

Compositionality: The Formation of a Learning Theory

Samir Karmakar

A learning theory should try to answer the following questions. What does it mean to learn? How is something learnt? How is the learnt information stored, processed and ultimately translated into the speech? And apart from these; one crucial question relating to the learning is how does one manipulate the meaning out of the sentential output? The main topic of this article is related with the last question. It is quite clear, that manipulation of the meaning is directly related with the question of interpretation. More briefly, how does the mind interpret an output? Therefore to give an answer to the previous question is basically an effort to define the nature and type of the mind. Since any kind of knowledge by its nature is essentially heterogeneous, the present field of study is also not an exception. Therefore, a short account of this multitude will not be an exaggeration in this context.

The two limiting conditions are set by two prominent schools of Behaviorist and Nativist. In both these two cases learner is considered as an isolated individual, but obviously in a contrastive way. A behaviorist will consider the learning as a behavior formation, under the notion of the conditioned response, whereas a nativist will consider learning under the general rubric of the biological trait and natural selection. And between these two extremes there are some other schools of thoughts, intersecting each other. Among them the school of Constructivists earns a good currency. To them the language is a consequence of the interaction. Interestingly, when a nativist claims the biological foundation of the language, a constructivist simply argues in favour of cultural evolution. Within this theoretical whole, a learner reaches its coherence in terms of compositionality by self-organizing himself, while interacting.¹ On the other hand side, the Cognitivist's view also radically differs from the others, since they argue that the learning a language is presumed to be the product of the general learning capacity. And if we include the view of Vygotsky's "social constructivism", the situation will become much more perplexing one. Vygotskyian view tells us that a matrix of social history and form triggers all kind of cognitive development, including learning. All these views constitute the background of the learning from different point of views.

Under this situation any kind of theory attempting to answer the manipulation process will maximally gain the status of hypothesis. Learning process should not imply only the accumulation of the linguistic ability but also relates the capacity to judge the well-formedness of an utterance. It should also explain the systematic productive character of the human mind. In consonance with this we should also argue that mind as a linguistic

¹In connection of this issue we can refer Voloshinov (1929), a much debated and politically exiled Marxist thinker. The stress on interaction, in case of developing a system of signification can also be noticed here. "Signs can arise only on *interindividual territory*. It is territory that cannot be called "natural" in the direct sense of the word: signs do not arise between any two members of the species *Homo Sapiens*. It is essential that the two individuals be *organized socially*, that they compose a group (a social unit); only then can the medium of signs take shape between them. The individual consciousness not only cannot be used to explain anything, but on the contrary, is itself in need of explanation from the vantage point of the social, ideological medium."

capacitance has some sort of interpretive capability. Once we can explain the nature and structure of this interpretive capacitance, the model of the mind as a linguistic module will automatically come into the picture. It is the structure of the interpretive capacitance of the mind which in itself can propose a more plausible answer to our initial queries. Because as a structure it reflects its diachronic formation and as well as an answer to its synchronic accession.

Here in this paper we will (i) examine two hypotheses – one proposed by Chomsky and another by Jackendoff. We will (ii) judge their basic theoretical architectures, (iii) logical congruity between these two proposals and its (iv) implication to attain the compositional aspect of the language. Finally, at the end we will try (v) to justify the importance of this paper in case of learning theory.

Both the theoretical architectures, proposed by Chomsky and Jackendoff, include a processor and an interpreter. The task of the processor is the central one, in the sense that it can generate an array of structures; whereas an interpreter can do basically two types of works - one, it can interpret the incoming information to decipher the meaning, and two, it can encode a meaning out of its constituent parts. It is quite interesting that in both of these cases information is considered as the structured one.

Let us try to explain the apparent differences of the postulations, specific to their authors. In brief the school led by Chomsky, express their view in support of a biologically endowed innate system, which is autonomous in contrast with other cognitive systems. On the other hand side, the linguists like Jackendoff have expressed their firm faith on the fact that the learning of the language should be treated as a part of whole cognitive development. These two distinct positions arise because of the nature and level of their scholastic enquiries. For the first group the major research question is to develop a theory of syntax, whereas, for the latter one, the prime interest is the semantics. But at the same time some things are common in these two approaches. In case of designing their theoretical tool they have given emphasis on determining the conceptual primes of their related fields, a set of rules and ultimately two different modules, such as argument structure and theta theory, are accepted to restrict the domain of well formed utterances.

All the proposed schemes, no matter whether it is proposed by Chomsky or Jackendoff or any other, share a lot in their corresponding theoretical architectures, in spite of their differences, because of the nature of the problem itself. Within their theoretical architectures, both the scholars have made a distinction between I-language and E-language, to demarcate their object of enquiry, though Jackendoff's approach to lexical semantics is an effort to go beyond the syntax by means of setting up the correspondence between the syntax and semantics. Both of them have proposed that the language as an E-language is the projection of the internally encoded information, which is popularly known as I-language in the literature.

It is the internal form of the language, which, one has to acquire to have mastery over his or her language. If this assumption is true then the question should be how it looks like. No doubt it is not like the language one encountered in the daily life. The internal form of

the language is the very interpretive capacitance about which we have mentioned earlier. This interpretive capacitance can be decomposed into a finite set of conceptual primes, a finite set of relation and ultimately into a cognitive architecture, in such a manner as it is done in other empirical sciences. One of the basic premises of this kind of reductionism is that each and every phenomenon or object of this world is made up of some primary elements. Whatever differences one can meet in the surface level is because of their idiosyncratic structural composition of those conceptual primes.² Or more explicitly, an object of the surface level, as a whole, is the extension of the intentional core. The first and foremost task for a researcher is neither to explain the extensional world nor to measure the abysmal depth of the intentional core, but to explore the causal path through which extensional comes into the existence out of the intentional one. And this is the way the saga of compositionality reveals itself, as the central issue of the learning theories.³

Developing a theory of the interpretive capacitance always needs apriori assumption of the intentional core, in the sense the physicist has to assume the atomic level, out of which extensional world projects itself. The task of the physicist never ends with the imagination of the primitive level, but the ultimate goal is to draw a causal path through which the primitives of the atomic level composed themselves into the complexity of the surface level. It is this causal path which in itself is the meaning of the entire existence. Same is also true for the learning theories. Let us see how it works in the theoretical architecture of Jackendoff.

To Jackendoff a lexical item is the collection of significant syntactic and conceptual information. In case of analyzing the conceptual structure, he assumes a repertoire of major conceptual categories, which includes following entities: Thing, Event, State, Action, Place, Path, Property, and Amount. In general the categories are rewritten in the following format:

$[X]_{\text{category}} \rightarrow [X f(x)^{(A_n, \text{ such as } 0 \leq n)}]$, where a conceptual category X is represented as the function of X, along with its optional argument(s) (A_n), if it is a verb. It can be exemplified by the following example:

1a. Bill ran toward the house.

1b. Syntactic structure: $[_S [_{NP} \text{Bill}] [_{VP} \text{ran} [_{PP} \text{toward} [_{NP} \text{the house}]]]]]$

1c. Conceptual structure: $[_{\text{Event}} \text{GO} ([_{\text{Thing}} \text{BILL}], [_{\text{Path}} \text{TO} ([_{\text{Thing}} \text{House}])])]$

The structure of the interpretive capacitance, in Jackendoff's approach, considers a lexical entry as a set of information Σ , such as $\Sigma = (\omega, \kappa, \delta, \sigma)$, where ω = word form, κ = grammatical category, δ = distribution of the word form in a syntactic framework and

²One good example in support of our claim can be produced from the organic chemistry. Consider the compound C_5H_8 , which has two different manifestations, depending on their idiosyncratic structural compositions, and they are as follows: (i) $H_3C-C \equiv C-CH_2-CH_3$, (ii) $H_2C=CH-CH_2-CH=CH_2$. It is the inter and as well as intra structural relations which determine their corresponding properties, though they are made up of same number of atoms of same primary elements.

³“Compositionality is the property that a system of representation has when (i) it contains both primitive symbols and symbols that are syntactically and semantically complex; (ii) the latter inherit their syntactic / semantic properties from the former.” – Fodor and Lepore.

finally σ = the conceptual structure of the word form. If it is so, then learning, to Jackendoff, is to acquire the knowledge of the lexical entries. Let us see what it means by lexical knowledge and how is it represented in the hypothetical world of Jackendoff.

2.

ω	toward	run
κ	P	V
δ	$_ NP_j$	$_ (PP_j)$
σ	$[\text{Path TO } ([\text{Thing } _j])]$	$[\text{Event GO } ([\text{Thing } _i, [\text{Path } _j])]$

Note that the index “i” is generally used to signify the first argument.

This kind of conglomeration of significant syntactic and conceptual information simply directs towards the fact that what ever conceptual structure Jackendoff has proposed have to always mediate and verified through the syntactic level, to attain the greater degree of regularity.⁴ Therefore the genesis of a learning theory is strictly depend on the syntactic counterpart, and further narrowing down the query will reveal the ultimate necessity of the compositionality issue. What Jackendoff actually proposed can be generalized in our following formal argument.

Let us assume that μ and σ are the set of meaning and the set of sentence respectively. Moreover, Π is a set of ordered pairs which represent the association between μ and σ . It is the association which is embedded within the Π , inferring the intentional core of the capacitance.

$\therefore \Pi \subseteq \sigma^* \mu, \forall s \in \sigma$ and $m \in \mu$; which means Π is set of all possible ordered pairs, such as (s, m) , and it can be read as “s composes m” and vice versa. If this statement is true, then the defining criteria (say f) of Π will be the composition of σ and μ , mapped over σ . This can be mathematically represented in the following form:

f: $\sigma^* \mu \rightarrow \sigma$; f can also be considered as the quotient of compositionality, whereas * represents “composition”.

In case of most of the learning theories, such as the Jackendoff’s one, the investigation starts with an effort to explore the nature of f, i.e. the compositional quotient; because it plays a crucial role by providing an explanation to the causal development of the extensional layer out of the apriori intentionality. Compositional quotient also plays a crucial role to block the generation of the ill-formed construction. The violation of the compositional property simply leads to the ill-formed sentences, as it happens in the following sentence:

3a. Sincerity ran into the room.

But how does it actually work within the theoretical periphery of Jackendoff’s Semantic Structure? Let us see. The conceptual structure of 3a will be as follows:

3b. Conceptual Structure: $[\text{Event GO } ([\text{Property SINCERITY}], [\text{Path TO } ([\text{Thing House}])])]$

⁴ “The first is that, without an appreciation of the syntactic structure of a language, the study of lexical semantics is bound to fail.” – Pustejovsky, James.

This type of conceptual structure (3b) for *run* as a lexical entry is not permissible, because a Property as a theme can not be the argument of a spatial motion. It should be the thing, as it is already indexed with “i” in table 2. Moreover the most crucial thing, what is to be noted here, is that in case of determining the criteria of semantically well-formed construction, Jackendoff has taken an account of the thematic roles, along with the argument structure in his theory, in a similar fashion as it is done by Chomsky by introducing θ -criterion in the level of LF. Therefore what ever sub-categorization rule is included as a conceptual structure for the lexical entry *run* has also its counterpart in the level of syntax, basically dealt with in the room of C-I interface, with the measuring tools of LF. We can substantiate this fact from our previous example of 1b and 1c. The NP in the subject position of verb *run*, is equivalent to the first argument of the Event function GO, which is marked with the index “i”, and PP has the corresponding Path function of TO. Not only that in case of 1b, it is quite clear from the θ -criterion that *run* can only take animate concept as its theme or agent. Therefore the lexical insertion rule automatically blocks the use of *sincerity* in the subject position in 3a.

Now the notion of compositionality is not always as simple as the above mentioned one. It is found that some of the time the meaning does not bear a one to one correspondence between the syntactic level and the conceptual level. To deal with this kind of problem Jackendoff has introduced the concept of explicit vs. implicit theta role. For example consider the following example:

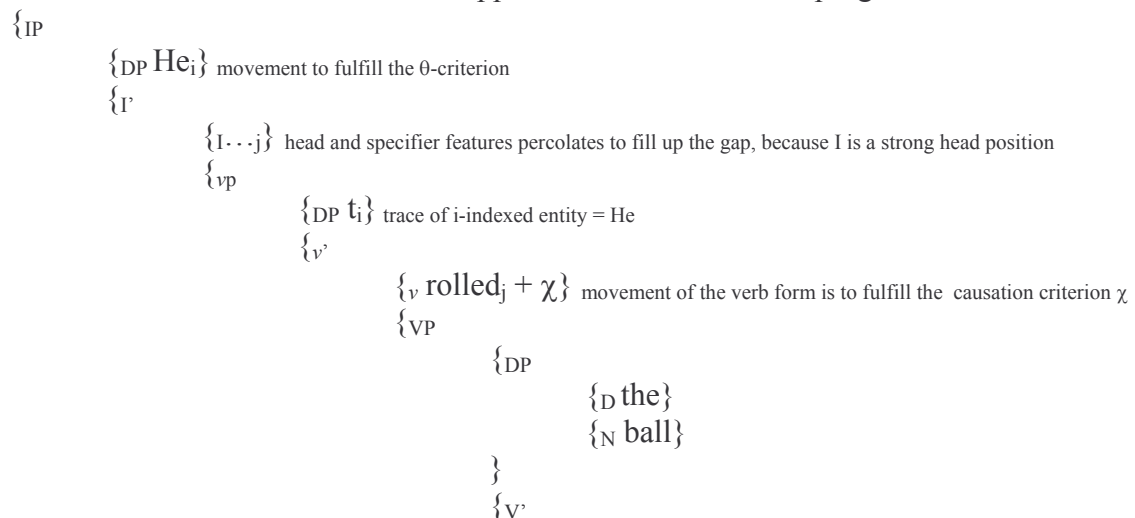
4a. He rolled the ball down the hill.

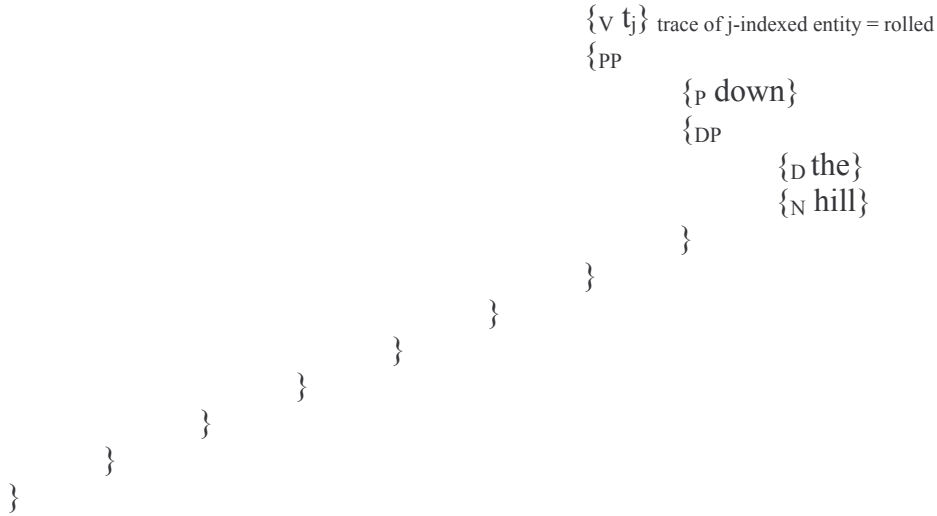
This sentence expresses an *Event* which tells us about the rolling of the ball down the hill. But as an inanimate object, the ball has hardly any capacity to roll down the hill, with out the act of an agent, which is the cause of this event. Therefore the hidden conceptual interpretation of 4a will be as follows:

4b. He caused the ball to roll down the hill.

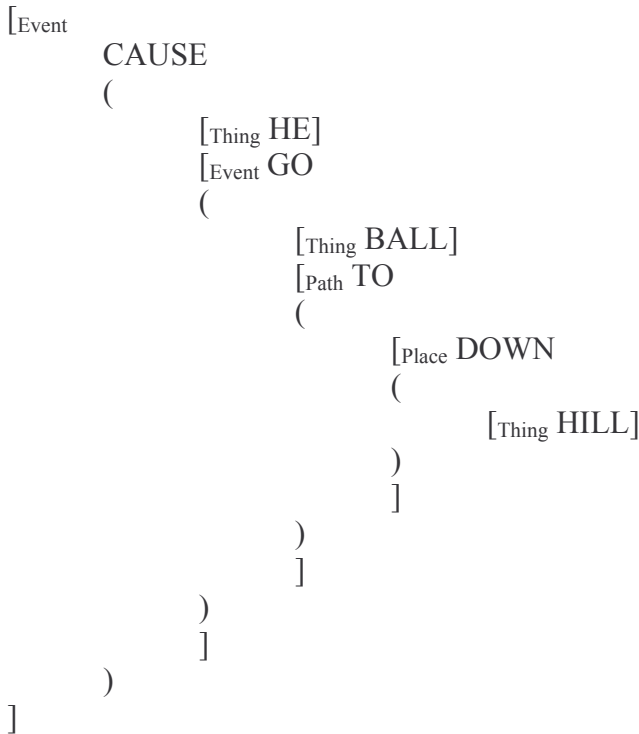
4c. Conceptual structure: [_{Event} CAUSE ([_{Thing} HE], [_{Event} GO ([_{Thing} BALL], [_{Path} TO ([_{Place} DOWN ([_{Thing} HILL)]))])])])]

Same kind of formalization is also available in the minimalist program of Chomsky, where the myth of causative marker (say χ) is introduced. Let us see how this sentence will be dealt with in the theoretical approach of the minimalist program.





But what is the need of this formal discussion? How are they basically related with the development of the learning theory? If learning a language means to acquire the mastery over the compositional property of the language it is quite obvious to know how it works to project the underlying thought into the surface level and vice versa. So Jackendoff is also taken a venture, resonating with the previous one, to highlight the compositional issue.



To deal with the notion of “compositionality”, Chomsky has introduced a level of C-I interface, where the underlying logical form (LF) of the sentence is dealt with, in terms of movement, trace, binding etc. to fulfill the criteria of theta theory and argument structure; whereas in case of Jackendoff, same purpose is attained by introducing a repertoire of

conceptual categories, along with the rule of argument fusion. In both these two cases all these theoretical apparatus are introduced to capture the notion of compositionality, which has a crucial role in case of learning a language. Particularly, as it is found even in the work of some other scholars that compositional aspect of the verb, in case of child language acquisition, emerges out due to the constraint, imposed by the corresponding argument structure.⁵

Learning a language is primarily about the way or the manner through which it comes into existence that means a gradual attainment of the compositional potentiality of a language – a mapping between the intentional and extensional issues, through a causal path, which as a *modus operandi*, ultimately projects itself in the status of *modus essendi*, i.e. compositionality. Therefore any kind of learning theories should reflect the compositional aspect of the language. Because it is the compositionality which relates thought with utterance.

-The End-

References:

- Chomsky, N. (1995). “Minimalist Program”. MIT Press. Cambridge. MA.
Cooper, William S.. (1978). “Foundations Of Logico-Linguistics”. D. Reidel Publishing Company. Holland.
Fodor, Jerry A. and Lepore, Ernest. 2002. “Why Meaning (Probably) Isn’t Conceptual Role”. Published in “The Compositionality Papers”. Clarendon Press. Oxford.
Grayling, A. C.. (2001) “Wittgenstein”. Oxford University Press. New York.
Jackendoff, Ray. (1990). “Semantic Structure”. MIT Press. Cambridge. MA.
Pinker, Steven. (1995). “Language Acquisition”. This excerpt from, “An Invitation to Cognitive Science”. 2nd edition. Edited by Gleitman, Lila R. and Liberman, Mark. MIT Press.
Pustejovsky, James. (1995). “The Generative Lexicon”. MIT Press. Cambridge. MA.
Voloshinov, V. N.. (1929). “Multiaccentuality And The Sign”. Published in “The Routledge Language And Cultural Theory Reader” (2000). Routledge. London.

⁵ “A general principle of argument structure is that the argument that is affected in some way (the particular way is specified by the verb) gets mapped onto the syntactic object. This is an example of a “linking rule”, which links semantics with syntax (and which is an example of the contingency a young child would have employed to use semantic information to bootstrap into the syntax).”- Pinker, Steven.