Language Combinatorics as Matrix Mechanics

Juan Uriagereka (University of Maryland), Roger Martin (Yokohama National University),
Ángel Gallego (Universitat Autònoma de Barcelona)

In this talk (and the following talk), we will propose an analysis of linguistic categories and their possible combinations by way of syntactic Merge in terms of matrix mechanics. We begin by considering Chomsky's (1974) distinctions for the "parts of speech" in (1).

(1) a Noun: [+N, -V]
b. Adjective: [+N, +V]
c. Verb: [-N, +V]
d. Adposition: [-N, -V]

To this, we apply the Fundamental Assumption in (2):

(2) a. ±N is represented as ±*1*
b. ±V is represented as ±*i*

It is easy to see that this assumption results in vectors as in (3), a minimal extension of which (substituting the vectors for the matrix diagonal) gives us the matrices in (4):

(3) a. [1, -*i*] b. [1, *i*] c. [-1, *i*] d. [-1, -*i*].

(4) a. $\begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}$ b. $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ c. $\begin{pmatrix} -1 & 0 \\ 0 & i \end{pmatrix}$ d. $\begin{pmatrix} -1 & 0 \\ 0 & -i \end{pmatrix}$

Matrices such as (4) are very well understood, and under common assumptions concerning the anti-symmetry of Merge can be shown to yield an interesting group. We assume the following representations of two types of Merge:

(5) a. First Merge (for head-complement relations) is represented as matrix multiplication.
b. Elsewhere Merge (for head-specifier relations) is represented as a tensor product.

Interestingly, applying (5a) reflexively (i.e., self-Merge) to any of Chomsky's matrices interpreted as in (4) yields the same result:

(6) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

(6) is the well-known 3rd Pauli matrix, Z. Further first-mergers among the Chomsky matrices in (4) or the output of these combinations yields three more matrices within the Pauli group:

(7) a. $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ b. $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ c. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

We further argue that the Pauli matrices in (8a) and the matrices in (8b), which include the two positive Chomsky matrices and their anti-variants, form a 32-element group in (9), which we tentatively dub the Chomsky-Pauli Group.

(8) a. $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

b. $C_1 = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}$ $C_2 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ $S_1 = \begin{pmatrix} 0 & 1 \\ -i & 0 \end{pmatrix}$ $S_2 = \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}$

(9) $G_{cp} = \{\pm I \pm X \pm Y \pm Z \pm iI \pm iX \pm iY \pm iZ \pm C_1 \pm C_2 \pm S_1 \pm S_2 \pm iC_1 \pm iC_2 \pm iS_1 \pm iS_2\}$

Just as the Chomsky matrices in (4) correspond to lexical categories, we might expect "grammatical categories" to correspond to other elements within the group. Moreover, only certain correspondences between "heads" (lexical items in the group) and "projections" (other elements in the group) yield "endocentric" structures. This is particularly so if we interpret "syntactic label" as in (10):

(10) The "label" of a matrix obtained by Merge is its determinant (= product of the items in the matrix's main diagonal minus the product of the items in the off-diagonal).

Given this notion of "label," the only first-mergers that yield "projected" head-complement structures are those in (11).

(11) a. The complement of a noun/adjective is a PP (or a related grammatical projection).
b. The complement of a verb/preposition is an NP (or a related grammatical projection).

We also demonstrate that the only elsewhere-mergers that yield similar symmetries, and can thus said to be "projected" head-specifier structures, are those described in (12).

(12) The specifiers of any type of category is an NP (or a related grammatical projection).

The tensor-product space resulting from multiplying all 32 matrices among themselves also has interesting properties. Among the 1024 tensor-products that the group allows, several matrix combinations are *orthogonal*. Orthogonal matrices, when added, yield objects with a "dual" character of the sort seen, for instance, in UP and DOWN situations of an electron's angular momentum – the probability of each relevant state being half – yielding a characteristic "uncertainty" in which the two states are not simultaneously realizable. We have a mathematically identical scenario, this time involving two orthogonal specifiers, which can be said to literally *superpose*. Thus the behavior of so-called copies in (13) – e.g., the fact that it can be pronounced UP (13b) or DOWN (13c), but not in both configurations (13d) (similar considerations obtain for interpretation) – can be deduced as superposition within this tensor-product space, so long as we define "chains" as in (14).

(13) a. [Armies [appeared [armies ready]]]
b. Armies appeared ready.
c. There appeared armies ready.
d. *Armies appeared armies ready.

(14) A chain C = (A, B) is the sum of two orthogonal tensor-product matrices A and B.

The conditions that yield chain superposition entail the possibility of entanglement of different chains into a super-chain, which is arguably what we see in (15).

(15) a. [Armies tried [armies to [appear [armies ready]]]]
b. Armies tried to appear ready.