, but transformations in the algebraic sense are variable only insofar as language varies, and changes, and possibly evolves (or is modified) to develop new capacities. Zellig developed a description of language as a mathematical object, and of linguistic information as its interpretation; Noam developed a formal system, the procedural steps of which produce (many, by intention all) sentences of a language, and advanced the hypothesis (couched as a necessary presupposition) that this system describes or corresponds to the cognitive means by which speakers of the language produce those sentences" (Bruce Nevin, 2010, p. 118-119, note 32).

There are, however, other areas in which the discrepancy is as deep and far-reaching. In this paper we will concentrate on one of those areas, namely on the concept of *trace*, which has been often overlooked in historical accounts, despite the significant bulk of work devoted to discussing its formal machinery. The present account

4. The issue then shifts to how to modify the equations governing a dynamical system in order to produce desired probability distributions over its state-space.

will focus, primarily, on the mathematical rationale behind the idea of trace and its usage in syntax⁵.

The beginning of the story goes back to Leopold Kronecker, who in 1870 proved the fundamental theorem of finitely generated abelian groups for the case of finite abelian groups⁶. In his definition, still lacking overt statements in the group-theoretic terms of our days, he proceeds as follows: he takes a specifically finite set $\phi', \phi'', \phi''', \dots$, such that from any two, a third can be derived by a specific method of abstract composition between the components, a method already determined in advance. If the result of such procedure is signalled by f , for the case where $\phi' \phi''$ are equal, there must be a third ϕ''' which amounts to $f(\phi', \phi'')$. Besides, it holds that

$$f(\phi', \phi''') = f(\phi'', \phi'),$$

$$f(\phi', f(\phi'', \phi''')) = f((\phi', \phi''), \phi''')$$

Accordingly, if ϕ' and ϕ'' are different, $f(\phi', \phi'') \neq (\phi' \phi'')$ (Wussing, 1969).

In his *Lehrbuch der Algebra* from 1895-96, Heinrich Weber defined a group of degree h to be a finite set. He required that from two elements of the system one can derive a third element of the system so that the following hold:

$$(\theta_r \theta_s) \theta_t = \theta_r (\theta_s \theta_t) = \theta_r \theta_s \theta_t$$

$$\theta \theta_r = \theta \theta_s \text{ implies } \theta_r = \theta_s.$$

In a modern approach one might write instead: $(\theta_r \theta_s) \theta_t = \theta_r (\theta_s \theta_t)$, so that either side may be denoted by $\theta_r \theta_s \theta_t$. What Heinrich Weber was defining is, in fact, a semigroup with cancellation and, given its finiteness, that ensured the existence of an identity and of inverses. Weber himself reckoned that this failed for infinite systems. From a historical standpoint, the concept of abstract group was the first case of emancipation by an algebraic structure⁷. In recent years, the idea of modeling semantic compositionality by means of Kronecker tensor product has been put forward (van de Cruys, Poibeau and Korhonen, 2013), leading, for instance, to a so-called Frobenius anatomy

5. The concept of *trace* of a matrix, as equivalent of the German term *Spur*, can be first found in the 1922 English translation of Hermann Weyl's book *Raum, Zeit, Materie* due to H. L. Brose. As the reader will probably know, the trace amounts to the numerical value resulting from adding all the components of the matrix main diagonal. It should be emphasized that the main feature of the trace in mathematics is that of invariance. Thus, even if an orthogonal matrix transformation is carried out, in which all the numbers are changed, the sum of the diagonal remains the same. This means that the trace keeps track of the eigenvalues, that is, of the elements which come up in the diagonal when the rest of the matrix is made up of zeroes. The trace is not the only invariant, though. Truth be told, the underlying characteristic polynomial of a matrix is also invariant.

6. Kronecker's proof was generalized to all *finitely generated* abelian groups by Emmy Noether in 1926.

7. I am relying here strongly on the content of a lecture on the topic given by Peter Neumann at the University of Sussex in 2001, celebrating the 90th birthday of Walter Ledermann, one of the main contributors to the realm of group theory.

of relative pronouns (Clark, Coecke, and Sadrzadeh, 2013). The reader eager for an intuitive representation of a group may just think of the famous Rubik's cube⁸.

What Zellig Harris notation stands for is nothing short of a tensor. He is basically abiding by the same operations Kronecker deployed in 1870 when defining the finite Abelian group, taking as basis a law of composition between its elements. Now, with regard to notation symbols, Harris uses the same as Cayley did. The only difference comes from the fact that both Cayley and Kronecker (who employed the Greek theta letter to indicate the generic elements of a group, instead of phi, as the former) was using prime and superprime indexes where Harris will be using numbers as subindexes and superindices. That very same notation extends to tensors once one considers that what is being combined are vector spaces. What Harris does is tensor contraction (or sort of).

This is reminiscent of the differences between the abstract index notation and the Ricci calculus with regard to contraction. The former indicates that a basis-independent trace operation has been applied, which reduces to a true summation whenever a specific basis is fixed. In the latter, contraction signals a literal summation, which requires numbers, thereby also demanding the choice of a specific coordinate system. In reality, the abstract index notation is faithful to the fact that almost all of the Ricci calculus remains intact if one does not choose a basis. In other words, there is a lot of meaning in the structure of the index expressions which does not depend on the basis.

A simple example to illustrate the point is an expression of the form t^a_a , where t^a_a is a tensor (or tensor field) of type (1,1), vulgarly known as an endomorphism. The expression t^a_a is interpreted in the usual index notation as the sum of the diagonal elements and, taken literally, is dependent on the basis. When interpreted as contraction in the abstract index notation, it results in a scalar (or scalar function) without ever having to choose a basis⁹.

3.3 Principles regulating traces

In the Principles and Parameters framework, traces abide by the so-called empty category principle (ECP), which states that all traces should be properly governed, that is, they should be either theta-governed or antecedent-governed. Consider the following examples:

- (1) Who did John say that Mary saw t ?
- (2) Who t said that?

8. A similar image, that of a cylinder, which became Rubik's next recreational toy, is explored in García Calvo (García Calvo, 1983) with regard to language types.

9. I owe this example to Mikhail Katz, Professor of Mathematics at Bar-Ilan University.

In (1), the verb “*see*” both governs and theta-marks the trace *t*, which is then theta governed. In (2), the *wh*-element governs the trace and is coindexed with it. The corresponding trace is therefore antecedent-governed.

It should be borne in mind, though, that intermediate traces are not subject to the ECP, since they are deleted at the Logical Form. Furthermore, they are distinguished from other empty syntactic categories, such as *PRO* and *pro* (Sag and Fodor, 1994).

These principles fundamentally diverge from those regulating traces in the Harri- sian approach, as we will have the opportunity to see at later sections of the present work (see § 3.5-3.7). It should suffice to mention now that Zellig Harris sets a dis- tinction between minimal traces and those which are composed from them. This, together with the claim that traces are a physical deposit, constitutes the core of the difference between Harris and Chomsky in that regard.

Three different stages can be found in Chomsky’s work with regard to the status of the trace: 1) Substitution transformations, 2) Trace Theory of Movement and 3) Copy Theory of Movement, the minimalist program representing a return from 2) to 3). This notwithstanding, traces are used *t* all over the place. So it remains a notational constant, despite a shift in its interpretation. The reasons why it has been so tempting to try to do away with the trace are not too difficult to fathom, since they relate to the usual considerations about economy in theoretical frameworks (being, as it is, that most of them substantially differ with regard to what economic principles are about):

Traces, therefore, have a unique theoretical status and are thus conceptually sus- pect: they are the sole grammatical constructs that are introduced in the course of the derivation to LF. By eliminating traces, a reduction in the number of theoretical pri- mitives is achieved, something that is clearly conceptually appealing given the *modus operandi* of the Minimalist Program (Kandybowicz, 2008, p. 4).

Movement in Chomsky can be drawn back to the substitution transformation in Harris (Kandybowicz, 2008). Whichever the type of movement, the trace does not change. Objections were raised within the generative tradition as to the particular status and nature of the trace: thus, criticism was made by Pullum and Postal (Pullum and Postal 1979) regarding *to* contraction in English (but see Radford 1999 for an ac- count of traces considering such contraction). Special attention was paid to iterations of movement, as in twice-displaced NPs (Lighfoot, 1976) or multiple *Wh*-Fronting of the kind found in German examples such as *Wovon glaubst du wovon sie träumt?* (Grewendorf, 2001, Boskovic, 2002). The study of apparent violations of the Proper Binding Condition, in which successive movements yield a configuration in which the antecedent no longer *c*-commands the trace, for the material extracted now linearly precedes the antecedent position, gave rise to the *remnant movement theory* (Koo- pman and Szabolcsi, 2000, Stabler, 1999, Cecchetto, 2004, Grewendorf, 2015). An extreme case is represented by multiple copy spell-out, as exemplified by repetition

phenomena as attested, for instance, in Nupe, a language spoken in Níger-Congo (Kandybowicz, 2008). Even if all of those may be rendered internal disputes within a school, with focus on tinkering many unsatisfactory details, while not questioning any of the principles of the paradigm¹⁰, the reality is that they lead in recent years, to formulations that drift away significantly from its original matrix, as illustrated by Borsley's latest work (Borsley, 2012), where his usage of the slash in an operator-like manner reminds of the dash for the trace in relation with free pregroups, as put forward by Lambek (see Morrill, 2011, for its interpretation within a general framework of displacement logic for grammar).

In linear algebra and functional analysis, a generalization of the trace is developed under the label of *partial trace*. While the trace is a scalar valued function on operators, the partial trace is to be construed as an operator-valued function. The most relevant applications of the latter pertain to the study of quantum information and decoherence (Nielsen and Chuang, 2000). To what extent, if at all, such partial traces relate to some of the subcategorizations of traces Harris advocated for, remains a topic for future research.

3.4 Some chronological perspective

From a philological standpoint, the following landmarks can be established in the history (and prehistory) of the concept.

As far as our knowledge goes, no reference to traces in the field of generative linguistics is made before their emergence in the extended standard program back in 1971 (Chomsky, 1971, included in Anderson and Kiparsky, 1973). However, a significant precursor must be pointed out, namely *Aspects of the Theory of Syntax*. In

10. Critical revision of the cornerstones of trace theory began as early as 1980 (Pullum and Borsley, 1980). With regard to proposals focusing on sideward movement (Corver and Nunes, 2007), under the assumption that Move is not any longer to be taken as a primitive in the minimalist program, but rather as the result of a complex fusion of Merge, Copy, and different linearization commands, under the names of Form Chain and Chain Reduction, my stance can be formulated in the following terms, however harsh they may sound to the naive reader: 1) By ignoring or postponing a serious discussion on Noether ascending and descending chain conditions, as well as on the interplay of chain intersection of compact and connected spaces and Hausdorff spaces, linguists will entertain (or deceive) themselves in an eternal detour or loop; 2) Chomsky had already proposed the computational command Form Chains as a solution to the paradox between two notions of economy he himself presents (Chomsky, 1995, p. 181-182), namely, shortest move versus fewest steps in a derivation. He might as well have cried out, "Join Paths!", as an inverse image of Biblical Moses, for what is being taken for granted is nothing but a fundamental group, in the sense the term has in algebraic topology. The two other options that come to my mind as plausible are either considering that the fundamental substratum of language is akin to Hausdorff space - which hardly comes to grips with the undisputable fact that most of the times the results from operations such as Merge do not comply with the required associativity linked to that space - or to introduce a little intervals operad, a mathematical concept Jon Peter May has worked on extensively.

that seminal work, Noam Chomsky introduced the notion of *transformation marker* (Chomsky, 1964, p. 131-132) and, in connection with it, that of *transformational history*, by which a given well-formed sentence that is derived from a basis made up of a sequence of phrase-markers can be represented diagrammatically¹¹.

A major epistemological requirement, part of today's research standards in disciplines such as physics, is the conservation of readability in jumping from one theory to another (to which, I dare say, one must add the ability to (re-)formulate any in natural language, in the "plain language"). To put it another way, it is a criterion of intersemiotic translatability, as already required by the likes of Charles Sanders Peirce or Roman Jakobson. The concept of *trace* should therefore retain its readability when hopping from the grass of one theory to the green pastures of another. Such a condition is arguably satisfied in the drift from GB to minimalism, despite the fact that in the latter, as it has been already seen above, traces are reduced to some kind of notational relic. However, no attempt has been made, to my knowledge, to show and prove the transversality of the concept in the shift from Harris to Chomsky.

Summarizing, Chomsky introduced the term trace in 1971, well after Harris's usage in 1965 and 1968. This might be well be construed as another case of "adaptation in a different context", as it has been just seen when dealing with the drift and fate of transformations. The effect in both cases, one might argue, was that of obscuring for Chomsky's disciples the true sense and value of anything by Harris that they might happen to read.

11. It should be noted that such chronology is accurate with regard to the subjective perception by the public (that is, by the readers and the research community of the time), but the picture changes if one considers the actual order of events, as they came to be known later. Thus, Chomsky was already making use of T-markers in his *The Logical Structure of Linguistic Theory* from 1955, which was only accessible to a wide public twenty years later, in 1975.

In that framework, to borrow Howard Lasnik's retrospective formulation, "T-marker is the interface with semantic interpretation, and the final derived syntactic representation is the interface with morphophonemics." (Lasnik, 2011, p. 16).

In his Introduction to *LSLT*, Chomsky acknowledged that his "transformations are understood in a very different sense; it probably would have been preferable to select a different terminology instead of adapting Harris's in this rather different context." (Chomsky, 1975, p. 43, also reproduced in Nevin 2009, p. 474).

3.5 Some basic ideas

In his work, Zellig Harris distinguishes two types of transformational traces:

- 1) The trace as the concatenation of the trace with its operand
- 2) The trace as permutation to some other point of the operand¹²

Let us further probe into the basic contrast at stake here, namely that of concatenation versus permutation. The former is to be understood in terms of contiguity of elements¹³, whereas the latter presents its usual mathematical meaning, as given in combinatorial theory. The term *operand* adopts here its standard mathematical meaning. The latter encompasses the idea of movement (or displacement, if a more theory-agnostic term is preferred)¹⁴.

12. Even if the first edition of *Mathematical Structures of Language* dates back to 1968, I will be always citing the second edition from 1979, due to Robert E. Krieger Publishing Company. This is the relevant paragraph:

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, p. 206).

13. See the concluding chapter of the present paper for some crucial quotes from Harris on that issue.

14. As one anonymous reviewer correctly points out, this formulation is somewhat elliptical, and requires some clarification. In his or her view, what appears to be a distinction between types of traces does in fact concern a different location of traces. Indeed, in the passage at stake, Harris states that the locations of transformational traces (understood as the physical deposit in one set of sentences which is absent from the other set under the mapping) are linearly ordered, and that the location of the trace can be viewed first as “the concatenation of the trace with its operand, 7.1.2, and in some cases the permuting of the trace to some other point of the operand”. In the referenced section 7.1.2, Harris goes on to say: “Each [function] *f* is a finite set of operators; each operator introduces material which is concatenated to its operand, or introduces changes in its operand”. The trace is therefore the introduced material or the introduced change, usually a morphophonemic change of word shape (even to zero), but “in some cases” permutation of previously introduced material occurs. An illustrative example may be yielded by the permutation of an adjective (a stative operator) shown in the relationship between *A book is red, a book which is red, and A red book*.

All this reflections are pertinent; however, I prefer to keep the metalinguistic statements neutral with regard to movement (or dislocation), in the event of it all proving to be just a representational mirage due to the dimensions of the substratum space the units are operating on within the theory (as alluded to as a somewhat scary possibility at the end of the present text).

It is important to note, as Harris himself does, that syntactic information is reactivated in the trace, so, strictly speaking, we are not dealing with an empty category here, but rather with doubled features (or features that need to be double-checked in order for the computation to proceed smoothly). Other than for some constraints regarding the features the trace may or may not bear, we are not so far here from the conception underlying the discontinuous morphemes within the structural tradition¹⁵.

Let us now briefly focus on tensor construction via tensor products. The first thing which has to be pointed out is that the rank, dimension and the number of degrees of freedom linked to its parametrization do not suffice to classify it in a fully-fledged manner. A given p-q tensor is “p” times covariant and “q” times contravariant, depending on the number of components from each type it presents. The rank of the tensor results from the sum of the quantities expressed in both components. A tensor is said to be antisymmetric in some of its components if, when interchanging two of them, it undergoes a change in sign, then gets the sign back upon the next interchange, and so on and so forth. The Levi-Civita symbol or the electromagnetic field tensor from Maxwell constitute two salient examples of that.

A *lower index* or subscript indicates covariance of the components with regard to that index: $A\alpha\beta\gamma\dots$

An *upper index* or superscript stands for contravariance of the components with respect to that very index: $A^{\alpha\beta\gamma\dots}$

Finally, a tensor may as well have mixed components, that is, both upper and lower indices: $A\alpha^{\beta}\gamma^{\delta}\dots$

The ordering of indices is highly significant, even when of differing variance. Two indices (one upper and one lower) with the same symbol within a term are summed over. That is known in multilinear algebra as a tensor contraction¹⁶. Such operation is incorporated into the renowned Einstein notation. As a result, one gets another tensor with order reduced by 2. Tensor contraction can be interpreted as a generalization of the trace. As it has been seen before, in the abstract index notation the indices are mere placeholders, labels of slots by means of letters. They are thus non-numerical and, therefore, not related to any fixed basis. That sets that notation apart from the Ricci calculus. Let V be a vector space, and V^* its dual. Consider, for example, a covariant tensor $h \in V^* \otimes V^*$, with rank two. It can be construed as a bilinear form on V , or, in other words, as a function of two arguments in V which can be represented as a pair of *slots*:

15. Somewhat diverging formulations may be found in the works of Zellig Harris, André Martinet or Charles F. Hockett.

16. Additionally, more than one index may each occur exactly twice within one term.

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

A contraction between two tensors is represented by the repetition of an index label. Thus, for example, $t_{ab}{}^b$ is the trace of a tensor $t = t_{ab}{}^c$ over its last two slots. This way of representing tensor contractions is formally similar to the Einstein summation. However, given that the indices are non-numerical, no real summation is at stake here. What we have is, rather, the abstract basis-independent trace operation (or duality pairing) between tensor factors of type V and those of type V^* . Generally speaking anytime one contravariant and one covariant factor occur in a tensor product of spaces, there is an associated *contraction* (or *trace*) map. For instance,

$$\text{Tr}_{12}: V \otimes V^* \otimes V^* \otimes V \otimes V^* \rightarrow V^* \otimes V \otimes V^*$$

is the trace on the first two spaces of the tensor product, whereas

$$\text{Tr}_{15}: V \otimes V^* \otimes V^* \otimes V \otimes V^* \rightarrow V^* \otimes V^* \otimes V$$

stands for the trace on the first and last space. These trace operations are signalled on tensors through the repetition of an index. Thus, we have:

$$\text{Tr}_{12}: h^abc^de \rightarrow h^ac^de \quad \text{and} \quad \text{Tr}_{15}: h^abc^de \rightarrow h^abc^da$$

Unfortunately, dealing with some related issues, such as the braiding associated to both the Riemann tensor and the Bianchi, identity goes far beyond the scope of the present paper. It will suffice to briefly refer on passing to the fact that braided monoidal categories are at the basis of semantics as proposed by André Joyal and Ross Street (Joyal and Street, 1986, 1991).

Summarizing, what is at stake here is a trace class operator. The multiple indices of a tensor are split into covariants and covariants and are usually written in two lines, as superindices and subindices. A trace always consists of a covariant index and a contravariant one. As noted above, Zellig Harris does not make use of either superindexes or subindexes. He just provides two numbers in the subindex slot, pretty much the way in which matrix rows and columns are annotated. Those two numbers are also meant to indicate which two objects from the list the difference signalled by the trace is relating to. Let us illustrate this with one particular example: given a tensor T with contravariant indices a, b, c, d and covariant indices z, x, v, n , then Tr_{23} would be the trace over the indices b and v and Tr_{32} would be the trace over the indices c and x . Consequently, “the importance of the trace is that it is a physical deposit in one member of the transformationally related pair of propositions” (Zellig Harris, 1979, p. 61).

Thus, a base operator like ϕ_{21} stands for elementary differences among sentences, in this case, for differences between transform 1 and transform 2, whichever they may be.

To conclude this section, it might be worth briefly mentioning some specific examples illustrating how Harris applied set-theoretic and algebraic concepts in his

analysis of language. Thus, one already finds a very thoughtful use of partially ordered sets (posets), which later became a landmark of formal semantics and Montague Grammar, in *Mathematical Structures of Language*. Harris's usage is aimed - as it will become customary after him - at dealing with sets of sentences and their decomposition into primes (kernels and carriers, in his terminology; constituency nodes or constituents themselves in other approaches). Such decompositions are partially ordered, as it can be empirically tested with some ease: Let us take the English utterance *A young boy's beginning to walk is slow*. Sentence A. *A boy walks* necessarily precedes Sentence B. *A boy is young*, in the factorization, since the reverse ordering would yield the undesired outcome *The boy who walks is young*. Non commutative products are at the basis of his treatment of pairs like *I began to believe that he left* as opposed to *I believe that he began to leave*.

As it has already been shown in the present paper, Harris paid special attention to group theory. Thus, groups and semigroups belong to his main tools of analysis. The absence of a proper inverse element (which groups demand) in many linguistic phenomena makes semigroups (and monoids) particularly suitable for the study of language. Thus, the reductions of certain elements to others lies on an underlying semigroup structure: the explanatory replacement of *wh* (*A book which I bought has disappeared*) and other conjunctions by *and* with a suitable metasentence of the form CS, a metatoken formulation which shows that referential sameness for individuals is not expressed by the primitive terms acquiring a new power of reference in each occurrence, but by the deployment of discourse linearity, namely, [1, *A book has disappeared*] and [2, *I bought a book*] and [3, *In 1 and 2*, 'book' in post-verb position of 2 refers to the same individual as 'book' in pre-verb position of 1].

Besides, the concept of "groupoid", forerunner of today's magma in many respects, is employed in order to account for the set of sentences under the (nonassociative) binary operators, as exemplified by conjunctions or subordinating connectives.

With regard to monoids, which Harris uses for free products and transformations, and which will become, as free monoids, the cornerstone of Joaquim Lambek's calculus years later, I would like to note on passing that their graphical representation in mathematics corresponds to what Hegel defined as *Selbigkeit*, while exactly matching some of the developments in topology and cohomology theory regarding *étalé spaces* and *topoi* as conceived by Grothendieck, Deligne, and others.

Harris is, nonetheless, never captive of what we might call the hypnotic seduction of formalism. Rather, he keeps a sober and realistic attitude towards it, being fully aware that what separates the usage of some simple mathematical objects for linguistic analysis from utter triviality is the very meaning-bearing nature of language itself:

“...All of the structures as they now stand are of very limited mathematical interest. They are insufficiently regular, and in some cases have disturbing constraints. The mathematical interest may lie in specifying what are the essential points that make these structures depart from their nearest neighbors within mathematics, and how these essential disturbances are related to the semantic burden that natural language alone can carry” (Zellig Harris, 1979, p. 207).

A distinct trait of Harris’s approach is his deployment of graphs. That path will ultimately result, after decades of extensive work on that domain by an exponentially growing number of researchers, in the Caucal Hierarchy on graphs, analogous to (but not fully commensurate with) its Chomsky / Schützenberger counterpart for strings. This is of the utmost importance, as I show in some of my forthcoming work. Harris’s emphasis on factorization (a prerequisite, indeed, for the well-functioning of the whole language architecture at various levels) does imply its being also worked out for graphs. As a matter of fact, graph factorization is, in my view, not only one of the main areas of differentiation with Chomsky (who did not contemplate it; see the Epilogue on Section 9), but also, one of the richest fields of research on language at the present time, to which, unfortunately, most linguists are oblivious. Matilde Marcolli’s studies on *Graph grammars* at Caltech, which are to be reckoned as some of the most important work currently conducted, might be seen as somewhat indebted to Harris’s pioneering work on the analysis of language. Decomposition lattices and their straightforward by-product, ideals also find a relevant place in Harris’s analysis of language utterances. The connection to factorization into ideals in the sense of Kummer and Dedekind is explored in recent work by Javier Arias (Arias, 2015).

On a curious note, certain hierarchies for language modelling, frequently known as Stanford Hierarchies, might have to be relabeled, for the sake of poetic justice, UPenn Hierarchies, since they constitute a significant bulk of Harris’s syntactic analysis from the late 1960s.

Familiarity with Category Theory is also at hand, as in the usage of commutative diagrams (e.g. Harris, 1979, p. 156).

After this extensive analytic and propaedeutic *tour de force* by Harris, the door was somehow already left ajar for the upcoming birth of opetopes (somehow linked to the French branch of his disciples, like Maurice Gross and others, who happened to have fluid communication with Groethendieck and the Bourbaki group)¹⁷.

17. Interestingly enough (cultural history is full of such felicitous coincidences, with also its due share of calamities, most of which remain untapped and still await for exploration), a twist of fate due to life among Spanish exiles in Paris, which included some prominent banned linguists and mathematicians, lead to the fact that opetopes were used (albeit not mentioning the concept) in syntactic analysis of Spanish by Agustín García Calvo as early as 1983, while *Pursuing Stacks* by Groethendieck was still unpublished and only circulated within a very intimate circle of contacts.

Incidentally, one should never underestimate the fact that the early dawn of Natural Language Processing (NLP), as we have come to know it today, is due to the work by Aravind Joshi at the Computer Science Department at the University of Pennsylvania. Many of his closest collaborators (some of them with strong or almost exclusive background in linguistics) left, or were forced to leave, academics and went to direct the then emerging divisions at the likes of Cisco, Oracle, or IBM. Thus, some of the most remarkable and flashy applications for our everyday life, like search motors, sound interfaces or chatbots, are due to that core of people, and not so much to the die-hard syntacticians or semanticists involved in the so-called linguistic wars.

Regrettably, more detailed elaboration on these issues would require an entirely separate paper.

3.6 Some additional differences

Furthermore, Harris takes into account those cases, by no means scarce, in which a trace is made up of others, as a product of what he calls *minimal traces*. In Chomsky, though, differences between traces are limited to those differences mirrored in indexation. Ironically, it is the structuralist Harris, not Chomsky, who envisions the generative power of traces, while also bringing together universalism and typology: “Minimal traces are of very few types within a language, their types being quite consistent across languages” (Zellig Harris, 1979, p. 65).

Certain constants appear in many traces. If a constant appears in two or more traces, as in ϕ_{31} and ϕ_{51} , then there is some trace which consists of that constant, and both ϕ_{31} and ϕ_{51} are products that contain the constant as a component.

Some traces can clearly be obtained by the successive application of two or more other traces:

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 (Zellig Harris, 1979, p. 65).

The question of movement is tightly intertwined with that of the trace. We have already seen that in Chomsky the trace is the result of syntactic movement (of operators, quantifiers, and so on). Such a movement is always one of *fronting*. Now, the question can be legitimately posed as to why movement knows only one direction. This is a fundamental question. Replying, as it is customary, shifting the weight of the proof on empirical considerations -psychological computability, for instance- demonstrates, in our view, the typical naivety of epistemological reductionism, being as it is easily shown, that this is a consequence of mathematical representation system adopted and preferred for the case. It is a direct consequence of using Lukasiewicz’s

postfix (also known as reverse) Polish notation. It is indeed a computation problem, but it is not the subject of the theory of language that is concerned, but rather the internal structure of calculating machines. Fronting is the only movement that a postfix notation can allow if one wants to optimize the stack. For any tree of operations, there is always at least one postfix sequence that can be evaluated with the most simple use of memory, namely, with just a stack in which only the two topmost elements are read. There are indeed good practical reasons to use the reverse Polish notation instead of the infix or prefix ones. In the other notations, one must allow access to lower elements, which have been previously deposited. That is the reason why, formerly, an infix notation calculator had a “limit on open parentheses”, whereas a postfix one did not, the matter then being just whether the stack was deep enough.

There is, however, in my view, a deeper epistemological issue at play in the whole discussion on traces in linguistics, whether the actors involved are aware of it or not. An underlying analogy, or, if the reader so prefers, a syzygy will help me illustrate what I mean, namely what Michael Dummett (Dummett, 1995, p. 135) called, when dealing with Gottlob Frege’s strategy in *Grundlagen der Arithmetik*, a two-sorted theory, where an expanded language, involving reference to quantification over directions, can be translated into the original language, that involves only quantification over lines. Two choices result from the expansion adding the direction-operator: 1) identify directions with lines, or 2) differentiate them both. The former yields a one-sorted theory, in which the direction-operator is delineated explicitly¹⁸, without resorting to a contextual definition; the latter leads to the already mentioned two-sorted theory. Constructing a model of the new theory, given a model of the original, is made possible by the ontological parsimony of the theory of directions, which does not require more objects of the new kind - directions - than were already present of the old one - lines. Meanwhile, the theory of cardinal numbers is very far from being ontologically parsimonious: it demands the existence of $n + 1$ numbers, given n objects of the original kind. Even if it were no possible to eliminate the direction-operator, we could, by re-interpreting the quantifiers, translate statements involving directions into statements not involving them. This cannot be done for statements involving numbers. Summarizing, the crucial fact is that the cardinality operator is of second order, while the direction-operator is of first order. Given that, occurrences of the cardinality operator can be embedded within the scope of other occurrences in a much more complicated way than with the first-level direction-operator. That leads to what has been known as Hume’s principle (Boolos, 1998), one of Frege’s

18. Otherwise, it would be entirely contrary to Frege’s ideas to place any restrictions on the occurrence of terms for directions in the argument-places of predicates.

basic tenets. Accordingly, the introduction of the cardinality operator entails a much stronger ontological assumption: the domain of objects over which individual variables range must be infinite¹⁹. In *Grundgesetze der Arithmetik*, Frege allows for binary functions, thereby opening the door for arguments that skip levels, that is, for correlating units from different levels. They look startlingly similar to tensor contraction.

If we were to very succinctly present a mathematician who happened to be a layman in linguistics with the way in which movement and traces work in contemporary generative grammar, and then ask him or her about the mathematical rationale which, in his or her opinion, lies behind the trace of a move operation, he or she would probably search for inspiration among the mute or dummy variables of calculus. Crucially, dummy variables are local, so that they can be used several times for different goals without incurring in contradiction or name clashing²⁰. For example, in Fourier series n is customarily used to name two different dummy variables.

All along section 3 we pointed out the main abstract properties of a trace in mathematics and the claim was made that invariance (e.g. invariance with respect to the change of base) is perhaps the most fundamental of them all. For example, the trace comes in extremely handy when trying to classify Möbius transformations. Now, is that property of invariance pervasive through the notion of trace in Harris and Chomsky?

The quotation below can be found (and that is far from a random occurrence) in *Mathematical Structures of Language* under the epigraph “homomorphisms and subsets”, in which the claim is made that the set of sentences (or utterances) partakes in certain homomorphisms which preserve the transformational relation.

19. One could introduce stipulations to the effect of giving the one-sorted theory the force of a two-sorted one, or, as Dummett puts it, “to sterilise reiterations of the direction-operator” (Dummett, 1995, p. 136). A theory could then be obtained whose theorems might be translated into theorems of the original theory with no direction-operator. Yet that would not imply real elimination of the direction-operator, for, when conducting the mapping from theorems of one theory into theorems of the other, we could not leave the quantifiers intact, but would be compelled to translate them. I leave it to more competent logicians than myself to dilucidate the particular details of the connections to basic assumptions of current syntactic theory, but it seems clear to me that Move necessarily assumes the asymmetry of the operator space and requires a direction operator to come to the fore.

20. Mathematicians handle scoping poorly. Mathematical notation is often extremely unclear about the scopes of objects or even indicating that some construct is a binding form at all. That is, one might argue, one of the major factors that make calculus difficult for so many beginners; after all, in it scoping is used non-trivially. Some of these issues are tackled in Sussman and Wisdom’s book (Sussman and Wisdom, 2001).

There are interesting relations among the mappings which have been mentioned in preceding chapters. Thus, we take S_ψ as the set of Ψ -decomposable propositions, on which are defined the lattice-like Ψ -structures representing the Ψ -traces in each proposition. E_ψ is the set of elementary propositions in S_ψ . 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 Ψ -structures, and $S_\psi \rightarrow S_\psi E_\psi$ is the natural mapping mentioned ... (Zellig Harris, 1979, p. 92).

An essential and very distinctive aspect of the harrisian approach to traces regards its role pertaining language change. Traces themselves may be subject to substantial shifts over time. That is indeed an idea chomskyan formalism does not capture nor seems suited for:

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 (Zellig Harris, 1979, p. 92).

The claim can be made that traces are (at least to a great extent) an empirical question. Thus, as Harris puts it, "it is possible that identical traces may be produced by different successions of ϕ , especially if some of the ϕ are ϕ , which may be zero parts of the trace that is due to preceding ϕ of the succession" (Zellig Harris, 1979, p. 117). In such cases, an inspection of the ϕ list and ϕ product table for the specific language needs to be conducted.

4. Some fundamental ideas

According to Harris, syntactic information is a crucial factor in traces, so, strictly speaking, a trace is not an empty category. As a matter of fact, he distinguishes (at least) the following two types of traces:

- D – Trace,
- V – Trace,

which stand for determinant traces and verbal traces, respectively. A central tenet of the Minimalist Program is that both displacement and plain structure-building are set up by one single operation, that of Merge (Chomsky 1995, p. 226). A challenge to this view comes from the empirical observation that there are many well-known differences between movement types. Coppe van Urk has recently referred in his lectures at Queen Mary University of London to three major problems in the minimalist view of movement: (1) the distinction between A- and A'-movement, (2) the problem of intermediate successive-cyclic movement, and (3) differences between phrasal movement and head movement.

When confronted with issues of linearization and with multiple phases of syntactic movement, as well as with the (possibly) different categorical triggers these may have, it might be helpful to keep in mind the chess move *en passant*. *En passant* is the sole privilege of pawns, and it happens to be the only capture in chess in which the capturing piece does not occupy the original square of the captured pawn. It is also interesting to underline that in the algebraic notation, the capturing move is written as if the captured pawn advanced only one square²¹.

5. Discussion

One might say, stretching metaphors a little bit, that around the Harris family table, and thanks, to a great extent, to Bruria Kaufman's influence and intellectual stature, an almost unique reconciliation of two deep scientific analogies was taking shape and anchoring in linguistics through his husband Zellig²², in an effort reminiscent of André Weil's Rosetta Stone.

Interestingly enough, in physics, when dealing with the QED Feynman rules, one needs to keep track of the '-' signs arising from re-ordering of fermionic fields and creation / annihilation operators. And, while an individual Feynman diagram is not always gauge-independent, when one sums overall diagrams contributing to some scattering process at some order, the sum is always gauge invariant, which runs analogously to what we saw for the basis-independent abstract notation for tensors.

21. Truth to be told, it should be called Zehfuss matrix, after Johann Georg Zehfuss, who first described the operation in 1858.

22. The two referred major research programs, yoked in an analogy of analogies or syzygy, can be identified as the Dedekind-Langlands Program and the Onsager Program, respectively (Krieger, 2003, p. 14-15). This is, of course, no coincidence, since Bruria Kaufman was the author of a seminal paper, together with Lars Onsager, on the solution of the two-dimensional Ising model (Kaufman and Onsager, 1949). Perhaps the day will come in which linguists discover how much they are indebted to Spinor Theory and Toeplitz matrices ever since.

While Harris linkage to Mac Lane and Eilenberg's Category Theory has been already alluded to in this text (see section 3.1), Chomsky's requires some completion: the study of the connections between Eilenberg machines, Chomsky and cellular automata has experienced renewed interest in the last decade (Razet, 2008, McIntosh, 2009). Considering some of the claims made in 3.2, the fact that the scope of Carnap's influence on MacLane has been attested in some detail (Awodey, 1996) should shed additional light on this intertwined mesh of intellectual flow.

It should be clear by now that the two notions of trace are not commensurate. Chomsky's trace concerns only movement, being a placeholder at a node from which a subtree has been moved. Now, the limited array of phenomena for which Harris invokes movement does not intersect at all with any of those which concern Chomsky.

Chomsky's notational trace accounts for restrictions on what can be relativized and other "islands", thus enabling semantic computation of dependencies across discontinuities introduced by relocation of subtrees (given that, in his view, actual words are associated only by means of linkages through abstract tree structures). Harris's sentence-forms, on their turn, are not abstractions, but generalizations which conveniently serve as abbreviations for lists of homomorphous sentences.

6. Conclusion(s)

In Chapter 2 of *Mathematical Structures in Language*, Harris summarizes the central features of language structure a linguist should always take into account:

There is nothing in language like bars in music. The only elementary relation between two words in a word sequence is thus that of next neighbourhood.

Since no space or distance is defined between operator and operand, operations and their effects are to be rendered contiguous. Any separation will be yielded by just the effect of later operators intervening on a given resultant. A constructive grammar of language, if feasible, must be available based solely on some characterization of its sentences which is based on purely contiguous relations. The only property that qualifies a sequence a format of the grammar is that the objects are not arbitrary words, but words of particular classes (or particular classes of words). The sentence characterization requires defining well-formed subsequences or operators which will make up the word sequences that constitute sentences; it is to these subsequences or operators that contiguity applies.

Two questions might be raised in this regard: 1) What kind of mathematical space (Hilbert space, Banach space, whatever) characterizes the space for writing or talking as Harris describes it? It should be borne in mind that he describes it as a non-measured space. Distance between two words can only be established by means of the sequence of other words between them.

2) How could the contiguity operand-operator Harris refers to be characterized? It is crucial to note that such relation ought to preclude displacement, that is, movement of the elements. Or, in other words, the fact that operations are contiguous entails that movement will always affect one of two given contiguous elements.

With regard to question 1), the observation should be made that trace-class operators are compact nuclear operators on Banach spaces. Sometimes, the term refers to just a subset of those, namely to the operators on separable Hilbert spaces²³. Whichever the case, the trace is always independent from the choice of basis. Crucial contributions in this mathematical domain were made in the mid and late 1950s by Alexander Grothendieck and Victor Borisovich Lidskil (Grothendieck, 1955, Lidskil 1958), hinging on earlier developments by Erik Ivar Fredholm (Fredholm, 1900, 1903)²⁴. As to question 2), one might at first sight be tempted to prioritize graph theory, with the observation that a sentence is a path graph of words colored with word classes. Operands and operators would then refer to the way in which different words or classes act on each other, without a system of brackets separating phrases. Lattices and graphs are indeed explored in some detail in Harris's work²⁵. There are, however, other layers in the theoretical background that must be acknowledged and even prioritized, such as Lukasiewicz's Polish notation, or, more significantly, the algebraic spaces his development of operator grammar demands.

The notion of trace is tightly intertwined in Harris with the system of reductions that yields the other sentences of the language as paraphrases of base sentences. There are three operations within that system, namely, permutation of words (movement), zeroing, and morphophonemic changes of phonological shape. Each reduction leaves a trace, which entails that the underlying redundancies of the base can be recovered. As to linearization of the operator-argument dependencies, the following statement shows how aware Harris always was of the fact that "since the relation that makes sentences out of words is a partial order, while speech is linear, a linear projection is involved from the start" (Harris, 1988, p.24). In such a framework, then, traces constitute the mainstay for the reformulation of transformations as elementary sentence-differences.

23. To be more precise, the category at stake here is not the category of Hilbert spaces but its subcategory \mathbf{Hilb}_2 , as defined by Michael Barr.

24. A whole lot ought to be said as to whether or not that kind of spaces also constitutes the substrate for the explanatory metalanguage, such as, for instance, the conventions of autosegmental phonology regarding association lines, spreading, multiple linkage of a feature to the elements of another tier, etc. Unfortunately, that goal cannot be attained within the limited space of the present paper.

25. It should be noted, being reminiscent of some of the topics briefly mentioned in section 3.1, that a finitely generated group is context-free when its Cayley-graph has a decidable monadic second-order theory (Kuske and Lohrey, 2005). As to graphical means of performing computations, trace diagrams are of the utmost significance. They can be represented as (slightly modified) graphs in which some edges are labeled by matrices (Stedman, 1990, Cvitanović, 2008).

The information retrieval associated to sentence interpretation would then equal the process of dereferencing the pointer:

In fact, the formal mechanisms used by theoreticians are simply (within terminological changes) those used by professional programmers who specialize in the treatment of non-numerical data. For example, the dummy symbol Δ is essentially a reserved memory whose content is specified by program; the trace symbol t is an address pointer; the bar notation is an indexing device for the number of times a loop is entered, etc. Arguments about these mechanisms of abstract grammar are then isomorphic to those involved in optimization of the programming of any algorithm. The choice between two theories, e.g. between 'generative' and 'interpretative', is analogous to the choice between SNOBOL and PL/I for a given program - with the operational difference that a programmer for whom the result would be sufficiently important can always program his algorithm in both languages, and choose according to the performance of the program in each language (Gross, 1979, p. 874).

Some few words are due here with regard to the notion of pointer in computer programming. It is widely accepted that Harold Lawson invented the pointer in 1964 and integrated it in the programming language PL/I. It also became part of other programming languages (C, Pascal, C++ or Ada), founding a significant application in the operating system MULTICS, developed by MIT, Bell Laboratories and General Electric and first released in 1969.

The intimate connection between Gentzen discoveries and Computer Science is warranted by the so-called Curry-Howard isomorphism that gives expression to a profound link between mathematical proofs and computer programs. Such axis has been later extended to include category theory: thus a Curry-Howard-Lambek correspondence is originated.

As Joachim Lambek showed in the early 1970s, the proofs of intuitionistic propositional logic and the combinators of typed combinatory logic share a common theory, namely the equational theory of cartesian closed categories. Nowadays, the Curry-Howard-Lambek correspondence is used by some authors to refer to the three way isomorphism involving intuitionistic logic, typed lambda calculus and cartesian closed categories, with objects being interpreted as types or propositions, while morphisms are construed as terms or proofs.

A caveat to this: such correspondence works at the equational level, and is not the expression of a syntactic identity of structures as it is in fact the case for each of Curry's and Howard's correspondences. In other words, the structure of a well-defined morphism in a cartesian-closed category is not comparable to the structure of a proof in Hilbert-style logic or in natural deduction. Trace diagrams find additional use, via the Curry-Howard correspondence, in the framework known as Ludics dy-

namics²⁶, in which designs (*desseins*, in French) play an essential part (Faggian, 2002). Designs are to be seen as objects in between sequents and proofs, with one crucial trait, namely, they can be reconstructed from interaction traces.

In the final chapters of *Mathematical Structures of Language*, and in some of his most significant work after that, Harris touches upon applications and consequences of his analytic toolbox which were to remain, to this point, *terra ignota* for Chomsky. Becoming aware of that ought to be an exercise of historical and intellectual responsibility among linguists of all affiliations.

7. A glimpse into the days of future passed

Hinging on some seminal ideas from Gerhard Gentzen, such as his pioneer ideas on what should come to be known as Game Semantics, Jean-Yves Girard has developed Linear Logic and a Geometry of Interaction (GoI). Linear Logic can be obtained from the classical sequent calculus by restricting the structural rules²⁷. Out of that framework, Ludics was born in 2001 (Girard, 2001), and it did not take too long for its straightforward connection to linguistics to be explored (Faggian, 2002, Faggian and Hyland, 2002, Lecomte and Quatrini, 2009), the most comprehensive account of such attempt being the book *Meaning, Logic and Ludics* by Alain Lecomte (Lecomte, 2011)²⁸. In order to truly grasp all the implications derived from the idea of design one has to deal with the notion of *slice*. Another crucial idea is that of *locations*. Each formula to be decomposed receives an address. Most importantly, in this approach, traces become hypotheses (logical axioms). Examples like those at the beginning of section 3.3 of the present text are handled in the following manner: given that Move always displaces an element from the bottom to the top of the tree, it follows that not only t and t are co-indexed with *who*, but also that *who* c-commands its traces. As the reader may immediately notice, a similar logic goes for anaphora. Anaphora is seen in Ludics as a relation in which the same entity is shared between an antecedent and the anaphoric element: a trace or pronoun. Both traces and pronouns are construed as variables, the only significant difference being the emptiness of the phonological content in the former, as opposed to the lack thereof in the latter. The analysis dee-

26. The notion of design is akin to that of abstract Böhm tree, which is itself a generalization of lambda terms, as well as a concrete syntax for games.

27. Linear Logic particularly signifies itself for the elegant manner in which it handles Gentzen's *Hauptsatz*, namely, cut elimination.

28. Here I would like to heartfully thank Boris Eng, computer scientist at Université Paris Diderot (Paris 7) for his kind assistance on these technical issues. Needless to say, all eventual errors are of my own.

ply relies on Kayne's ideas of anti-symmetry and double constituency (Kayne, 1994, 2002)²⁹. As Lambek calculus teaches us, a proof tree may become a syntactic tree if being turned upside down. In the first tree, the hypothesis assumes the role played by the trace in the second one. An important difference has to be pointed out, though: in the proof tree corresponding to, say, the utterance *Mary likes a Japanese writer*, as opposed to what happens in its syntactic formalism, the quantified NP does not really move. Rather, "it is the hole ready to accept it which moves, from the complement position of the verb to the level of the whole sentence, before being absorbed by the expectation of the quantified nominal phrase" (Lecomte, 2011, p. 141).

A very powerful axiomatic framework has been proposed for GoI by Abramsky, Haghverdi and Scott (Abramsky, Haghverdi, and Scott, 2002) resorting to traced symmetric monoidal categories³⁰. A characterization of trace structures on cartesian monoidal categories, is provided in Hasegawa (Hasegawa, 1997). An equivalence between traces and parameterized fixed point operators was proved by Martin Hyland and has been exploited by this very same author and Nick Benton in several collaborative papers. A typical example of a traced symmetric monoidal category with a GoI Situation is the category of sets and partial injective functions. Such category is equipped with the tensorial structure defined by the disjoint unions of sets and functions.

8. ¿Some geometry, for a change?

The first pages of the present text dealt with the concept of abstract group as introduced by Leopold Kronecker in 1870. In a brief text which I strongly recommend to anyone interested in epistemology, the great Russian mathematician Vladimir Ígorevich Arnold reflected on that very same topic along the following lines:

What is a *group*? Algebraists teach that this is supposedly a set with two operations that satisfy a load of easily-forgettable axioms. This definition provokes a natural protest: why would any sensible person need such pairs of operations? "Oh, curse this maths" - concludes the student (who, possibly, becomes the Minister for Science in the future).

We get a totally different situation if we start off not with the group but with the concept of a transformation (a one-to-one mapping of a set onto itself) as it was historically. A collection of transformations of a set is called a group if along with any two transformations it contains the result of their consecutive application and an inverse transformation along with every transformation.

29. By considering doubling constituents into account, one can dispense with some binding principle like Condition C. Now, to me that is, in historical terms within our discipline, a perfect instantiation of the old Latin saying *naturam expellas forca, tamen usque recurret*.

30. Vladimir Ígorevich Arnold, "On teaching mathematics", *Russian Mathematical Surveys*, 53:1, 1998, p. 234.

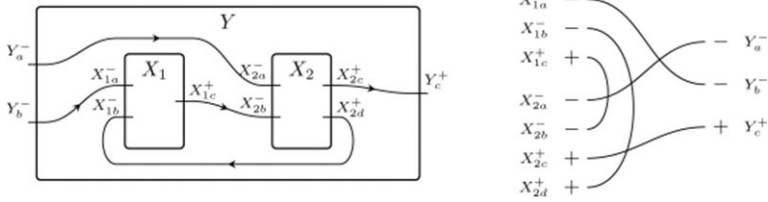
This is all the definition there is. The so-called “axioms” are in fact just (obvious) *properties* of groups of transformations. What axiomatisators call “abstract groups” are just groups of transformations of various sets considered up to isomorphisms (which are one-to-one mappings preserving the operations). As Cayley proved, there are no “more abstract” groups in the world. So why do the algebraists keep on tormenting students with the abstract definition? (Arnold, 1997).

Here I cannot but mention the significance of the Gelfand-Naimark-Segal Theorem, a theorem regulating the jump from algebra to geometry. It can be proven for such a case that to each topological space uniquely corresponds a commutative algebra, and to each commutative algebra uniquely corresponds a topological space. Crucially, the shift or leap from one domain to the other is information-preserving. However, one should never forget that what is at stake here is not an identity: there is nothing so different from a geometric extension as algebraic successions. Yet, in some cases it is possible for the former to codify the same information as the latter. That is precisely what the term *duality* intends to describe. The parallelism runs as follows: in geometry, points can be gathered to form sets, whose unions and intersections yield new sets. In algebra, the product of two functions takes the value zero on all points corresponding to zero either in one of them or in both. That is, set union in geometry corresponds to product of functions in algebra. We then make use of multiplication in order to track one of the operations on sets. In a separate paper, Irving Segal (Segal, 1947) showed that it is sufficient, for any physical system that can be described by an algebra of operators on a Hilbert space, to consider the irreducible representations of a C^* -algebra³¹.

A graphical calculus for traced categories has been put forward recently, in which compositions are represented by a diagram of the following sort (Spivak, Schultz and Rupel 2017)³²:

31. It should not shock anyone that a clear path from the study of language to that of physical systems is made clear. After all, it logically follows from the current trendy claim that language is to be seen as a natural object. Coherence with such assumption would then demand abandoning all endorsement of linguistics being associated to cognitive science, anthropology, or evolutive psychology (or biology and any form of neo-schleicherianism, for that matter). Instead it should be embedded in the supreme science of Reality, and, as part of *De rerum natura*. as in Lucretius phrasing, become a branch of Physics.

32. An online version can be read at <https://arxiv.org/pdf/1508.01069.pdf>. The referred diagram is to be found on page 3 of that article. There is always some more information in a string (or wiring) diagram for a traced category T like the one on the left side of the graphic than in the cobordism represented on the right part of the picture.



9. Epilogue

To conclude, I cannot help but refer to something I consider to be fundamental. What had started as feeble wanderings from my side found appeasing support in the somewhat casual discovery, which I owe to David Halitsky³³, that Samuel Noah Karp, a former associate and teacher at the world-renowned Courant Institute of Mathematical Sciences (the same in which Zellig Harris gave the lecture from which *Mathematical Structures in Language* originated), had clearly understood how, if one observes a two-dimensional tree on a four-dimensional space, relations between vertices or nodes are found that happen to be very useful for the study of natural and formal languages. Thus, a subtree occurring in two different places from a syntactic tree could indeed be in the same place, so that what one would be watching (with the cyclopean eye of the mind), is some sort of mirage, two different projections of a four-dimensional tree onto the dimensional plane. Interestingly enough, this, and no other, was the scenario that Chomsky’s transformational apparatus intended to capture in its origins. The reader should bear in mind that at that time the work by Dushnik and Miller (Dushnik and Miller, 1941) had not reached yet the wider audience Trotter’s book helped them achieve³⁴. A time at which Frank Harary would have allegedly told

33. Incidentally, he is the author of “A geometric model for codon recognition logic” (Halitsky, 1994), and proponent of a new classificatory scheme of the genetic code which shows the five codon symmetries on a two-dimensional table (as opposed to the usual complicated illustrations in 3D), as well as creator and advocate of the notion of the syntactic retina, which he instantiates with the help of Coxeter prisms. It has been a couple of years since I started developing certain analyses, which perhaps I will publish someday soon, of what I had come to label “detachment of language”, relying, precisely, on the retina metaphor. The reader may very well grasp how much rejoicement and self-assuredness I experienced upon discovering I was not alone in that intellectual path, for at least in its first stretches it could be travelled in good company.

34. One of the most salient developments in dimension theory since the concept was introduced by Dushnik and Miller about fifty years ago. A theorem by Schnyder characterizes planar graphs in terms of the dimensions of the associated posets.

Chomsky, according to the former's own account, that graph theory could not say anything about linguistics, for syntactic transformations were entirely arbitrary operations on trees³⁵.

Referencias

- Abramsky, Samson and Bob Coecke (2002). "Physical Traces: Quantum vs. Classical Information Processing". *Electronic Notes on Theoretical Computer Science*, 69, Special Issue: Proceedings of Category Theory in Computer Science.
- Abramsky, Samson, Haghverdi, Esfandiari and Philip Scott (2002). "Geometry of interaction and linear combinatory algebras". *Mathematical Structures in Computer Science*, 12 (5):1-40.
- Andrews, Mark William (2003). Language Learning and Nonlinear Dynamical Systems. Dissertation Presented to the Faculty of the Graduate School of Cornell University.
- Ađžaga, Nikola (2017). "Automated conjecturing of Frobenius numbers via grammatical evolution". *Experimental Mathematics*, 26 (2): 247-252.
- Arnold, Vladimir Igorevich (1997). "On teaching mathematics". Address at the discussion on teaching of mathematics in Palais de Découverte in Paris on 7 March.
- Arias, Javier (2015). "Preámbulo a un análisis de la relevancia de los estudios de Ernst Kummer sobre factorización para la historia de las teorías del lenguaje". *Eikasía*, 64, 53-80.
- Awodey, Steve (1996). "Structure in Mathematics and Logic: A Categorical Perspective". *Philosophia Mathematica*, 3, 209-237.
- Borsley, Robert D. (2012). "Don't move!". *Iberia: An International Journal of Theoretical Linguistics*, 4 (1): 110-139.
- Boskovic, Zeljko (2002). "On Multiple Wh-fronting". *Linguistic Inquiry*, 33 (3): 351-383.
- Boolos, George (1998). *Logic, Logic, and Logic*. Cambridge, Massachusetts, Harvard University Press.

35. Which at that time was fully true. Explaining to what extent that original scenario has changed along time, up to the current days, wildly exceeds the scope and goals of this paper. Neither will I value how faithful the report from Hans Hyttel, professor at Aalborg University, is, according to whom Harary told him and all participants at a seminar in 1985 that, during a stroll (which no doubt took place) Chomsky had lamented the lack of a mathematical theory that studied the connections between objects, representing the latter as vertices and the former as edges. *Se no è vero, è ben trovato*. Yet, an indisputable fact is that, in 1958, hardly a year after the publication of *Syntactic Structures*, Harary published a brief note on Carnap's reflections on asymptotic relative frequencies of ordered pair and their relation to linear graphs.

- Ceccherini-Silberstein, Tullio and Wolfgang Woess (2003). "Growth-sensitivity of context-free languages". *Theoretical Computer Science*, 307, 103 – 116.
- Cecchetto, Carlo (2004). "Remnant Movement in the Theory of Phases", in Luigi Rizzi (ed.). *The Structure of CP and IP*, Oxford, Oxford University Press, 166-189.
- Chomsky, Noam (1964). *Aspects of the Theory of Syntax*, Cambridge, Massachusetts, MIT Press.
- Chomsky, Noam (1973). "Conditions on Transformations", in Anderson and Kiparsky (eds.), *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston, 232–286.
- Chomsky, Noam (1995). *The Minimalist Program*, Cambridge, Massachusetts, MIT Press.
- Clark Stephen and Stephen Pulman (2007). "Combining Symbolic and Distributional Models of Meaning", *Papers from the American Association for Artificial Intelligence*, Stanford, Stanford University, 52-55.
- Coecke, Bob, Mehrnoosh Sadrzadeh and Stephen Clark (2010). "Mathematical foundations for a compositional distributional model of meaning", in Johan van Benthem, Michael Moortgat, and Wojciech Buszkowski (eds.), *Linguistic Analysis (Lambek Festschrift)*, 36, 345-384.
- Corver, Norbert and Jairo Nunes (2007). *The copy theory of movement*. Amsterdam, John Benjamins.
- Cruys, Tim van de, Thierry Poibeau and Anna Korhonen (2013). "A Tensor-based Factorization Model of Semantic Compositionality". *Conference of the North American Chapter of the Association of Computational Linguistics*, 1142-1151.
- Cvitanović, Predrag (2008). *Group Theory: Birdtracks, Lie's, and Exceptional Groups*, New Jersey, Princeton University Press.
- Dummet, Michael (1995). *Frege: Philosophy of Mathematics*. Cambridge, Massachusetts, Harvard University Press.
- Dushnik, Ben and E.W. Miller (1941). "Partially ordered sets". *American Journal of Mathematics*, 63 (3): 600-610.
- Samuel Eilenberg (1974). *Automata, Languages and Machines*, Vol. A. New York, Academic Press.
- Samuel Eilenberg (1976). *Automata, Languages and Machines*, Vol. B. New York, Academic Press.
- Faggian, Claudia (2002). "Travelling on designs: Ludics dynamics", *CSL'02 Proceedings of the 16th International Workshop and 11th Annual Conference of the EACSL on Computer Science*, 427-441.

- Faggian, Claudia and Martin Hyland (2002). "Designs, Disputes and Strategies", CSL'02 Proceedings of the 16th International Workshop and 11th Annual Conference of the EACSL on Computer Science, 442-457.
- Fredholm, Ivar (1900). Sur une nouvelle methode pour la resolution du probleme de Dirichlet", Stockholm, Kong. *Vetenskaps Akademieus Förk*, 39-46.
- Fredholm, Ivar (1903). "Sur une classe d'équations fonctionnelles". *Acta Mathematica*, 27, 365-390.
- Gelfand, Israel and Mark Naimark (1943). "On the imbedding of normed rings into the ring of operators on a Hilbert space". *Matematicheskii Sbornik*, 12 (2): 197-217.
- García Calvo, Agustín (1983). De la construcción (Del lenguaje II), Zamora, Lucina.
- Girard, Jean-Yves (2001). "Locus solum: from the rules of logic to the logic of rules". *Mathematical Structures in Computer Science*, 11, 301-506.
- Grewendorf, Günther (2001). "On Multiple Wh-Fronting". *Linguistic Inquiry*, 31(1): 87-122.,
- Grewendorf, Günther (ed.) (2015). Remnant Movement. Berlin, De Gruyter Mouton.
- Gross, Maurice (1979). "On the Failure of Generative Grammar". *Language*, 55(4): 859-885.
- Halitsky, David (1994). "A geometric model for codon recognition logic". *Mathematical Biosciences*, 121 (2): 227-34.
- Grothendieck, Alexander (1955). "Produits tensoriels topologiques et espace nucléaires", Memoirs of the American. *Mathematical Society*, 16, 140.
- Grothendieck Alexander (1956). "La theorie de Fredholm". *Bulletin de la Société Mathématique de France*, 84, 319-384.
- Harris, Zellig Sabbetai (1979). *Mathematical Structures of Language*, New York, Krieger Publishers (1st edition from 1968).
- Harris, Zellig Sabbetai (1988). *Language and Information*. New York, Columbia University Press.
- Hasegawa, Masahito (1997). "Recursion from Cyclic Sharing: Traced Monoidal Categories and Models of Cyclic Lambda Calculi", in Philippe de Groote and J. Roger Hindley (eds.), *Typed Lambda Calculi and Applications*, Berlin, Springer, 196-213.
- Joyal, André, Ross Street and Dominic Verity (1996). "Traced monoidal categories", *Mathematical. Proceedings of the Cambridge Philosophical Society*, 119 (3): 447-468.
- Kandybowicz, Jason (2008). *The Grammar of Repetition: Nupe grammar and the syntax-phonology interface*, Amsterdam, John Benjamins.

- Kayne, Richard S. (1994). *The Antisymmetry of Syntax*, Cambridge, Massachusetts, MIT Press.
- Kayne, Richard S. (2002). "Pronouns and their antecedents", in Samuel David Epstein and T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, Oxford, Blackwell, 133–166.
- Koopman, Hilda and Anna Szabolcsi (2000). *Verbal Complexes*. Cambridge, Massachusetts, MIT Press.
- Krieger, Martin (2003). *Doing Mathematics: Convention, Subject, Calculation, Analogy*, New Jersey, World Scientific Publishing.
- Kaufman, Bruria and Lars Onsager (1949). "Crystal statistics. III. Short-range order in a binary Ising lattice". *Physical Review*, 76, 1244-1252.
- Kronecker, Leopold (1870). "Auseinandersetzung einiger Eigenschaften der Klassenzahl idealer complexer Zahlen", *Monatsbericht der Königlich-Preussischen Akademie der Wissenschaften zu Berlin*, 881-889.
- Kuske, Dietrich and Markus Lohrey (2005). "Logical Aspects of Cayley-Graphs: The Group Case". *Annals of Pure and Applied Logic*, 131, 263-286.
- Lasnik, Howard (2011). "The notion of derivation in linguistics: Syntax". *50 Years of Linguistics at MIT*, MIT.
- Lecomte, Alain (2011). *Meaning, Logic and Ludics*. London, Imperial College Press.
- Lecomte, Alain, and Myriam Quatrini (2009). "Ludics and Its Applications to Natural Language Semantics", in Hiroakira Ono, Makoto Kanazawa and Ruy de Queiroz (eds.), *Logic, Language, Information and Computation*, Berlin, Springer, 242-255.
- Lentin, André (2002). "Reflections on references to mathematics in the work of Zellig Harris", in Bruce Nevin (ed.), *The Legacy of Zellig Harris: Language and Information into the 21st Century*, vol. 2: Mathematics and computability of language, Amsterdam, John Benjamins, 1-9.
- Lidskil, Vladimir Borisovich (1958). "Theorems on the completeness of a system of characteristic and adjoint elements of operators having a discrete spectrum". *Doklady Akademii Nauk SSSR*, 119, 1088-1091.
- Lightfoot, David (1976). "Trace Theory and Twice-Moved NPs". *Linguistic Inquiry*, 7 (4): 559-582.
- McCarthy, John (1981). "History of Lisp". Paper presented at the ACM SIGPLAN History of Programming Languages Conference in June 1978, then included in Richard Wexelblat (ed.), *History of Programming Languages*, New York, Academic Press, 173-197.
- McIntosh Harold V. (2009). *One Dimensional Cellular Automata*, Frome, Luniver Press, 3-4.

- Morrill, Glyn (2011). "The Displacement Calculus". *Journal of Logic Language and Information*, 20 (1): 1-48.
- Nagy, Benedek (2010). "Graphs of Grammars – Derivations as Parallel Processes", in *Computational Intelligence in Engineering, Studies in Computational Intelligence*, SCI 313, Springer, 1-13.
- Neumann Peter (2001). Introduction to the theory of finite groups. Conference given at the University of Sussex on 19 March.
- Nevin Bruce (2010). "Noam and Zellig", in Douglas A. Kibbee (ed.), *Chomskyan (R) evolutions*, Amsterdam, John Benjamins, 103-168.
- Nevin, Bruce (2009). "More Concerning the Roots of Transformational Generative Grammar". *Historiographia Linguistica*, 36 (2): 461-481.
- Nielsen, Michael A. and Isaac L. Chuang (2000). *Quantum Computation and Quantum Information*, Cambridge, Cambridge University Press.
- Postal, Paul Martin (2004). *Skeptical Linguistic Essays*. Oxford University Press, New York.
- Pullum, Geoffrey K., Robert D. Borsley (1980). "Comments on the two central claims of 'trace theory'". *Linguistics*, 18, 73-104.
- Pullum, Geoffrey Keith (2010). "Creation myths of generative grammar and the mathematics of Syntactic Structures", in Christian Ebert, Gerhard Jäger and Jens Michaelis (eds.), *The Mathematics of Language*, Berlin, Springer, 238-254.
- Radford, Andrew, Martin Atkinson, David Britain (1999). *Harald Clahsen and Andrew Spencer, Linguistics: An Introduction*, Cambridge, Cambridge University Press.
- Razet, Benoit (2008). "Finite Eilenberg Machines" in Ibarra Oscar H. and Bala Ravikumar (eds.), *Implementation and Application of Automata*, Berlin, Springer, 242-251.
- Sag, Ivan A. and Janet Dean Fodor (1994). "Extraction without Traces", in Raul Aronovich, William Byrne, Susanne Preuss, and Martha Senturia (eds), *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. Stanford, California, CSLI Publications, 365-384.
- Segal, Irving (1947). "Irreducible representations of operator algebras". *Bulletin of the American Mathematical Society*, 53, 73-88.
- Stabler, Edward (1999). "Remnant movement and complexity", in Gosse Bouma, Erhard Hinrichs, Geert-Jan Kruijff, and Richard Oehrle (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, Stanford, CSLI Publications, 299-327.

- Stedman, Geoffrey E (1990). *Diagram Techniques in Group Theory*, Cambridge, Cambridge University Press.
- Spivak, David, Schultz, Patrick and Ruper, Dylan T. (2017). "String diagrams for traced and compact categories are oriented 1-cobordisms." *Journal of Pure and Applied Algebra*, 221(8): 2064-2110.
- Sussman, Gerald Jay, and Jack Wisdom (2001). *Structure and Interpretation of Classical Mechanics*.
- Trotter, William Thomas (1992). *Combinatorics and Partially Ordered Sets: Dimension Theory*, Baltimore, The Johns Hopkins University Press.
- Weber Heinrich (1895-1896). *Lehrbuch der Algebra*, Braunschweig, F. Vieweg und Sohn (two volumes).
- Wussing, Hans (1969). *Die Genesis des abstrakten Gruppenbegriffes*, Berlin, VEB Deutscher Verlag der Wissenschaften.

Sobre el autor

JAVIER ARIAS NAVARRO es profesor de Lingüística, Universidad de Lisboa, Doctor en lengua española. Correo Electrónico: jariasnavarro@yahoo.es

CUHSO. CULTURA-HOMBRE-SOCIEDAD

Fundada en 1984, la revista CUHSO es una de las publicaciones periódicas más antiguas en ciencias sociales y humanidades del sur de Chile. Con una periodicidad semestral, recibe todo el año trabajos inéditos de las distintas disciplinas de las ciencias sociales y las humanidades especializadas en el estudio y comprensión de la diversidad sociocultural, especialmente de las sociedades latinoamericanas y sus tensiones producto de la herencia colonial, la modernidad y la globalización. En este sentido, la revista valora tanto el rigor como la pluralidad teórica, epistemológica y metodológica de los trabajos.

EDITOR
Matthias Gloël

COORDINADORA EDITORIAL
Claudia Campos Letelier

CORRECTOR DE ESTILO Y DISEÑADOR
Ediciones Silsag

TRADUCTOR, CORRECTOR LENGUA INGLESA
Aurora Sambolin Santiago

DESARROLLADOR DE SISTEMAS
Laura Navarro Oliva

SITIO WEB
cuhso.uct.cl

E-MAIL
cuhso@uct.cl

LICENCIA DE ESTE ARTÍCULO
Creative Commons Atribución Compartir Igual 4.0 Internacional