

***Co-experiencing prosody:
The distributed motives and effects of prosody in languaging processes***

The speaking activity can be described in terms of vocal action: sequences of articulated segments of all kinds (words, morphemes...), trains of templates to be filled with phono-semantic matter (“syntactic structures”), some of them governed by lexical heads (“constructions”). Speech occurs when a voicer embarks on causing a hearer to construct meaning, giving the impression that intended meaning or representations (with due reference to “reality”, both situational and intersubjective, including the discourse itself) are “encoded” before emission and “decoded” in the aftermath, with acoustic forms symbolizing cognitive events on both sides (production and / or interpretation).

Yet, this conception does not resist a closer investigation into the theoretical constraints of the empirical object: not only is the vocal output also perceived by the very voicer (through auditory and tactile proprioception), but “languaging” (in Maturana’s sense) is also used privately for verbal thinking, with at least some of the prosodic features of orality (melody, rhythm, intensity, periods, word and sentence stress, tonemes...), and with no possibility to invoke encoding and expression.

A reasonable alternative hypothesis is that to speak is to direct the sense-making process, the thought-generating dynamics through a semi-rigid, semi-flexible sequence of vocal actions that are reminiscent of similar vocal and auditory occurrences of experience in which the same processes have already been implemented in varying contexts and associative networks: a speaker will either guide an external’s addressee’s cognitive path and “adventure” vocally (through the acoustic perturbation of the environment), or guide his *own* ideating process (through the projection of what the perceived signal would sound like if it were actually voiced). The “thought” achieved by verbal means is not the “final representation” eventually obtained, but rather, the dynamic sequence or “film” of semantic events activated by the vocal keys and matched against both idiosyncratic and conventional psychological backgrounds, with the versatile dialectics of convergence and divergence this entails for “interpretation”.

In this framework, lexical and grammatical units are redefined in terms of vocal keys activating profiled patterns (retrieving lexical bundles of features of all sorts formed in the course of the personal experiences streamlined by social life and “culture”, linking them with grammatical profiles implying psychological positioning (space, time, modality, intersubjectivity...). Language-specific categories and syntax evidence a typological diversity of scenarii that vary according to both the nature of the operations developed in the course of the language’s history and the order in which they are carried out. Syntax, in particular is versatile: it predetermines an “orthosyntactic agenda” like the canonical NP-VP sequence in English but it is possible to improvise all sorts of deviations (pre-initial addenda like adverbials, transitional addenda like insertions, and so on), and each template can be matched with varying instantiations, either ready-made ones (pronouns...) or improvised ones (phrases, clauses), with specific entries enacted by introductory markers (determiners, prepositions, complementizers, auxiliaries).

This model can operate if and only if prosody is given a key role in the emotional control of syntactic decisions: after a subject, improvising to deviate from the planned route (NP) and open up a parasyntactic excursus requires both the interruption of the predetermined path (the VP) and its resumption. Some prosodic actions are the vocal counterpart of emotion-driven metalinguistic and syntactic decision in real time (with the working memory this requires *for both*), and they are perceived by both interlocutors, who are both trained (by regular practice) to rely on them as *symptoms* of the ongoing process – which does not make the latter intentionally controlled coded signs in spite of their relative conventionality. Thus, prosody conflates three forms of emotion-based dynamics : subjective emotion (mood...), intersubjective emotion (requiring, seducing, directing...), and metalinguistic emotion (structural planning, current execution and retrieval, plus the proprioceptive feedback and rerouting it enables). These proposals will be illustrated and located in the context of current trends (auto-poiesis and enaction, embodiment and distributed cognition).