

Sentencehood*

Isso Nakamura

University of Tokyo/Japan Society for the Promotion of Science

This paper discusses the concept of sentencehood in Minimalist syntax. Although this problem has long been neglected in generative grammar, the intuition of sentencehood is deeply rooted in human innate syntax. We present several pieces of evidence confirming the sentence as a syntactically defined object, including Subject-Aux Inversion, Extended Projection Principle, and Empty Category Principle. We also offer a biolinguistic explanation why the sentencehood innately exists in human language.

Keywords: sentence, Subject–Aux Inversion (SAI), Extended Projection Principle (EPP), Empty Category Principle (ECP), syntactic typology

1 Introduction

If we define science as a truth-seeking activity, we consciously or unconsciously presuppose that humans often hold superfluous conceptions. Thus, science has two mutually interacting sides: one is *philosophical*, which simplifies scientific theories modeling some aspects of the world eliminating erroneous conceptions by logic and observation, as Descartes did in his *Meditation*; the other is *cognitive-scientific*, which tries to explain why humans seem to have unreasonable cognitive capacities and do not intuitively hold correct conceptions of the world, as Hume did in his *Treatise of Human Nature*.

The same holds for theoretical linguistics. At first glance, human language appears to contain superfluous functions such as grammaticalness, linear order, and projection. One of the most important desiderata of the Minimalist Program is to offer the broadest possible explanations to syntactic phenomena without appealing to biologically inadequate assumptions. Grammaticalness, linear order, and projection are all abandoned as superfluous conceptions. Minimalist Program can be regarded to employ a kind of philosophical thinking.

In many cases, the important consequences given by the Minimalist Program are at the same time cognitive-scientific. Grammaticalness, for example, is not just abandoned with the genesis of the Minimalist Program but is reduced to the legibility conditions holding at the interfaces. This explains why humans appear to have an unnecessary capacity for distinguishing grammatical sequences from ungrammatical ones. The proposal of interface conditions not only simplifies the syntactic theory as a whole, but also reveals the nature of unnecessary capacities of the human cognitive capacity. In fact, giving at the same time philosophical and cognitive-scientific explanations to a single phenomenon is the most natural way by which science proceeds.

Nevertheless, there are some concepts that are merely abandoned as superficial and have never been explained in the study of language. Kuroda (1969) was the first to notice this kind of problem. Through careful observations of Japanese, he argued that Universal Grammar is equipped with the concept of “subject,” which Frege and subsequent logicians regarded as unnecessary. This indicates that grammar is not logic. Once generative grammar is defined as the investigation of the extent to which extent the theory of grammar is close to the Strong Minimalist Thesis (SMT), we must begin to not only abandon superfluous concepts but also *explain* them.

*The present paper is a revised version of my MA thesis submitted to the University of Tokyo in 2022. I am grateful to Chizuru Nakao, Akira Watanabe, and two anonymous reviewers for helpful comments. This work was supported by Grant-in-Aid for JSPS Fellows (No. JP22J21631).

“Sentence” is another concept that traditional grammarians and logicians have long assumed but structuralists¹ and generative syntacticians do not regard as a formally defined concept of grammar. Surveying generative grammatical studies in the 2010’s, Chomsky et al. (2019) comment that—

- (1) “[T]he informal notion ‘sentence’ [is] now abandoned in favor of hierarchically structured objects.” (Chomsky et al. 2019:232)

The point of this statement is “informal.” Is the notion “sentence” truly independent of formal syntax? No. Rather, the concept of sentence heavily depends on syntax, and it should be treated in the field of Minimalist syntax, as we will see in this paper.

The remaining paper is organized as follows: section 2 surveys the history of research concerning the definition of sentence and, based on insightful work by Motoki Tokieda, argues that sentencehood is a kind of linguistic intuition just as grammaticalness; section 3 provides some pieces of evidence confirming the existence of sentencehood as a linguistic intuition dependent on syntax, suggesting some further consequences to Minimalist syntax; section 4 turns to the problem of sentence “beyond explanatory adequacy,” i.e., “why does sentence exist in human language?”; section 5 concludes the paper and suggests some implications for cognitive science.

2 “Sentence” as an arbitrary unit in human language

2.1 Sentencehood and phrase structure: why has the problem of sentencehood been neglected?

The goal of generative grammar is to understand the innate linguistic capacity of human beings. It consists of various modules such as the lexicon, phonology, and semantics. In natural sciences, complex objects are dealt through simplification. For instance, the falling body does not completely obey the law of free fall owing to air resistance. Physicists thus consider the ideal situation (i.e., vacuum) by eliminating the superfluous effect of air resistance, in order to unify physical phenomena in a systematic fashion. Similarly, theoretical linguists study pure linguistic knowledge by abstracting away from performative factors such as memory load, utterance, and speech recognition. Take an example of center-embedding in Japanese: “Taro-ga Jiro-ga Saburo-ga Shiro-ga hasi-tta-to omo-tta-to kangae-ta-to i-tta [Taro said that Jiro imagined that Saburo thought that Shiro ran.]” is unacceptable due to memory load, but it does not cause any grammatical problems. In fact, sentences with only one or two embeddings, such as “Taro thought that Jiro ran” and “Taro thought that Jiro imagined that Saburo ran,” are acceptable, strongly suggesting that grammar (i.e., the linguistic faculty in narrow sense) and working memory need to be considered separately. Along with memory load, etc., grammaticalness is just one factor that determines acceptability (Chomsky 1965).

The essence of the linguistic faculty lies in grammar, which produces an infinite range of linguistic sequences. Non-human animals are also equipped with linguistic capacity in the broad sense such as working memory and speech recognition. Parrots, for example, can use those mental organs to repeat utterances by humans. However, as Descartes points out in *Discours de la méthode*, parrots cannot produce an infinite number of sentences, which humans easily can, but only imitate them. Then, what makes the infinite production of sentences possible? It is the hierarchical arrangement of the lexical items that is essential to the infinity of language, since the number of the lexical items stored in the brain is finite. There is no limit of the length of a sentence thanks to embedding (“[Taro said [that Jiro said [that Saburo thought that ...]]]”), and core syntax produces infinitely many sentences. The syntactic object (SO) formed by syntactic computation is sent to the conceptual-intentional system and the sensorimotor system, mental organs adjacent to the core syntax. The former interprets the meaning, and the latter converts SOs into the phonetic forms we actually hear.

The simplest way to form a hierarchical structure is to take elements a and b and create a set {a, b}. The set-forming operation is called MERGE.

- (2) MERGE(a, b) = {a, b}

¹See Tokieda (1941) and the discussion in section 2.2. In fact, the notion of “sentence” is simply undefinable under the structuralist method.

Note that there is no ordering relationship between a and b. This means that MERGE does not determine the word order, but it only forms hierarchical relationships. The set formed by MERGE is called a phrase. For example, the phrase “Taro’s book” is generated as follows.

- (3) a. $\text{MERGE}(\text{Taro}, s') = \{\text{Taro}, s'\}$
- b. $\text{MERGE}(\{\text{Taro}, s'\}, \text{book}) = \{\{\text{Taro}, s'\}, \text{book}\}$

If MERGE is applied repeatedly, the phrase becomes increasingly longer.

How is the “sentence” defined in this framework? Interestingly, since Chomsky (1995), who first defined Merge, although in a way different from (2), the question “what is a sentence?” has seldom been discussed in the framework of generative grammar. This is not only because researchers were unaware of the existence of the problem, but also because the basic assumptions of generative grammar since the 1950’s unintentionally concealed it. In the LSLT-model (Chomsky 1975), the hierarchical structure is generated by a set of phrasal structure rules, a “rewriting system” between linear sequences such as the following:

- (4) $XYZ \rightarrow XWZ$

This reads “rewrite the element Y between X and Z as W.” X and Z need not be specified, in which case (4) is called a context-free phrase structure rule. If X or Z is specified, (4) is called a context-sensitive phrase structure rule. For example, the fact that a noun phrase (e.g., *the man*) consists of a determiner and a noun is described by the following (context-free) phrase structure rule:

- (5) $\text{NP} \rightarrow \text{Det N}$

The phrase structure rule can be applied repeatedly.

- (6) $\text{VP} \rightarrow \text{V NP}$

Rule (6) generates verb phrases with a transitive verb as a head. If (5) is applied after (6), a verb phrase such as *criticize the man* is generated. By repeatedly applying phrase structure rules, we can construct sentences of infinite length, as in the case of MERGE. Since phrase structure rules rewrite a (set of) symbol(s) into another, it is necessary to posit a starting symbol. In the 1950’s, the starting symbol was considered “sentence.”

- (7) $\text{S} \rightarrow \text{NP VP}$

For example, the sentence “the woman criticized the man” is generated by applying the phrase structure rules in the order of (7), (5), (6), (5). Any sequence generated by phrase structure rules with S as the starting symbol is, of course, a sentence. In other words, in the phrase structure grammar model, the question “what is a sentence?” cannot arise. As long as we assume an adequate set of phrase structure rules, no sequence that is not a sentence can be generated.

However, there is a fatal flaw in the phrase structure rules: they are too complex. Consider the rules for rewriting verb phrases in Japanese.

- (8) a. sono otoko-o hihan-suru
 the man-Acc criticize
 ‘criticize the man’

b. $VP \rightarrow NP V$

The relationship between V and NP in (8a) shows that the object/goal of an action is a certain noun phrase, just as in *criticize the man*. In the MERGE model, we can offer a deep explanation for the semantic relations common to “criticize the man” and “sono otoko-o hihan-suru,” independent of word order, since MERGE only forms unordered sets; in other words, the semantic relation between a verb and its object is formed by the merger of a verb and a noun phrase in any language, and the “superfluous” factor of word order, which varies from language to language, is abstracted away. Phrase structure rules, however, do not allow for such generalization, because they have to specify linear order, requiring specific rules to be formulated for English (6) and Japanese (8b), respectively. Moreover, under the phrase structure schema, infants acquiring a specific language would have to memorize a large number of phrase structure rules that are different for each individual language. Such an assumption is inconsistent with the fact that children acquire a language in a relatively short period of time with poor linguistic stimuli.

Thus, under the Minimalist Program, which removes as many devices as possible from core syntax, the MERGE model was introduced as the simplest operation to build hierarchical structures. Unlike phrasal structure rules, which construct hierarchical structures in a top-down manner, MERGE builds them up in a bottom-up fashion; phrase structure rules have a starting point (“sentence”), but MERGE does not have an end point and can be iterated without limit. Therefore, the notion of “sentence,” which is self-evident in the phrase structure rule model, is eliminated as an informal notion. In fact, in a “state-of-the-art” paper dealing with the 2010’s minimalist syntactic research, Chomsky et al. (2019) says “[t]he informal notion ‘sentence’ [is] now abandoned in favor of hierarchically structured objects,” (p. 232). Few studies exist on this topic under the Minimalist Program, and it is no exaggeration to say that the question of “what is a sentence?” is now completely forgotten.

2.2 Rereading Motoki Tokieda’s *Principles of Japanese Linguistics*: why does the sentence matter to linguistics?

It is worth noting that the situation surrounding the problem of the sentence is very similar to that in the early 20th century. Saussure says:

- (9) If we picture to ourselves in their totality the sentences that could be uttered, their most striking characteristic is that in no way do they resemble each other. (Saussure 1916:106)

Saussure concludes that “it[sentence] belongs to speaking[parole], not to language[langue]” (Saussure 1916:124). Hjelmslev sums up the controversy over the definition of sentences:

- (10) The sentence, by the way, does not seem to be a concept of linguistic order. It is very difficult to define a sentence. However, judging from the definitions given by H. Paul and Wundt in their examples, it seems that one must choose between the linguistic concept of a mere segmental series, and the much narrower concept of logical proposition, which is much closer to that of sentence but has nothing to do with linguistics. John Ries clearly demonstrated that morphological analyses give only three kinds of elements: phonemes, words, and series of words (*Wortgefüge*). He correctly adds: “but morphological analyses do not find sentences either. If we suppose that, by deeper observations, we can recognize morphological features of the series of words called sentences—the opposite cannot be denied a priori—then we will have a more certain clue to the more detailed division of a large number of series of words.” This is the current state of our research. (Hjelmslev 1928:25-26, translated by IN)

Clearly, it is very difficult to define a sentence with a formal method. The resignation and indifference to defining a sentence as a linguistic concept in the early 20th century is very similar to that of generative grammarians today.

However, it should not be overlooked that we humans can certainly discern sentences from phrases, even if we are not taught the formal definition of it. Native speakers of English intuitively know that (11a–c) are sentences and that (12a–c) are not.

- (11) a. John ate the cake.
 b. Did John eat the cake?
 c. Which cake did John eat?
- (12) a. that John ate the cake
 b. if John ate the cake
 c. which cake John ate

The contrast between (11) and (12) cannot be explained by merely eliminating the notion of sentence as formally unformulable. Rather, linguists should try to understand what lies behind the intuition of sentencehood.

Motoki Tokieda, one of the most famous traditional Japanese linguists known for his great originality, was likely the first to point out the importance of the concept of sentence for linguistics. Tokieda takes a positive view of the concept of sentence as a linguistic reality.

- (13) First, I would like to reject the attitude of pre-determining the limits of sentences and non-sentences or the reality of sentences by certain criteria. Rather, I would like to consider first what our intuition of the sentence as a unit is, and on what basis we think of the sentence. (Tokieda 1941:328, translated by IN)

Tokieda denies the structuralist view that treats utterance as if it were a natural object and classifies it into sentences and non-sentences according to external criteria. Rather, it is our internal language instinct that enables us to determine what kind of sequence is a sentence. Tokieda also says²:

- (14) When a word in language is called a unit, it is not in the sense of an atomic unit that is achieved as a result of analysis, but it must be in the sense of a qualitative unit. The qualitative unit is a concept defined by our subjective consciousness as a single totality and as a unified body. With this concept presupposed, we can divide a given sequence of phonemes into a combination of two or three words. The same holds for sentences, which can be recognized as one or two sentences just because there is a concept of a sentence as a predetermined unity. The problem is how the essence of a word or sentence as a unit subjectively defined can be explained academically. (Tokieda 1941:220)

The attempt to define sentences in terms of structural linguistics does not work because it treats language as if it were an object separate from the speaker's language instinct; rather, the concept of sentence can only be understood by considering the speaker's linguistic intuition. In the course of his thought, Tokieda pays attention to Yoshio Yamada's definition of sentence by the concept of predicates.

- (15) In fact, the main point distinguishing words and sentences is, as stated above, the presence or absence of the activity of the focus of consciousness. In other words, the reason why a word or a series of words can become a sentence is because of the power of thought that exists within it. (Yamada 1936:901, translated by IN)
- (16) There must always be a unifying effect in every thought. Call it apperception effect. The apperception is indeed the life of thought. The thought that is integrated by the apperception and expressed in the form of language is the sentence. (Yamada 1936:901)

²As is seen from the following quotation, wordhood is regarded to form a pair with sentencehood in Tokieda's theory. Although Tokieda does not offer a very interesting hypothesis on wordhood, it is important to recognize that the same kind of problem as sentencehood exists in treating words in generative grammar.

- (17) The apperception effect is a cover-term of the combination by consciousness, and it includes all thoughts such as explanation, imagination, questioning, command, prohibition, desire, and emotion. (Yamada 1936:917-918)

Yamada argues that the essence of a sentence is not merely to express a proposition, but to unify the situation with the subject. In this respect, his theory is close to Austin's speech act analysis (Austin 1962).

Tokieda further develops Yamada's theory and introduces new concepts of *ji* and *shi*. *Ji* is a word that expresses purely subjective attitudes, while *shi* expresses objective materials and concepts. Nouns, adjectives, and verbs are classified as *shis*, while particles and auxiliaries are classified as *jis*. The connection of *shi* and *ji* is the essence of the sentence. In Tokieda's words, *ji* "expresses a subjective unity by wrapping an object with a subjective thing."

More precisely, Tokieda summarizes the conditions that a sentence must meet in the following two points.

- (18) a. Unity
b. Completeness

Unity (18a) refers to the unification of objective *shis* by subjective *jis*, as mentioned above. *Completeness* (18b) is the condition that the phrase that unifies the sentence must be a thematically complete phrase. For example, "Hana-wa [flower-Top]" is not considered a sentence, even though it is an expression unified by a *ji*, because it is thematically incomplete.

In terms of modern generative grammar, the two conditions can be redefined as follows:

- (19) a. A modal element is in the highest position in the phrase
b. The phrase is complete in terms of θ -theory and predication.

The essence of the two conditions can be unified as follows:

- (20) *The Condition of Sentencehood:*

A thematically complete SO is recognized as a sentence if and only if a tense/modal feature is in its Minimal Search (MS) domain.

Here, hypothesis (20) seems adequate, and in fact, offers rich consequences to the theory of grammar, as we will see in what follows.

2.3 Study of the sentence in generative grammar

Before beginning the minimalist research of sentencehood, let us discuss some methodological preliminaries.

The first one is about the definition of a sentence: what do we exactly mean by the word "sentence"? Note that sentencehood is different from grammaticalness. Consider the following examples we saw in section 2.2:

- (21) a. John ate the cake.
b. Did John eat the cake?
c. Which cake did John eat? (= (11))

- (22) a. that John ate the cake
b. if John ate the cake
c. which cake John ate (= (12))

Here, (21) and (22) are all grammatical, but only sequences in (21) are sentences. Thus, we have some intuition distinguishing sentences from non-sentences, and the intuition is different from that which enables grammaticality judgment.

It is important to review how the notion of “grammar (I-language)” is introduced. In the 1950’s, with the careful observations of English, Chomsky confirmed the presence of a formal device in our brain that enables infinite generation of hierarchical structures called (generative) grammar. Grammar is considered to be the source of our judgment of grammaticality, with ungrammatical sequences deviating standard derivations or representations provided by a set of grammatical rules. The judgment of grammaticality is in turn the sole channel through which we can access our mental grammar.

Grammaticality, however, has never been clearly defined; it is formally undefinable. This is because the sole intuition we can directly access is that of acceptability; grammaticality is just one factor determining acceptability, and we do not have a formal method that precisely distinguishes grammaticality from other related intuitions. Take an example from Chomsky (1965):

- (23) a. I called up the man who wrote the book that you told me about.
- b. quite a few of the students who you met who come from New York are friends of mine.
- c. John, Bill, Tom, and several of their friends visited us last night.
- (24) a. I called the man who wrote the book that you told me about up.
- b. the man who the boy who the students recognized pointed out is a friend of mine. (Chomsky 1965:10-11)

According to Chomsky, the sentences in (24) are less acceptable than (23). As Miller and Chomsky (1963) and others show, these effects are due to the limit of memory capacity and caused by nested/self-embedded structures³. Important is that, without Miller and Chomsky (1963) and the related studies, these examples would not be considered as “unacceptable but not ungrammatical,” since we only have the intuition of acceptability; in other words, what is taken to be grammatical changes as the research in syntax and the related fields proceeds. Such holistic nature of research is common to scientific studies in general⁴.

Given this, what we can do now is to define “sentencehood” as what we intuitively judge to be a “sentence.” It might seem to be circular, but in fact no problem arises. As we have seen in this section, we do have some intuition of sentencehood in mental grammar, the exact nature of which remains to be explored. Just like grammaticality, the notion of sentencehood will be made more sophisticated as the study advances.

Note that what one takes to be a sentence differs from person to person, probably because of non-linguistic factors such as the individual’s degree of education. In fact, Tokieda judges a wider range of phrases to be a sentence than Yamada, whose intuition reflects that of mine; according to Tokieda, “kaji” (“fire”) and “tae-naru hue-no ne yo” (“oh, beautiful sound of flute”) are sentences when uttered in particular contexts. However, such vagueness of the concept is a peripheral phenomenon, and there is a class of phrases the sentencehood of which nobody denies, just as in the study of grammaticality. We should investigate the nature of such core phenomena before turning to the peripheral ones, the status of which will be precisely determined as a consequence of the theory of the “kernel.”⁵.

Another issue is the relationship between “sentence” and “proposition.” Actually, the notion of “proposition” is similar to that of “sentence,” but they should not be confused. “Proposition” is a concept determined by logic: it is a type of logical entity complete in terms of its thematic properties. “Sentence” is a mental object defined by our innate linguistic intuition. Thus, there exist propositions that are not sentences:

- (25) that John read the book.

³Though these contrasts might be explained in terms of some syntactic device like the A-over-A condition, I ignore such a possibility for expository purposes.

⁴Chomsky (2021) also mentions the relation between modern linguistics and the Duhem–Quine holism.

⁵For related issues, see section 4.2.

The *that*-clause is not a sentence, but it is thematically complete, assuming that *that* is a meaningless object (Lasnik and Saito 1984). Although there does not seem to be any “core” instance of “nonpropositional sentences,”⁶ it is due to what Tokieda (1941) calls (the condition of) “completeness,” which we saw in the previous subsection.

Lastly, let us briefly discuss how to deal with embedded sentences. As we saw in the last paragraph, the *that*-clause is not itself a sentence. However, the complement of *that* seems to be a sentence. The problem is whether we should treat “embedded sentences” to be actual sentences, belonging to the same type of phrases as matrix sentences.

Related here is the point we discussed in the previous subsection, that is, the modality is a crucial factor defining sentencehood. Hence, if an embedded sentence itself has the modal meaning independent of the matrix predicate, it can be regarded as a “true sentence.” In fact, there are examples indicating this:

- (26) a. I like (it) that Alice plays chess.
b. I like for Alice to play chess. (Ransom 1986:18)

As Ransom (1986) points out, the embedded clauses of (26a) and (26b) have the same propositional content but differ in what he calls “propositional modality,” that is, how the speaker conceives the propositional content. In (26a), the speaker presupposes the propositional content “Alice plays chess”; however, in (26b), the speaker does not expect the propositional content of the embedded clause to be true, and it simply acts as the object of the matrix verb.

Since the matrix verb is the same in (26a) and (26b), the embedded sentence in (26a) should by itself add something to the proposition, which the infinitival clause in (26b) does not, and it is the attitude of the speaker to take the propositional content to be true. Sentences in fact express varieties of attitudes in cooperation with modals, sentential adverbs, and the matrix predicate if embedded. Yet, as indicated by (26), sentences certainly have their own meanings, regardless of whether they are embedded. We will thus assume that embedded sentences (\neq embedded clauses/CP) are sentences, essentially the same type of objects as matrix sentences⁷.

3 Evidence for the sentence as a syntactically defined object

3.1 Subject–Aux Inversion

As we saw in the previous section, we define sentence as follows:

- (27) *The Condition of Sentencehood:*
An SO is recognized as a sentence if a tense/modal feature is in its MS domain. (= (20))

According to Chomsky (2021), tense is not an independent head but a feature of v . Hence, suppose that the locus of tense/modal-feature is v , where modals are base-generated.

Assume further that MS is defined with the proper subset relation:

- (28) a. MS finds a target in the optimal way via the shortest possible path, where the path of α is the set of all SOs of which α is a term.
b. The path of α is shorter than the path of β iff the path of α is a proper subset of that of β (Epstein et al. 2020:4)

For example, given $SO = [\alpha [\beta a b] [\gamma [\delta c d] e]]$, the paths of each term are defined as below:

- (29) a. $\{\alpha, \beta\}$

⁶But, see the discussion in section 4.2.

⁷Of course, as Hooper and Thompson (1973) show, certain types of embedded sentences display some properties different from matrix sentences, but it is not the effect of the (non-)sentencehood itself.

- b. $\{\alpha, \beta\}$
- c. $\{\alpha, \gamma, \delta\}$
- d. $\{\alpha, \gamma, \delta\}$
- e. $\{\alpha, \gamma\}$

a, b, and e can be detected via MS, but c and d cannot: the path of e is a proper subset of the paths of c and d. The paths of a, b, and e are equivalent in length, since there is no proper subset relation between them.

In this section, we will comb through several pieces of evidence for the sentencehood hypothesis and consequences thereof.

3.1.1 Sentencehood and Subject–Aux Inversion

Subject–Aux Inversion strongly suggests that (20) holds cross-linguistically, which means sentencehood depends on syntax. As far as I know, under the current Minimalist schema, no one has explained why SAI occurs in matrix questions but not in embedded questions⁸. Sentencehood condition (20) forces SAI in matrix questions. Assume that an interrogative must be headed by Q with a thematically complete complement. Suppose also that Q-feature can appear not only by itself (as a complementizer) but also as a part of v . Then, $v^{(*)}_{[T]}$ must move to the highest position of an interrogative sentence with a thematically complete SO its complement. (As we will see in section 3.2, what (A-)moves to SpecINFL is not NP, but vP . $v^{(*)}_{[T]}$ raises above Q from inside the upper $v^{(*)}P$.)

- (30) a. $[[_{vP} \dots v_{[T,Q]} \dots] [INFL \text{ } \overline{v_{[T,Q]} \dots}]]$ (sentencehood condition satisfied, but the v -head with Q does not have a thematically complete complement) \rightarrow
- b. $[v_{[T,Q]} [[_{vP} \dots v_{[T,Q]} \dots] [INFL \text{ } \overline{v_{[T,Q]} \dots}]]]$ (satisfies all conditions, being an interrogative thanks to $v_{[T,Q]}$ with a thematically complete complement)

In embedded questions with Q-complementizer *if*, $v^{(*)}_{[T]}$ does not raise above *if*; otherwise, *if* would not in the highest position of the embedded clause. It could not be selected by the matrix verb and hence no longer regarded as a question, regardless of whether the raised v has Q-feature:

- (31) a. I wonder if John read the book.
- b. $[I \text{ wonder } [if_{[Q]} [John \text{ read the book}]]]$

The matrix verb *wonder* selects the interrogative CP headed by *if*_[Q], which need not be a sentence.

The hypothesis also explains the difference of meanings between matrix and embedded questions. Holmberg (2016) points out that matrix questions have the effect of requiring the hearer's answer, which embedded questions do not have. Although Holmberg tries to deduce it by assuming what he calls Q-force operator above TP, it is simpler to regard it as the effect of the tense/modal feature on v (*read* in (31))⁹.

The same analysis holds for WH-questions. In direct WH-questions, SAI must take place, since v with tense/modal features has to be on one of the highest positions in the whole SO owing to the sentencehood condition. As can be seen from the case of indirect WH-questions, the WH-phrase requires a Q-feature on the top of its sister phrase:

- (32) a. $[[What] [did_{[T,Q]} [John \text{ read}]]]?$

⁸Chomsky (2015) offers an analysis for the asymmetry between the subject NP and T', although it is not an explanation for the driving-force of SAI itself; when SAI takes place, what is inverted is always T, but not N. He tries to resolve this problem by assuming that NP moves to the subject position after SAI takes place. However, such a derivation is obviously counter-cyclic, unmaintainable under the simplest assumption of Minimalism. Moreover, as Carstens et al. (2016) points out, his analysis causes the same problem in $v^{(*)}P$. These kinds of problems do not occur under the sentencehood theory.

⁹For related discussion, see section 4.2

- b. I wonder [[what] [Q [John read_[T]]]].

SAI can occur even in embedded clauses as observed in Belfast English (Henry 1995), since it still satisfies selectional requirements for Q imposed by the matrix verb. However, Standard English does not employ verbs that select phrases with Tense- and Q-features, and hence it does not exhibit SAI in embedded clauses. In Belfast English, on the other hand, verbs have selectional features for Tense and Q, and SAI in embedded clauses is grammatical.

Why direct WH-questions require SAI has long been a mystery; WH-fronting appears to be enough to mark a question, and thus SAI in direct WH-questions seems redundant in terms of learnability. However, as we saw above, the sentencehood condition offers an elegant solution to the problem.

3.1.2 Consequences for head movement in general

The traditional GB theory regards head movement as an adjunction operation by a head, which apparently violates cyclicity condition (Chomsky 2001, 2019). Chomsky (2001) thus proposes that head movement is a PF-operation. However, the PF-movement hypothesis fails to explain the performative difference between direct and indirect questions Holmberg (2016) observes; if SAI were a PF-operation, direct questions would have the same interpretation as indirect questions.

Roberts (2010) gives another example where SAI seems to have a semantic effect.

- (33) a. * Which one of them does anybody like?
 b. Which one of them doesn't anybody like?
 c. * They succeeded in finding out which one of them anybody liked.
 d. * They succeeded in finding out which one of them anybody didn't like.
 e. They succeeded in finding out which one of them wasn't liked by anybody. (Roberts 2010:10)

In (33b), *doesn't* licenses the negative polarity item *anybody* via c-command, in contrast to (33a). This could not be explained if head movement were a purely PF-phenomenon, which should not affect interpretations. These examples sufficiently confirm that at least some cases of head movement take place in narrow syntax.

Then, how should we define syntactic head movement without causing the cyclicity problem? Fukui and Narita (2017) try to avoid the problem by assuming an SO distinct from the ordinary one, which represents head movement:

- (34) a. [[T C] [t_T vP]]
 b. SO1 (SO for cyclic merger): {C, {T, vP}}
 c. SO2 (SO for head movement): {T, C}

Fukui and Narita argue that the seemingly counter-cyclic structure with head movement (34a) in fact consists of two SOs (34b) and (34c). (34b) is an ordinary phrase structure without head-adjunction. (34c) represents the merger of head-level elements, indicating T and C constitute a single head.

Yet, Fukui and Narita's characterization still violates the condition of Resource Restriction (Chomsky 2019), a sophisticated version of cyclicity which is necessary to minimize syntax by eliminating countercyclic operations such as Parallel Merge, Sideward Movement, and Late Merge.

(35) Resource Restriction:

Each application of MERGE cannot generate more than one additional accessible element in a workspace.

For instance, given $WS = [\{a, b\}]$, there are three accessible elements, that is, a , b , and $\{a, b\}$. If b (internally) merges with $\{a, b\}$, the workspace is transformed to $WS' = [\{b, \{a, b\}\}]$. WS' has four accessible elements, that is, a , (the upper copy of)

b, {a, b}, and {b, {a, b}}. The lower b is not in the MS domain of WS because the upper b intervenes. The application of MERGE transforming WS into WS' thus satisfies the condition of Resource Restriction.

Fukui and Narita's solution is problematic according to the Resource Restriction. When head movement applies, WS = [{C, {T, vP}}] is transformed into WS' = [{C, {T, vP}}, {T, C}]. Putting aside the internal structure of vP, WS has five accessible elements (vP, T, C, {T, vP}, {C, {T, vP}}), but WS' has eight (vP, T, C, {T, vP}, {C, {T, vP}}, T, C, {T, C}), because T and C in SO2 are in the MS domain as well as T and C in SO1, and hence visible to the calculation of the Resource Restriction.

The most plausible way to characterize head movement is reducing it to ordinary (Internal) MERGE operation (Matushansky 2006). In SAI, for instance, *v* with tense/modal features internally merges above IP/⟨*ϕ*, *ϕ*⟩P:

(36) [*v* [_{IP} ... *v* ...]]

In reality, head-level MERGE is a natural consequence of MERGE-only hypothesis; since both lexical items and phrases are SOs, there is no a priori reason to ban head-level merger. Roberts (2010) correctly points out that head movement cannot be distinguished from phrasal movement in Bare Phrase Structure schema (Chomsky 1994), though he still assumes head movement by adjunction operation.

One might maintain head movement by adjunction to explain the well-known complementary distribution of inverted auxiliaries and complementizers:

- (37) a. * if did John read the book.
b. * did if John read the book.
c. * that did John read the book.
d. * did that John read the book.

If head-movement only occurs via head-level adjunction, these facts are easily explained, since *if*, *did*, and *that* cannot constitute a head-level amalgamate due to morphological reasons. However, the above contrasts need not be attributed to narrow syntax. It is more natural to derive them from semantic properties of Q-feature and complementizer. Hamblin (1958) argues that an interrogative is a set of propositions:

- (38) a. if John read the book
b. {John read the book, ¬John read the book}

The complementizer *if* and the Q-feature on *did* can be regarded as a function mapping a proposition to a set of propositions. They cannot take a set of propositions as an argument, which explains why (37a) and (37b) are ungrammatical. Similarly, the ill-formedness of (37c) and (37d) is due to the selectional features of *that* and Q, which only take a proposition, not a set of propositions or a (that-)clause, as its complement.

Once syntactic head movement is reduced to (Internal) MERGE as its special case targeting a head, we need not assume a "target position" for each application of head movement; there is no reason to suppose SAI to be a movement operation targeting C; it is just a case of merger of *v*_[T] and IP (in traditional terminology) containing *v*_[T] inside. Similarly, *v*-raising does not take INFL as its target position. As we will see in section 3.2, it is sometimes remnant movement of vP, and sometimes merger of *v* above IP.

More generally, we can abandon "S-selection by functional categories," which Chomsky (2000) introduces to explain the hierarchical order of functional categories (C-T-v). The C-T-v order is undefinable under the sentencehood theory and is in fact superfluous¹⁰:

¹⁰However, we might admit that C-T/INFL's adjacency is required by feature inheritance.

(39) [$v_{[T]}$ [$_{IP}$... $v_{[T]}$...]]

In this example, there are at least two vP forming a chain, in traditional terms. The structure causes no problem, considering the duality of semantics: v 's T-feature does not play any role in the lower copy, and it works only in the highest position in a phrase; v as a θ -assigner, on the other hand, works only in the lowest position but not in the highest position. It is a well-established fact that every SO cannot play one semantic role in more than one syntactic position, and syntactic v -raising is no exception.

The Head Movement Constraint (Travis 1984) is also abandoned: if MERGE is not driven by syntactic factors such as Attract by syntactic features, which is the simplest assumption, then no constraint should be imposed to head movement. In fact, the literature reports many cases wherein head movement does not obey Head Movement Constraint (Rivero 1991, 1994, Lema and Rivero 1991, 1992, Borsley et al. 1996, Roberts 1994, 2010).

(40) *Breton*:

- a. Lenn a ra Anna al levr
 read-INF.PRT does Anna the book
 'Anna reads the book'
- b. Lennet en deus Anna al levr
 read-PPRT has Anna the book
 'Anna has read the book'

Note that not all cases of head movement are reducible to MERGE. It is true that all syntactic head movements are a subcase of (Internal) MERGE, but PF-head movement may also exist. Cliticization and incorporation are representative cases:

(41) *French*:

Je l_i 'aime t_i
 I him/her-Cl-love

It is conceivable that both cliticization and incorporation move an element to a position already filled by another morpheme. They are thus irreducible to MERGE. In fact, there are two types of head movement, as Roberts (1994) suggests: one is syntactic MERGE, including non-local head movement, and the other is PF-merger obeying head movement constraint, reduced to the adjacency condition in Externalization. Cliticization and incorporation thus belong to the latter type.

3.1.3 Question-formation without SAI

SAI is a straightforward support for the existence of sentencehood defined in the sense of this paper as a linguistic intuition dependent on syntax. However, in many other languages, other strategies are employed to form interrogatives. In Japanese, for instance, a question is formed by adding a question particle at the end of sentence:

- (42) a. Taro-wa hon-o yon-da
 Taro-Top book-Acc read-Perf
 'Taro read a book.'
- b. Taro-wa hon-o yon-da no?
 Taro-Top book-Acc read-Perf Q
 'Did Taro read a book?'

If (42b) is analyzed as (43) assuming *no* to be a Q-complementizer, it apparently violates sentencehood condition:

(43) [[Taro-wa [hon-o yon-da]] no]

However, (42b) need not be analyzed as a CP. Rather, the question particle *no* is attached to the main verb:

(44) Taro-wa [hon-o yon-da-no]

This means that *no* is not a [+Q] complementizer working as a head separate from the matrix verb, but that it phonologically realizes the [+Q] feature of the matrix verb. In fact, as Fukui and Speas (1986) propose, there does not seem to be any functional “head” in Japanese. The nature of agglutinative languages is that their formal features are represented by (bound) morphemes, rather than by (relative) positions, which is the nature of fusional and isolating languages (English does not have a special [+Q] morpheme).

3.1.4 CP-level movement without SAI

It is well known that A'-movement without SAI is possible:

- (45) a. John read this book.
b. This book, John read.

Rizzi (1997) analyzes Topicalization as A'-movement of the topic to SpecTop.

(46) [This book [Top [John read ~~this book~~]]]

However, there is no overt topic marker in English, and hence its status as a head is dubious¹¹. Rather, as Narita and Fukui (2022) claim, the topic-comment relation is enabled by Generalized Predication holding at SEM-mapping:

- (47) *Generalized Predication at SEM-mapping:*
For each SO of the form $\{\alpha, \beta\}$, with α and β both selectionally complete and α D-linked (given in the discourse), $\{\alpha, \beta\}$ can be optionally mapped to CI instructions to the effect that β is the comment (rheme) about the topic (theme) α . (Narita and Fukui 2022:211)

Under Narita and Fukui's hypothesis, the English topic construction is analyzed without assuming superfluous functional heads¹².

(48) $[_\Sigma [This\ book_i] [_{TP} [John] [_{T'} read\ t_i]]]$ (Narita and Fukui 2022:216)

This configuration yields two interpretations, according to Generalized Predication (47) (SEM(Σ) indicates the semantic value of Σ at SEM):

- (49) a. SEM(TP) ($\approx T'_k$ is about John_{*j*}):
The proposition that k is an event in the past in which some entity j read some entity i is about j , which is John (where the index i is provided by the copy/trace t_i of *this book_{*i*}*).
b. SEM(TP) ($\approx John_j$ is about T'_k):
The proposition that j is John is about k , which is an event in the past in which j read some entity i (where the index i is provided by the copy/trace t_i of *this book_{*i*}*). (Narita and Fukui 2022:216)

¹¹For detailed criticism of the cartographic approach, see chapter 5 of Narita and Fukui (2022)

¹²See also Lacerda (2020) for a similar “contextual” approach to the information structure.

If (47) is applied to the set $\{\{\text{This book}\}, \text{TP}\}$, we get (where — is filled either by (49a) or (49b)):

(50) SEM(Σ):

The proposition that — is about i , which is the entity denoted by *this book_i*. (Narita and Fukui 2022:216)

This is what we mean by “topic-comment” relation. No additional category is necessary for topic construction. As there is no head c -commanding $v_{[T]}$ in (48), SAI is not needed to meet the sentencehood condition¹³.

Another possible counterexample to the analysis given to SAI in English would be asymmetry in question formation in Italian, where SAI need not occur when *perché* “why” or *come mai* “how come” is WH-fronted.

- (51) a. I tuoi amici hanno già fatto il lavoro
‘Your friends have already done the work’
b. I tuoi amici già hanno fatto il lavoro
‘Your friends already have done the work’
- (52) a. Che cosa hanno già fatto?
‘What (they)-have already done?’
b. *Che cosa già fatto?
‘What already (they)-have done?’
- (53) a. Voi siete già andati a Milano
‘You have already gone to Milan’
b. Voi già siete andati a Milano
‘You already have gone to Milan’
- (54) a. Dove siete già andati?
‘Where (you)-have already gone?’
b. *Dove già siete andati?
‘Where already (you)-have gone?’
- (55) a. Perché Gianni è venuto?
‘Why Gianni has left?’
b. Come mai Gianni è partito?
‘How come Gianni has left?’
- (56) a. Perché (i tuoi amici) già hanno finito il lavoro?
‘Why (your friends) already have finished the work?’
b. Come mai (voi) già siete tornati a Milano?
‘How come (you) already have come-back to Milan?’

These cases may be explained if we suppose a difference in word order between Italian and English: Italian *come mai* and *perché* left-adjoins to a propositional element (sentence), whereas English *why* right-adjoins to it. Given that *perché* and *come mai* are IP-adjuncts, it is not unreasonable to suppose that they are in the clause-initial position *in situ*¹⁴:

- (57) a. [WH Sentence] (WH = *perché*, *come mai*, *how come*)
b. * [Sentence WH] (WH = *why*)

¹³(48) does not cause labeling problems under the vP -raising analysis we will later develop. See the discussion in section (3.2).

¹⁴Though Italian does not allow WH-in-situ in general (Rizzi 1982:51), it may be due to some phonological factors that require every WH-phrase to be on the clause-initial position. For prosodic accounts for WH-movement, see Richards (2010).

- c. [WH [$v(*)_{[T]}$] [Sentence ~~WH~~]] (WH = *why*)

(57a), where WH-phrases left-adjoin to the sentence, does not violate the phonological condition of WH-fronting. In (57b), on the other hand, *why* right-adjoins to the sentence and violates the condition. Hence, *why* must raise to “SpecC” position after the raising of $v(*)$, as in (57c).

Other WH-phrases are generally not base-generated in the clause-initial position. They must move to a position higher than the one where *perché* and *come mai* are generated. As movement to an adjunct position is generally prohibited, the merger of a lexical item which makes “distinct” SpecC position (SAI) is necessary. English *why*, on the other hand, right-adjoins to a propositional element. If it stays in situ as in (57b), it violates the phonological condition which requires a WH-phrase to be linearly the first element. Thus, *why* must raise to a distinct position. $v(*)$ with T-feature is merged, and then *why* raises to the position distinguished by $v(*)_{[T]}$. *How come*, which does not require SAI, would left-adjoin to a sentence, like Italian *perché* and *come mai*. For a similar approach, see Collins (1991).

3.1.5 Some seemingly exceptional cases

There are some seemingly exceptional cases which would pose a problem to the theory we have developed so far:

- (58) *Paduan*:

che roba che l magna!
what stuff that he eats

‘The things he eats!’ (Zanuttini and Portner 2003:49)

This configuration does not meet the sentencehood condition. One way to solve the problem would be analyzing (58) as a nonsentential, that is, an NP with a relative clause (*che l magna*). Then, (58) has essentially the same structure as (59):

- (59) What an amazing car!

Since these phrases are nonsententials, they need not satisfy the sentencehood condition¹⁵.

In fact, there is a type of exclamatives that should be analyzed as accompanying a relative clause (concealed exclamations):

- (60) It’s amazing the big car he bought. (Grimshaw 1979:298)

This sentence is semantically equivalent to the one with an overt WH-phrase.

- (61) It’s amazing what a big car he bought. (Grimshaw 1979:298)

In (60), a WH-phrase carrying E(xclamative)-feature need not occur, since the predication property of the matrix predicate indicate the c-commanding phrase to be an exclamative (see also Grimshaw (1979)). As a relative clause is an adjunct of an NP, it might be omitted:

- (62) It’s amazing the difference. (Michaelis and Lambrecht 1996:215)

The same analysis holds for some problematic cases of a question:

¹⁵Although (59) might be regarded as a sentence under Tokieda’s theory, our hypothesis of sentencehood (20) does not consider a thematically incomplete phrase to be a sentence.

(63) *Québec French*:

- a. Qui que [tu as vu t]?
Who *que* you have seen
- b. Qui qui [t est venu]?
Who *qui* has come (Rizzi and Shlonsky 2007)

These examples do not seem to observe the sentencehood condition. However, questions display similar semantic properties to exclamatives (concealed question):

(64) John refused to tell the police the fellows who had been involved. (Baker 1968)

This semantically parallels with the following:

(65) John refused to tell the police which fellows had been involved. (Baker 1968)

Considering the parallelism, (63) can be analyzed to have the following structure:

(66) [NP [NP Qui] [RC_(adjunct) que tu as vu t]]

Québec French thus employs a completely different strategy for question formation, though it looks similar to English matrix questions when externalized.

Another way to treat the exceptional cases would be to suppose that there are parametric differences in peripheral cases of sentencehood, from person to person or from language to language. In fact, as we saw in section 2.3, the notion of sentence seems to vary owing to syntactic or non-syntactic reasons. If so, (58) is similar to interjections, which some people (or perhaps languages) treat as a sentence. Such parametrization may be affected by the fact that undoubtedly nonsentential exclamatives such as (59) exist. Nevertheless, I would like to emphasize that the “core” condition of sentencehood is universal, despite some variations in peripheral phenomena.

3.2 The Extended Projection Principle (EPP)

3.2.1 The prevailing view on EPP: Labeling Algorithm

EPP is one of the most mysterious phenomena in syntax. Since its first definition in Chomsky (1981) as Principle P, it has long been assumed, and not explained (with an exceptional attempt by Rothstein (1983), which is nevertheless unsuccessful). Bever (2009) even claims that EPP is not a property of UG but results from general cognitive mechanisms. However, Bever’s proposal is no more than to give up seeking a true explanation.

Chomsky (2013) abandons the pessimistic view and tries to explain the phenomenon in terms of Labeling Algorithm (LA), a part of Transfer. LA is driven by Minimal Search (MS), a third factor mechanism¹⁶ looking for the highest and most prominent feature in the Spelled-out domain.

(67) Labeling Algorithm:

- a. if H is an LI, {H, XP} is labeled H.
- b. if YP is a lower copy of a dislocated element, {XP, YP} is labeled XP.
- c. if X and Y agree in terms of F, {XP, YP} is labeled F.

¹⁶Whether Minimal Search is an operation is an important question. Narita et al. (2017) define MS (0-Search in their terminology) as an operation. Kitahara (2020), on the other hand, suggests that it is a structural relation defined on structural descriptions. The latter view seems to me to be simpler, and thus preferable.

(67a) seems to be unproblematic. (67b) is required by the principle of Dynamic Antisymmetry (Moro 2000), motivated by linearization. If Internal Merge and External Merge are not distinguished (Chomsky 2008), (67c) is indispensable. However, note that all those are built on the assumption that every syntactic object (SO) must be labeled, which is never safe (Fukui and Narita 2017).

Under the LA approach, EPP is explained as below:

- (68) John ate the cake. (John=E(xternal)A(rgument), eat=V, the cake=I(nternal)A(rgument))
- a. $[_\alpha \text{ EA } v^* [_{\text{VP}} \text{ V IA}]]$ (v^* P-level; Transfer of VP)
 - b. $[_\beta \text{ EA T } [_\alpha \text{ EA } v^* [_{\text{VP}} \text{ V IA}]]]$ ($\langle \phi, \phi \rangle$ P-level)
 - c. $[_{\langle \phi, \phi \rangle \text{P}} \text{ EA T } [_{v^* \text{P}} \text{ EA } v^* [_{\text{VP}} \text{ V IA}]]]$ (Transfer of the whole clause)

EA in its base position must be dislocated; otherwise, it would cause labeling conflict with its sister $[v^* [_{\text{VP}} \text{ V IA}]]$. The highest copy of EA must be in SpecT, since it is the only position where EA can be accommodated by (68c); T and EA agree in terms of ϕ -feature, and the highest copy of EA and its sister together form a $\langle \phi, \phi \rangle$ P.

3.2.2 Problems of Labeling Algorithm

At first glance, Chomsky's argument is without any drawbacks. However, there are both theoretical and empirical problems.

Narita and Fukui (2022) point out that Labeling Algorithm actually contains six assumptions unintegratable to a single conception of minimal search:

- (69)
- a. $\{H, \text{XP}\}$, H an LI, is labeled H
 - b. Traces (lower copies) of IM are invisible to LA
 - c. $\{\{X_{[F]}, \alpha\}, \{Y_{[F]}, \beta\}\}$ is labeled $\langle F, F \rangle$
 - d. Weakness/invisibility of R
 - e. Strengthening of R and invisibilization of v by pair-Merge of v to R
 - f. Weakness/invisibility of T [parametrized] (Narita and Fukui 2022:143, slightly modified)

(69a) is the basic definition of minimal search, which seems reasonable. However, the other assumptions are not only irreducible to (69a), but also unreasonable to assume. (69b) is incompatible with the copy theory of movement and (simplest) MERGE, which make no distinction between internal and external Merge; (69c) requires X and Y not only to have a common feature F but also to agree in terms of F. However, the AGREE operation violates No Tampering Condition, adding some information to an already formed SO. (69d) might be interpreted as the nature of roots, but is incompatible with (69e), making R searchable. (69f) also needs further explanation for why minimal search “skips” morphologically poor T.

As Narita and Fukui argue, all these additional assumptions are necessary for Chomsky to maintain the hypothesis that all SOs must be labeled. Therefore, once the labeling requirement is abandoned without losing important empirical results of Labeling Algorithm (*e.g.*, EPP), minimalist syntax will make a great advance toward the Strong Minimalist Thesis. In fact, it is possible to gain a broader explanation of EPP without appealing to the labeling requirement, as we will see below.

Let us turn to empirical problems. The most serious empirical problem with the Labeling Algorithm is that it cannot explain the existence of $\{\text{XP}, \text{YP}\}$ structures where X and Y do not agree. For example, in (70), *there* does not seem to agree with T.

- (70)
- a. There is a man in the room.
 - b. There are three men in the room.

LA predicts not only (70b), but also (70a) is ungrammatical, since (67c) requires agreement (\neq sameness in value) between T and the nominal in SpecT.

More problematic case is locative-/predicate-inversion. Contrary to *there*-construction, they cannot be explained even if we assume EPP to be a requirement of D-feature in SpecT, as Chomsky (1995) does.

- (71) a. Into the room walked John. (Rochemont and Culicover 1990:69)
 b. Sitting in front of the camera smiling was Bill. (Rochemont and Culicover 1990:79)

One might argue that there is an invisible nominal in SpecT, and that locative PP is adjoined to it. Such an argument fails to predict the agreement pattern in locative inversion in (72); T agrees with post-verbal NP, not preverbal invisible NP, just like in *there*-construction.

- (72) a. In the swamp {was/*were} found a child.
 b. In the swamp {*was/were} found two children. (Bresnan 1994:95)

Moreover, it cannot handle predicate inversion, where the preposed phrase is a predicate, not an adjunct.

- (73) a. Voting in favor were three women. (Bresnan 1994:75)
 b. Coiled on the floor lay a one-hundred-and-fifty-foot length of braided nylon climbing rope three-eighths of an inch thick. (Birner 1992:58)

3.2.3 $v(*)P$ -raising analysis

Considering three problems outlined thus, $v(P)$ -raising is the best way to explain the EPP effect. Under this hypothesis, the derivation of unergative sentences goes as follows.

- (74) a. [John v^* [_{VP} eat the cake]] (Transfer of VP)
 b. [INFL [_{VP} eat the cake] [John v^* [~~_{VP} eat the cake~~]]] (VP-movement out of v^*P)
 c. [[_{v^*P} John v^* [~~_{VP} eat the cake~~]] INFL [_{VP} eat the cake] [~~_{v^*P} John [~~eat the cake~~]]] (v^*P -movement to SpecINFL and Transfer of the whole sentence)~~

VP moves out of $v(*)P$ before $v(*)P$ -raising owing to some morpho-phonological conditions which require $v(*)_{[T]}$, V, and/or INFL to be linearly adjacent. Since *there*, locative PPs, and predicates are all in vP , vP -raising can explain residual EPP phenomena.

There are still more reasons for believing in vP -raising analysis. Rochemont and Culicover (1990) give several pieces of evidence for vP -raising analysis of locative inversion. For example, in (75), adjectives predicating semantic subjects (*John* in (75a), *Bill* in (75b)) in vP are fronted together with PP.

- (75) a. Into the room nude walked John.
 b. In front of her smiling stood Bill. (Rochemont and Culicover 1990:71)

The secondary predicates *nude* and *smiling* are fronted together with PP, though they are not included in them. Therefore, locative inversion is not PP-fronting but actually vP -raising.

vP -raising to SpecT is a commonplace phenomenon in languages other than English. Based on data from Hindi, Mahajan (2003) argues that vP overtly or covertly moves to SpecT in all languages¹⁷. Müller (2004) makes a similar claim in German.

¹⁷Mahajan suggests that English only employs covert vP -movement, contrary to my account here.

Since their motivation is reducing head movement to remnant phrasal movement, they do not apply their analysis to the account of EPP. Unsurprisingly, English may display the same property, with the difference in positions where subject NP is pronounced.

Note that, though traditionally considered as impossible, “projection” of the moved element does not cause any problem in the current theory of grammar. Internal and External MERGE are in fact the same operation and are not distinguished at least for Minimal Search¹⁸. Once the informal notion “projection” is reduced to Minimal Search, it is similar to displacement, except that the latter includes MERGE operation (Narita et al. 2017). As the movement of a moved phrase is generally possible, Minimal Search targeting a moved element causes no problem. For similar arguments, see Hornstein and Uriagereka (2002) and Donati (2006).

3.2.4 EPP-fulfillment by head movement

Alexiadou and Anagnostopoulou (1998), surveying parametric variation in EPP-effect in several Indo-European and Semitic languages, show that DP-movement to SpecT and *v*-to-T movement exhibit complementary distribution. How should such “extended EPP-effect” be explained?

As we saw in section 3.1.2, head movement is reducible to ordinary MERGE operation. It means that verb-raising is reanalyzed as a subcase of *v*P-raising. Take an example of Greek:

- (76) eftase ena pedi /o Jorgos /kathe filos mu.
 arrived a child George every friend of-mine
 “A child/George/every friend of mine arrived” (Alexiadou and Anagnostopoulou 1998:512)

According to Alexiadou and Anagnostopoulou, there is no null-expletive in (76), because (76) does not exhibit the definiteness effect as sentences with an expletive do. Alexiadou and Anagnostopoulou claim (76) is a case of verb raising to T, but it is also conceivable that *v* with T-feature raises above INFL:

- (77) a. [*v* [INFL [*v* NP]]]
 b. [eftase [INFL [eftase ena pedi/o Jorgos/kathe filos mu]]]

For now, assume *v*P-raising with an expletive to be the cause of definiteness effect, although the exact mechanism remains unknown. Importantly, under the sentencehood theory, *v*-raising languages can be analyzed in parallel with *v*P-raising languages, the difference of which is reduced to the lexical idiosyncrasy requiring/prohibiting pied-piping of *v*(*)P.

Though traditionally analyzed as a case of head movement by adjunction, English *be/have*-raising is not *v*-to-T/INFL raising; otherwise, there would not be any adjunction site for adverbs (which cannot adjoin to “bar-level” SOs) between *be/have* and the subject in English, and hence (78) would be ungrammatical:

- (78) The fact that John probably has made several mistakes is well-known. (Kayne 1989)

However, it is also possible that *be/have* moves just above NegP (hence below INFL) by Internal Merge and undergoes PF-merger with INFL¹⁹:

- (79) a. [[_{*v**P} John *t_i* *t_k*]_{*j*} [INFL [probably [has_{*i*} [VP_{*k*} *t_j*]]]]] →

¹⁸It is true that Internal MERGE is more “economical” than External MERGE because the former picks out an element in Workspace, a domain close to the SO, while External MERGE must access the lexicon, which is more difficult to search (Chomsky 2019). However, such distinction is only related to the problem of locality in movement and does not matter to the problems of projection.

¹⁹Kayne (1989) was the first pointing out this analysis, though he assumes head-adjunction, contrary to the analysis here.

- b. $[[[_{v*P} \text{ John } t_i t_k]_j [\text{INFL} [\text{probably} [\text{INFL}+\text{has}_i [\text{VP}_k t_j]]]]]$

The “landing site” of the PF-merger is just below *probably*²⁰. Note that the notion “landing site” requires hierarchical structure to remain in Externalization. Such an extra level is necessary, or cliticization and incorporation, two representative cases of head movement that cannot be reduced to Internal Merge, would violate No Tampering Condition.

Observe the French counterpart of (78):

- (80) * Le fait que Jean probablement ait fait plusieurs erreurs est bien connu. (Kayne 1989)

Although it is commonly assumed that the landing site of PF-merger between a free and a bound morpheme is generally where the free morpheme is, there is no a priori reason thereof. In this case, the landing site of PF-merger of the verb and INFL is INFL, contrary to English.

- (81) $[[[_{v*P} \text{ Jean } t_i t_k]_j [\text{INFL} [\text{probablement} [\text{ait}_i [\text{VP}_k t_j]]]]] \rightarrow$

- (82) $[[[_{v*P} \text{ Jean } t_i t_k]_j [\text{INFL}+\text{ait}_i [\text{probablement} [\text{ait}_i [\text{VP}_k t_j]]]]]$

Thus, (80) is ungrammatical, since there is no adjunction site for *probablement* between SpecINFL and the Spec of the upper *v*.

3.2.5 Sentencehood as the driving force of *vP*-raising

Then, what drives *vP*-raising? It is the sentencehood condition. The intuition of sentencehood cannot be attributed to semantics or phonology, but to syntax, the core system of the Faculty of Language.

- (83) *The Condition of Sentencehood:*

An SO is recognized as a sentence if a tense/modal feature is in its Minimal Search domain. (= (20))

Recall also that the structural “depth” of an SO is defined in terms of proper subset relations between the paths of SOs:

- (84) a. MS finds a target in the optimal way via the shortest possible path, where the path of α is the set of all SOs of which α is a term.
b. The path of α is shorter than the path of β iff the path of α is a proper subset of that of β (Epstein et al. 2020:4, =(28))

$v(*)P$ itself is not a legitimate sentence: it does not satisfy the Case requirement. The INFL element inherited from C is necessary, and INFL is merged to $v(*)P$. Note that INFL, not T, is merged. Chomsky (2008) claims that T inherits formal features from C, building on cases where C seems to have ϕ -feature. However, to the best of my knowledge, there is no evidence indicating C has Tense-feature. If *v* has Tense-feature as Chomsky (2021) suggests, T does not exist as a head.

²⁰ Although *have* and *be* typically move above VP, they can further raise across adverb(s), i.e., between INFL and the adverb. This movement is ordinary Internal Merge, and it does not cause any cyclicity problem. In contrast, verbal expletive *do* does not move above the adverb:

- (1) i. Nora probably did not open the letter. (Baker 1991:398)
ii. * Nora did probably not open the letter. (Baker 1991:399)

This is probably because an expletive is inserted where it is needed by Externalization. Since raising of an auxiliary above *probably* is optional, where it is needed is the position just below *probably*.

This assumption is also important from the viewpoint of the sentential projection line. Why C is just above T and T is just above ν remains unexplained. Chomsky (2000) attributes this peculiarity to S-selection, but it is unfalsifiable and thus illegitimate as a hypothesis. In our theory, ν requires INFL to be a legitimate sentence. INFL inherits formal features from C, and must be below C. The sentential projection line is thus explained.

Returning to the point, if the INFL head is merged to ν P, then ν does not occupy the highest position in the whole SO. It must be raised higher; therefore ν P-raising occurs in finite sentences.

- (85) a. [$_{\nu^*P}$ EA ν^* [$_{VP}$ V IA]] (satisfies the sentencehood condition, but EA does not agree in terms of Case-feature)
 b. [INFL [$_{VP}$ V IA] [$_{\nu^*P}$ EA ν^* [$_{VP}$ V IA]]] (EA agrees with INFL in terms of Case-feature, but the whole SO does not satisfy the sentencehood condition)
 c. [[$_{\nu^*P}$ EA ν^* [$_{VP}$ V IA]] [INFL [$_{VP}$ V IA] [$_{\nu^*P}$ EA ν^* [$_{VP}$ V IA]]]] (satisfies all conditions)

In infinitives, ν P-raising is not needed, since they are not a sentence. The merger of *to* with ν P is required instead; otherwise, ν P becomes an illegitimate sentence. In other words, *to* functions as a “non-sentence” marker.

Of course, if pied-piping of ν P is not required, the verb-first order results, as above-mentioned:

- (86) [ν^* [INFL [EA ν^* V IA]]] (satisfies all conditions)

Our theory also explains the non-existence of the EPP-effect in Japanese. ν (P)-raising is difficult to justify in the language, because of its strict head-final order; ν -raising causes illegitimate T- ν order. ν P raising results in ν -EA-T order, since NPI-subjects are allowed in Japanese.

- (87) a. ν -raising
 [$_{TP}$ [$_{\nu^*P}$ EA [$_{VP}$ IA V] ν^*] T] ν^*] (T- ν order)
 b. ν P-raising
 [$_{\nu^*P}$ EA [$_{VP}$ IA V] ν^*] [EA [$_{\nu^*P}$ EA [$_{VP}$ IA V] ν^*]] T] (ν -EA-T order)

In Japanese, ν (P)-raising is not needed: owing to Case-particles, argument NPs are legitimate at interfaces without the aid of INFL, unlike in English. Then, ν , a modal or a modal final-particle is the highest position in the whole sentence, satisfying the sentencehood condition.

- (88) Japanese:
 [$_{\nu^*P}$ EA+Case [$_{VP}$ IA+Case V] ν^*] (satisfies all conditions; cf. (85))

This structure is almost the same as what Fukui and Speas (1986) claim. In fact, their idea of attributing the presence/absence of A-movement to the presence/absence of agreement/functional categories seems still essentially correct. The ν (*)P-raising analysis is actually a kind of modification of their argument, discarding superfluous assumptions such as Spec-head agreement.

3.2.6 Analysis of Subject-island

Returning to English, there is another important consequence given by ν (*)P-raising analysis. Chomsky (2008) offers an interesting contrast concerning the condition of extraction domain (CED) effect:

- (89) a. It was the CAR (not the TRUCK) of which [they found the (driver, picture)]

- b. Of which car did [they find the (driver, picture)]?
- (90) a. * It was the CAR (not the TRUCK) of which [the (driver, picture) caused a scandal]
 b. * Of which car did [the (driver, picture) cause a scandal]? (Chomsky 2008:147)

The CED effect exemplified above has some exceptional cases, as Chomsky points out:

- (91) a. It was the CAR (not the TRUCK) of which [the (driver, picture) was found]
 b. Of which car was [the (driver, picture) awarded a prize] (Chomsky 2008:147)

These examples indicate that the Subject-island effect is caused by “D-structure” properties, which Chomsky explains in terms of ν -/ ν^* -distinction; the former, introduced for unaccusatives/passives, does not form a phase, while the latter, for unergatives, does.

Therefore, the contrasts between (89), (90), and (91) cannot be explicated with the standard derivations as follows:

- (92) a. $[_{\nu^*P} [_{EA} \dots WH\dots] \nu^* [_{VP} V IA]] \rightarrow$
 b. $[_{TP} [_{EA} \dots WH\dots] [T [_{\nu^*P} e \nu^* [_{VP} V IA]]]] \rightarrow$
 c. $[_{CP} WH [C [_{TP} [_{EA} \dots e\dots] [T [_{\nu^*P} e \nu^* [_{VP} V IA]]]]]]$ (corresponds to (90))
- (93) a. $[_{\nu P} \nu [_{VP} V [_{IA} \dots WH\dots]]] \rightarrow$
 b. $[_{TP} [_{IA} \dots WH\dots] [T [_{\nu P} \nu [_{VP} V e]]]] \rightarrow$
 c. $[_{CP} WH [C [_{TP} [_{IA} \dots e\dots] [T [_{\nu P} \nu [_{VP} V e]]]]]]$ (corresponds to (89) and (91))

Since the extraction site of the WH-phrase is the same in the two derivations, namely inside SpecT, there does not seem to be any distinction possible between them. Chomsky, alternatively, proposes that the movements of the argument containing WH and WH itself proceed *in parallel*:

- (94) a. $[_{\nu^*P} [_{EA} \dots WH\dots] \nu^* [_{VP} V IA]] \rightarrow$
 b. $[_{CP} WH [C [_{TP} [_{EA} \dots e\dots] [T [_{\nu^*P} e \nu^* [_{VP} V IA]]]]]]$ (corresponds to (90))
- (95) a. $[_{\nu P} \nu [_{VP} V [_{IA} \dots WH\dots]]] \rightarrow$
 b. $[_{CP} WH [C [_{TP} [_{IA} \dots e\dots] [T [_{\nu P} \nu [_{VP} V e]]]]]]$ (corresponds to (89) and (91))

The WH-phrase moves directly from its base-position to the scope-position when the argument containing it raises to the SpecT position. Although the exact mechanism of phase conditioning improper movement still remains unclear, this successfully reduces the difference between (90) and (91) to the “D-structure” properties. Yet, it obviously violates cyclicity strictly redefined by Chomsky (2019). The operation that generate (94b) and (95b) from (94a) and (95a) obviously goes beyond what MERGE can do under Resource Restriction, which is necessary for excluding counter-cyclic variants of MERGE.

The exceptional behavior of subject extraction is, in fact, derived without appealing to counter-cyclic operation under the $\nu^*(*)P$ -raising analysis. As we have seen, what raises to SpecINFL is not an NP but a $\nu^*(*)P$ so that the “D-structure” property of $\nu^*(*)P$ is retained after “A-movement” (in what follows, I assume that, in unaccusatives and passives, IA adjoins to νP in order to make νP visible; for the visibility condition of MERGE, see section 3.3.1):

- (96) a. $[_{\nu^*P} [_{EA} \dots WH\dots] \nu^* [_{VP} V IA]] \rightarrow$
 b. $[[[_{\nu^*P} [_{EA} \dots WH\dots] \nu^* \nu P] [INFL [_{VP} V IA] \nu^* P]] \rightarrow$
 c. $[WH [C [[[_{\nu^*P} [_{EA} \dots WH\dots] \nu^* \nu P] [INFL [_{VP} V IA] \nu^* P]]]]$ (corresponds to (90))

- (97) a. $[_{vP} v [_{VP} V [_{IA} \dots WH\dots]]] \rightarrow$
 b. $[[[_{vP} [_{IA} \dots WH\dots] v] [INFL [_{VP} V IA] v^*P]] \rightarrow$
 c. $[WH [C [_{vP} [_{IA} \dots WH\dots] v] [INFL [_{VP} V IA] v^*P]]] \rightarrow$ (corresponds to (89) and (91))

Contra Chomsky's analysis, A'-extraction takes place after A-movement, as was traditionally assumed. However, subject-extraction out of v^*P is prohibited owing to Phase Impenetrability Condition. No *ad hoc* principles or operations are added.

3.2.7 Eliminating syntactically-driven NP-movement

One important consequence of the vP -raising analysis of EPP is that it reduces A-movement to vP -movement. It is true not only for subject-raising, but also other cases of A-movement.

Passives and unaccusatives are known to have two significant properties in common: (a) loss of external argument/theta-role and (b) inability of assigning accusative Case (Burzio 1986). They are thus ideally treated as a single phenomenon, caused by lexical properties of v and distinguished from its transitive counterpart, v^* .

Under sentencehood theory, most part of the traditional explanation of passives and unaccusatives are maintained. The derivation goes as follows:

- (98) John arrived./John is criticized.
 (99) *The derivation of unaccusatives/passives:*
 a. $[v [arrive John]] \rightarrow$
 b. $[John [v [arrive John]]] \rightarrow$
 c. $[INFL [[arrive John] [John [v [arrive John]]]] \rightarrow$
 d. $[[John [v [arrive John]]] [INFL [[arrive John] [John [v [arrive John]]]]]$

Suppose (non-transitive) verbalizer v blocks Case-agreement. Firstly, NP *John* moves above v (99b) to make the raised vP visible in the sense we will see in section 3.3.1; VP raises owing to some morphological factor forcing $v_{[+T]}$, INFL, and V to be adjacent (99c), just as in the case of transitive sentences we saw above; then, vP undergoes (remnant) movement to "SpecINFL," satisfying the sentencehood condition.

Yet, another derivation is possible with unaccusatives:

- (100) There arrived three men (at the hotel)
 (101) *The derivation of unaccusatives with an expletive subject:*
 a. $[there [v [arrive three men]]] \rightarrow$
 b. $[[arrive three men] [there [v [arrive three men]]]] \rightarrow$
 c. $[INFL [[arrive three men] [there [v [arrive three men]]]] \rightarrow$
 d. $[[there [v [arrive three men]]] [INFL [[arrive three men] [there [v [arrive three men]]]]]$

The difference from (99) is that the argument NP *three men* does not move out of vP . In (101), expletive *there* is in "Spec v ," making raised vP visible in (101d), instead of *three men*. For the visibility condition on MERGE, see section 3.3.1, where the explanation will be given why *there* must be present in the subject position.

Raising to Subject (RTS) is also elegantly explained in terms of sentencehood condition and the Case-feature agreement, the former being a C-I interface condition while the latter a condition imposed on Externalization.

- (102) John seems to have won. (RTS)
 (103) *The derivation of RTS:*

- a. [$v_{[+T]}$ [seem [to [John have won]]]] \rightarrow
- b. [John [$v_{[+T]}$ [seem [to [~~John~~ have won]]]]] \rightarrow
- c. [[seem [to [~~John~~ have won]]] [John [$v_{[+T]}$ {seem [to [~~John~~ have won]]}]]] \rightarrow
- d. [INFL [[seem [to [~~John~~ have won]]] [John [$v_{[+T]}$ {seem [to [~~John~~ have won]]}]]]] \rightarrow
- e. [[John [$v_{[+T]}$ {seem [to [~~John~~ have won]]}]] [INFL [[seem [to [~~John~~ have won]]] {John [$v_{[+T]}$ {seem [to [~~John~~ have won]]}]]]]]

Suppose infinitival *to* occupies the lower INFL and blocks Case-feature agreement. The derivation goes as follows: John moves across *to* to be Case-marked (103b); the upper VP also moves before INFL is introduced as in (103c) due to morphological reasons (103c, d); then, vP is raised to “SpecINFL,” satisfying all conditions of grammaticality and sentencehood (103e). Though (103b) is an operation moving an NP, it is not driven syntactically. Since agreement itself is countercyclic and thus untenable as a syntactic operation, movement of NP to a Case position is rather driven by Externalization. Also, Case-assignment by some form of Spec-Head agreement (Fukui and Speas 1986) need not be assumed; to be c-commanded by INFL or v^* without any potential interveners is enough to be Case-marked.

Lastly, let us give an analysis to Raising-to-Object (RTO) under the sentencehood theory, without appealing to syntactically-driven NP-movement.

(104) Mary believes John to have won.

(105) *The derivation of RTO:*

- a. [to [John [have won]]] \rightarrow
- b. [Mary [$v^*_{[+T]}$ [believe [John [to [~~John~~ [have won]]]]]]] \rightarrow
- c. [[believe [John [to [~~John~~ [have won]]]]] [Mary [$v^*_{[+T]}$ {believe [John [to [~~John~~ [have won]]}]]]]] \rightarrow
- d. [INFL [[believe [John [to [~~John~~ [have won]]]]] [Mary [$v^*_{[+T]}$ {believe [John [to [~~John~~ [have won]]}]]]]]] \rightarrow
- e. [[Mary [$v^*_{[+T]}$ {believe [John [to [~~John~~ [have won]]}]]]] [INFL [[believe [John [to [~~John~~ [have won]]]]] {Mary [$v^*_{[+T]}$ {believe [John [to [~~John~~ [have won]]}]]}]]]]]

First, John moves above infinitival *to* in (105b) before $v^*_{[+T]}$ and Mary are introduced; the matrix VP also raises for the above-mentioned morphological reason (105c), and INFL is merged (105d); then, the matrix v^*P raises to “SpecINFL” position, satisfying the sentencehood condition.

3.2.8 Eliminating categorizer *n*: some considerations on argument structure

Once A-movement is reduced to vP -raising, there is no longer any operation targeting N or NP. Although N'-ellipsis appears to refer to N(P), deletion is not a syntactic operation, because it violates the No Tampering Condition; Case-assignment takes place in the course of Externalization, no matter what kind of strategy is used.

This indicates that the categorizer *n* should be eliminated at least from narrow syntax, as Kayne (2008) in fact argues in terms of Antisymmetry. Rather, referring to concrete or abstract objects is a function of roots itself. Moreover, if a predicate is formed from a categorizer (*v* or *a*) and a root, its semantic properties should be deduced as a combination thereof. Here, I will briefly sketch the semantics resulting from the sentencehood theory.

Consider the totality of a speaker's knowledge of the world Σ . Σ contains all entities (e_i) of the world the speaker knows:

$$(106) \quad \Sigma = \{e_1, e_2, e_3, \dots, e_n\}$$

Every e is represented by a root. Important here is that a root does not simply indicate an entity e_i , but it does so with relation to the whole Σ ; as Saussure correctly pointed out, a word acquires meaning only by being distinguished from others. Assume

further that referring to an entity in Σ is the function of a root; *i.e.*, it does not have any other functions such as θ -assignment, which is the null hypothesis.

θ -roles are assigned by v ; relating roots/entities and the world is v 's sole function. Since the class of argument structures defined by v is limited, v is a functional category, and makes a closed class, as Kayne argues. What we have treated to be a (lexical) verb in the generative tradition is in fact a composite of v and a root, $\{v, R\}$, where R modifies v (not vice versa). Hence, R is not needed for v to function in syntactic computation. Elementary verbs such as *be*, *have*, *get*, and *give* are good examples. They are semantically bleached and seem to represent purely θ -theoretic relations between arguments.

The theory we have developed above also predicts that:

- (107) a. There is no verb that does not assign any θ -role²¹.
b. There is no element other than roots (nouns) that can be assigned a θ -role.

Thought which has no relation to the world is not only meaningless but also impossible. Relations are something connecting entities, not relations. They are no longer relations if we think about them, *i.e.*, objectify them as part of the world (nominalization of verbs).

These considerations give a new insight into the study of the argument structure. VP is traditionally analyzed as a two-layered object consisting of VP and vP :

- (108) $[_{v^*P} EA [_{v^*} [_{VP} V IA]]]$
where EA = external argument, IA = internal argument, V = (verbal) root

However, (108) is incompatible with the theory we have considered so far. Verbal root V cannot assign θ -role to its internal argument. Moreover, v^* S-selects V or VP in (108). That is, V is classified in some way, just as C (see Grimshaw (1979), a.o.), which is redundant because v is also classified (v for unaccusatives and passives, v^* for unergatives). Therefore, it would be theoretically plausible if V's classification is eliminated from narrow syntax, reducing it to v 's property.

Under the theory developed so far, on the other hand, vP is derived as follows:

- (109) a. $\{v^*, R\} \rightarrow ((\text{strong}) \text{ Transfer of } R)$
b. $v^* \rightarrow (\text{Merger of } v^* \text{ and } IA)$
c. $\{v^*, IA\} \rightarrow (\text{Transfer of } IA)$
d. $v^* \rightarrow (\text{Merger of } EA)$
e. $\{v^*, EA\}$

Assume strong Transfer (Narita 2014), which is a necessary assumption to keep to θ -assignment by head-complement relation. First, v^* and modifying R are merged, which is not necessary in the case of elementary verbs (see Marantz (2007) for word-level Transfer). Then, R is spelled-out, and v^* is merged with IA . In order to assign external θ -role, IA is Spelled-out, and finally EA is merged. As is obvious from (109), v^*P is not two-layered, except for ditransitives (Larson 1988).

3.2.9 Summary

Evidently, EPP is the consequence of $v(*)P$ -raising forced by the sentencehood condition. This even explains the variations of phrase structures of languages without the EPP effect, including verb-first and agglutinative/isolating languages:

²¹Note that weather verbs also assign theta-roles to nouns. Chomsky (1981), for example, claims that *it* in *It is raining* is a quasi-argument, not an expletive.

(110) Variations in the phrase structure under the sentencehood theory

	English French Italian	Greek Irish	Japanese
INFL	yes	yes	no
pied-piping of $v(*)P$	yes	no	no
type of movement	$v(*)P$ -raising	$v(*)$ -raising	no movement

The presence of INFL necessitates $v(*)P$ -raising. Whether $v(*)$ pied-pipes $v(*)P$ is parametrized; if it is necessary, the English-type structure results; otherwise, the verb-first order does. If INFL is not present, the Japanese-type structure is available, though it does not exclude the possibility of vP -raising not driven by the sentencehood. For the distinction between pro-drop/non-pro-drop languages (*i.e.*, the reason why English not only *allows*, but also *requires*, the subject position to be filled), see the next section.

3.3 The Empty Category Principle (ECP)

3.3.1 The Visibility Condition of MERGE: explaining ECP

We have so far seen the rich consequences the condition of sentencehood has, showing Labeling Algorithm is in fact a superfluous assumption for the explanation of SAI and EPP. Yet another problem remains which Chomsky tries to solve by LA, namely, Empty Category Principle (ECP). In this subsection, we will determine whether the ECP effect can be easily explained by the condition of sentencehood, adding a general condition of visibility, which is a natural consequence of considerations on learnability.

Syntax abounds with invisible elements: empty categories, implicit arguments, unpronounced copies, and so on. However, their distribution is tightly constrained, for otherwise children acquiring a language would not learn where such invisible items are and what role they play in a sentence. Thus, the condition (111) is a natural assumption:

(111) Visibility Condition on MERGE

Every application of MERGE must be visible in some way.

We must clarify what “visibility” means. In the most primitive sense, it means that, when $MERGE(X, Y)$ forms $\{X, Y\}$, (at least some part of) X and Y must be phonetically realized. However, that is too simple a condition, which fails to account for all cases of merger of invisible elements. Thus, we must define “visibility” more loosely.

(112) When $MERGE(X, Y)$ forms $K = \{X, Y\}$, K is visible if and only if

- both X and Y contain phonetically realized elements, or
- X is phonetically realized and Y 's presence can be inferred, or
- X is phonetically realized and Y need not be interpreted.

(112c) is required to handle intermediate copies of successive cyclic movement. They do not have to be learned but are independently derived from general conditions such as phase.

Then, in what cases Y 's presence is inferred (*i.e.*, learnable)? There are two representative cases:

(113) a. Which book did John read e ? (the lower copy of movement)

- e_i Every man loves some $_i$ woman. (the highest copy of covertly raised elements)

(113b) is a self-evident case of movement; once an element is learned as an operator by its actual use in ordinary conversations, it is self-evident that it moves to the scope position. (113a) is more important: the gap left by movement is hard to identify, and it is natural to assume that its presence is visible only when a selectional relationship is established between the gap and the predicate. Thus, we have:

- (114) Y's presence is inferred when
- a. Y's head (*i.e.*, an LI in the MS domain of Y's sister) is selected, or
 - b. Y is part of a chain of an operator and in the scope position.

Note that Y may have more than one *structural*, or, put it more precisely, "potential" head according to Epstein, Kitahara and Seely's definition of MS. However, considering the duality of semantics, every occurrence of Y has exactly one *semantic* head, which must be one of the structural/potential heads. Take an example of WH-movement:

- (115) Which book did John read?

In this case, the head of the lower copy of *which book* is *book*, selected by *read*. The head of the higher copy is, on the other hand, *which*, satisfying the scopal condition imposed on the WH-/Q-feature. *Book* is just pied-piped due to some phonological reasons and does not play any semantic role in the higher position.

Returning to the point, let us examine how the core cases of the ECP effect are explained in terms of the sentencehood and the visibility condition.

- (116) a. Who do you think read the book?
b. * Who do you think that read the book? (Chomsky 2015:10)

(116a, b) are representative cases of "that-t effect." (116b) has the following structure:

- (117) [Who_i do you think [e_i [that [[e_i v*_[T] e_j]_k [INFL VP_j e_k]]]]]

What is problematic here is that no part of the higher copy of v*P is phonetically realized there; the merger of [INFL VP_j e_k] (X in (112)) and v*P (Y in (112) and (114)) does not satisfy the condition (112). Assuming S-selection is a strictly local relationship held between heads, the matrix verb *think* selects *that*, not v*_[T]; although *that* and INFL in the embedded clause are connected by inheritance relationship, *that* does not select v*_[T]. The upper copy of v*P does not meet the conditions (114a, b), and therefore it fails to satisfy the Visibility Condition on MERGE. The ungrammaticality of (116) is thus explained.

(116a), on the other hand, has the structure:

- (118) [Who_i do you think [[e_i v*_[T] e_j]_k [INFL VP_j e_k]]]

Here, it is v* that the matrix verb *think* selects. There might appear two candidates for the head in the MS domain of the sister of *think*, namely v* and INFL. However, INFL cannot be selected because of the lack of semantic features. v* with T-feature is thus the most prominent element when the whole clause is interpreted as a sentence. Hence, (116a) meets the condition (114a).

3.3.2 Italian: absence of the ECP-effect

Another important point of ECP is that it is not observed in Italian:

- (119) Chi_i credi che e_i verrà?
 who think that will-come (Rizzi 1982:117)

The embedded clause of (119) has the following structure:

- (120) [_{CP} [_{CP} che [_{CP} [_{CP} verrà] [INFL *v*P]]]]

The difference between English and Italian is reduced to the parameter concerning the realization of verbs. As we have already seen, English and French do not allow main verbs to remain inside *v*P. In fact, this is the reason why these two languages exhibit the EPP effect. On the other hand, Italian *forces v*(*) (main verb) to remain in situ. It is exemplified by the following²².

- (121) a. Rita pag-a sempre tutto.
 Rita pay-PRES.IND.3SG always all
 ‘Rita always pays all’
 b. ?* Rita sempre pag-a tutto. (Murakami 2013:127)

- (122) The fact that John probably has made several mistakes is well-known. (Kayne 1989) (=78)

Since Italian does not exhibit the EPP effect, its derivation should be different from that of French. It also contrasts with English, where the auxiliary raises to a position between NegP and INFL; if this were the case in Italian, (121b) would be grammatical as well as (122). In fact, Italian main verbs stay in situ in *v*(*) position and raises to SpecINFL pied-piping the whole *v*(*)P. Then, it forms an amalgam with INFL by PF-merger. The landing site of this PF-merger is *v*(*), contrary to French.

- (123) The derivation of (121a):
 a. [[Rita pag(=*v**_[T]) e_i]_j sempre [-a(=INFL) [V tutto]_i e_j]] → (PF-merger of *pag*(*v**) and -a(INFL))
 b. [[Rita pag-a(=*v**_[T]) e_i]_j sempre [INFL [V tutto]_i e_j]]

The higher copy of *v*P in Italian is phonetically visible thanks to the in-situ *v*, and the condition (112a) is successfully met, even if the External Argument is not phonetically realized. The difference in EPP/pro-drop parameter and ECP between English and Italian is thus explained.

3.3.3 French *que-qui* alternation

French offers another interesting phenomenon concerning ECP. It is a well-known fact that French behaves like a “mixture” of English and Italian: it is apparently not a pro-drop language, but it does not exhibit the *that-t* effect as in English when *qui* is used as a complementizer.

- (124) a. * L’homme [Op que [t est venu]]
 ‘The-man *que* is come’
 b. L’homme [Op qui [t est venu]]
 ‘The-man *qui* is come’
 (125) a. L’homme [Op que [tu as vu t]]
 ‘The-man *que* you have seen’

²²However, (51b) and (53b) are grammatical, in contrast with (121b). *Already*-type adverbs might have an adjunction site inside *v*(*)P. See Cinque (1999).

- b. * L'homme [Op qui [tu as vu t]]
 'The-man qui you have seen' (Rizzi and Shlonsky 2007)

Rizzi and Shlonsky (2007) attempt to explain the *que-qui* alternation by appealing to the Subject Criterion, assuming that *qui* is an amalgam of *que* and an expletive-like nominal *-i* with the value of its number feature unspecified. The assumption of *qui* as an amalgam directly excludes (125b), which leaves no room for the subject expletive *-i*. It still correctly accommodates (124b), with *-i* in the subject position. Finally, the difference in grammaticality between (124a) and (125a) is explicated in terms of Subject Criterion, banning elements in the subject position from further moving upward.

Although their treatment of *qui* seems to me to be correct, Criterial Freezing is not an explanation but is just a generalization, which, if correct, needs further explanation. Alternatively, let us analyze the above contrasts with the condition of sentencehood. Following Rizzi and Shlonsky, assume PF-merger between *que* and *-i* forming *qui*, which takes place only if *-i* adjoins to the upper $v(*)P$ and it is adjacent to *que*. Then, in (124b), even though $v(*)$ and VP are not phonetically realized, the upper $v(*)P$ remains visible, in contrast with (124a). The sentencehood condition is satisfied, and the contrast between (124a) and (124b) is explained.

- (126) [L'homme [Op [que [[-i [$\Theta_P v$]] [INFL est venu]]]]]

3.3.4 Dialectal Variations in the ECP-effect

PF-merger analysis even extends to the problematic cases in English that Sobin (1987) points to through an experiment in which students are asked to judge the acceptability of the sentences presented.

- (127) Who did you say that kissed Harriet? (Sobin 1987:33)

The result was confusing; sentences with a *that-t* configuration like (127), generally considered to be a representative case of the ECP effect, were rejected only by 17.5% of informants, indicating some parametric variation between English speakers or dialects concerning the *that-t* effect.

However, the problem can be solved in exactly the same way as in the French *que-qui* alternation. Suppose that, for speakers who allow *that-t* configuration, *that* is a phrase-level adjunct of the embedded $v(*)P$, and it undergoes PF-merger with the adjacent covert C. This is not an unreasonable assumption, since *that* is used not only as a complementizer but also as a determiner NP; it has been generally assumed that vP has an adjunction site (escape hatch) for a phrase-level object, and *that* adjoins there, just like French expletive-like *-i*²³.

Sobin reports another interesting fact:

- (128) Who did you ask whether kissed Harriet? (Sobin 1987:58)

Here, 97.6% of informants rejected sentences of this type. The result sharply contrasts with that of *that-t* configuration. However, it is no surprising fact under our analysis so far. Contrary to *qui* and *that*, *whether* can only be used as a complementizer, and does not function as an NP. Thus, it cannot adjoin to the embedded vP , and the phonetically unrealized vP violates the visibility condition²⁴. The asymmetry between *whether* and *that* is parallel with that of *que* and *qui*.

²³This analysis may extend to the absence of *that-trace* effect in a relative clause:

- (1) the man that killed Mary

That is parametrized as to its usage. When used in a relative clause, it can adjoin to $v(*)P$ as an adjunct, and then forms an amalgam with C by PF-merger, even in Standard English.

²⁴It is difficult to attribute the ungrammaticality of (128) to WH-island or Relativized Minimality. As Sobin points out, Icelandic sometimes allows extraction out of WH-island, but the asymmetry between *that* and *whether* is observed.

3.3.5 The ECP effect without agreement

The sentencehood theory can even explain what LA cannot, that is, the ECP effect with locative inversion, which does not involve agreement between locative PP and INFL:

- (129) a. It's in these villages that we all believe can be found the best examples of this cuisine.
 b. * It's in these villages that we all believe that can be found the best examples of this cuisine. (Bresnan 1994:96)

Since LA only covers cases where agreement takes place in {XP, YP} structure, it cannot account for why (129b) is ungrammatical.

Our theory correctly predicts (129b) to be ungrammatical (irrelevant details are omitted):

- (130) [that [[_{EPP} $v_{[T]}$] [can be found the best examples of this cuisine]]]

The structure is essentially the same as (117), and the same explanation holds for (129b): v is not selected by *that*, its local c-commander, failing to meet the condition (114a). (129a), on the other hand, is grammatical because v with T-feature is selected by the matrix verb, and it does not violate the Visibility Condition.

Note that we have not added any rules or conditions to narrow syntax. The visibility condition on MERGE is imposed on Externalization by the general condition of learnability. The theory we have established so far significantly simplifies syntax, reducing unnecessary assumptions such as LA.

3.3.6 Summary

The contrasts between English, French, and Italian are summarized as below:

- (131) [[EA $v(*) e_i$]_j [INFL [(Adv) [α [Neg [[V IA]_i e_j]]]]]]]

	English (main verb)	English (auxiliary)	French	Italian
EPP	yes	yes	yes	no
ECP	yes	yes	yes/no	no
base position	V	v	v	v
stay in situ	yes	no	no	yes
final position	V	α	INFL	v

French exhibits the ECP effect in clauses headed by the complementizer *que*, but it is possible to extract the subject from inside the $v(*)$ P, adjoining an expletive-like element *-i* to the upper $v(*)$ P, which forms an amalgam *qui* with *que*.

Languages that do not require $v(*)$ P-raising exhibit no ECP effect, since $v(*)$'s visibility is always guaranteed. For example, Japanese, a language without INFL, allows a subject to be extracted, as well as an object.

- (132) *Japanese*:

- a. dare-ga minna-ga kitaku-sita to omotteiru-no?
 who-Nom everyone-Nom go-home-did C think-Q
 'Who does everyone think went home?'
 b. dare-o minna-ga Taro-ga nagu-tta to omotteiru-no?
 who-Acc everyone-Nom Taro-Nom hit-Past C think-Q
 'Who does everyone think Taro hit?'

Irish, a $v(*)$ -raising language, behaves similarly:

(133) *Irish*:

- a. an duine a mheas tú a chonaic tú t
the person aL thought you aL saw you
'the person that you thought that you saw' (McCloskey 1979)
- b. an t-ainm a hinnseadh dúinn a bhí t ar an áit
the name aL was-told to-us aL was on the place
'the name that we were told was on the place' (McCloskey 2002)

4 Why does the sentence exist?

4.1 Rereading Saussure: the nature of signs

Before turning to the problem of the existence of sentence in human language, let us consider the nature of Externalization and signs. Chomsky has repeatedly claimed that human language is not optimally designed for Externalization.

- (134) a. John or Bill is in the room (one or the other)
b. John or the men **is/*are* in the room (either John is in the room or the men are in the room) (Chomsky 2021)

(134b) is ungrammatical regardless of the choice of copulas. Important here is that the intended semantic structure is unproblematic as the parallel sentence (134a) illustrates. This simple fact indicates the existence of a certain class of semantically well-formed sentences that cannot be uttered. Chomsky thus concludes that language is a system of thought, not that for communication.

It is worth noting that under Chomsky's schema of biolinguistics, the emergence of Externalization is regarded as an "accidental" phenomenon in the evolution of language. It has also seldom been discussed what role the sign (lexical item) plays under the Minimalist Program, despite their significance in the traditional philosophy of language. In fact, Externalization and the sign, which are closely related phenomena of human language, are essential factors for the explanation of the concept of sentence.

Saussure defines signs as follows:

- (135) I call the combination of a concept and a sound-image *sign*, but in current usage the term generally designates only a sound-image, a word, for example (*arbor*, etc.). One tends to forget that *arbor* is called a sign only because it carries the concept "tree," with the result that the idea of the sensory part implies the idea of the whole. Ambiguity would disappear if the three notions involved here were designated by three names, each suggesting and opposing the others. I propose to retain the word *sign* [*signe*] to designate the whole and to replace *concept* and *sound-image* respectively by *signified* [*signifié*] and *signifier* [*signifiant*]; the last two terms have the advantage of indicating the opposition that separates them from each other and from the whole of which they are parts. (Saussure 1916:99)

From (135), several consequences can be drawn according to Saussure. One is the arbitrariness of signs:

- (136) The bond between the signifier and the signified is arbitrary. Since I mean by sign the whole that results from the associating of the signifier with the signified, I can simply say: *the linguistic sign is arbitrary*. (Saussure 1916:100)

Maruyama (1981) warns that (136) should not be confused with what Plato said in *Cratylus*:

- (137) People certainly have misunderstood the concept of ‘arbitrariness’ for a long while; that is, they have been liable to take it to mean “there are no necessary or natural connections between things or concepts and words, their names.” However, it has been almost one century since the view which treats language as a nomenclature was denied, and it is obvious that such arbitrariness cannot be either linguistic or semantic problem. And, just reading the following passage quoted from Saussure’s manuscript 14 (perhaps written in 1897) is enough to see how irrelevant the above double-quoted thesis is to Saussure’s view of language:

“Zoologists, anthropologists, and linguists, when they speak to the public about the so-called articulated language, do so as if it were a concept self-evident for everyone. There, they confuse ‘articulation’ with a mental entity given to language, like ‘a set of ideas.’ (3302)”

In the above quotation, no such words as arbitrariness or necessity are found, but one can observe Saussure’s fundamental idea “the unity of linguistic expression and meaning”; in other words, his negative attitude against so-called intellectualism that detaches language and meaning and focuses on the latter is clearly expressed in the view of ‘articulated language.’ (Maruyama 1981:297, translated by IN)

Then, what is the “concept” connected with the “sound image”? Benveniste (1966) offers an important argument concerning this point:

- (138) One of the components of the sign, the sound image, makes up the signifier; the other, the concept, is the signified. Between the signifier and the signified, the connection is not arbitrary; on the contrary, it is *necessary*. The concept (the “signified”) *bœuf* is perforce identical in my consciousness with the sound sequence (the “signifier”) *böf*. How could it be otherwise? Together the two are imprinted on my mind, together they evoke each other under any circumstance. There is such a close symbiosis between them that the concept of *bœuf* is like the soul of the sound image *böf*. The mind does not contain empty forms, concepts without names.

“Pure concepts” that have emerged without any interaction with non-linguistic systems are often assumed in the generative literature, but such prelinguistic concepts are difficult to justify. Saussure also says:

- (139) *mouton* does not have the same value as *sheep* in English. One reason is that when we talk about *mouton* on the dinner table, we have to say *mutton* [in English].

Even if “lexical” empty items exist, they are at most “imitation” or “abstraction” of true lexical categories, like “barefoot waterskiing,” which requires us to practice with skis before skiing barefoot (Dennett 2017). Invisible elements would be legitimate only when they are generated by the interaction between syntax and Externalization.

Now it is obvious that signs with the *signifier* and the *signified* are indispensable items in order that they have meanings. In other words, Externalization is a crucial factor for MERGE to work fully. Without Externalization, no lexical items would appear in the human brain, and hence no linguistic computation would be possible. It is in fact a departure from the Chomskian view of language. Chomsky regards the emergence of Externalization as an accidental phenomenon irrelevant to core syntax. However, words in syntax without any connection to SM-systems are simply an empty object, and hence SOs formed of them are, too. Although it is generally assumed that MERGE can apply to null elements and work as a successor function (Chomsky 2008), it does not suffice as a “system of thought” in Chomsky’s words.

Returning to the point, let us take a closer look at the relationship between the sign and Externalization, setting aside its implication to syntax.

In Saussure's terminology, *langue* and *parole* are dependent on each other: *langue* is a system of signs formed by *parole* (each individual's use of language²⁵), and *parole* is regulated by the social code which *langue* determines. We cannot derive one from the other.

The sign and Externalization are thus necessarily *social*; that is to say, there would be no signs or utterance without conversation; they must be used, or they will not play any role in us. Here, we should limit the meaning of "use" to "use with intention" or even "use to affect the world." Of course, we use some physical or mental devices of our body without any intention: we often mutter something unconsciously, without intention to affect the world. However, it does not matter to the core problems of signs. We can mutter anything in almost any situation, unless forbidden by some social codes. As we learn or modify our knowledge of signs by connecting their actual uses with the situations where they are used, mutterings cannot be the key to determine or affect the meanings of signs.

Because of the sociality of the sign and Externalization, an utterance as a set of signs inevitably has illocutionary force in conversation. It is in fact the source of sentencehood, as we will see in the next subsection.

4.2 Externalization and sentencehood

When signs are used with intention, they inevitably carry illocutionary force. As Austin (1962) correctly points out, there is no "constative proposition,"²⁶ whose function is just "stating" its meaning, as in a proposition in syllogism. Austin's argument also holds for words: we do not utter words (nonsentential utterances) without any intention to affect the hearer; otherwise it would not be part of conversation. Sentences and words always have some illocutionary force, and the hearer is expected to answer properly (not necessarily by a linguistic act). In fact, if language were unable to express illocutionary force, nobody would try to use language: it would be a simply meaningless act.

In terms of the evolution of language, it is conceivable that linguistic signs evolved from animal's chattering. Animals also communicate via signs with illocutionary force, though their speech does not have hierarchical structures. Once MERGE appeared, it became possible to divide signs into smaller units (Wray 1998, Kirby and Christiansen 2003)²⁷:

- (140) a. /abcdef/ (a, b, c, etc. are phonemes) →(MERGE)
b. [ab [c [d [e f]]]]

If illocutionary force itself is detached from a sign via MERGE, it functions as an independent head (IF in the following example):

- (141) a. /abcdef/ →(MERGE)
b. [IF [abcdef]]

Then, what is the IF head? The most plausible candidate is *v* with tense/modal-feature. In fact, as Austin observes, the type of illocutionary force a sentence has largely depends on its verbs, modals and their grammatical moods.

- (142) a. Shut it, do.
b. I order you to shut it. (Austin 1962:73)
- (143) a. You may shut it.
b. I give permission /I consent to your shutting it.

²⁵The term *parole* does not solely mean "utterance" of each sentence; it also covers individual's selection and merger of signs. Maruyama (1981) says that the latter sense has far more important in considering the nature of *parole*.

²⁶For related arguments in syntax, see Ross (1970), though Ross's performative deletion rule is difficult to justify.

²⁷MERGE is a *decomposing* operation when applied to an already formed object; it is usually treated as a *building* operation, but setting aside processing, both are two sides of one thing.

- (144) a. You must shut it.
 b. I order you /I advise you to shut it. (cf. (142))
- (145) a. You ought to shut it.
 b. I advise you to shut it. (Austin 1962:74)

(b) sentences are what Austin calls “explicit performatives.” Their types of illocutionary force are determined by the matrix verbs with the particular form [+1st person, +singular, +present, +active]. (a) sentences are called “primary performatives.” They have the same functions as (b) sentences when uttered, but their illocutionary force is signified by the modal or mood of the verb. In both cases, $v(*)_{[T]}$ plays a crucial role in determining their types. Similar examples can even be found in “constative” sentences:

- (146) a. I state/argue/suggest/bet that he did not do it.
 b. He did not do it. (Austin 1962:133-134)

Uttering (146b) functions as a “constative” act, which is enabled by the *absence* of a modal in the v position. The same holds for the epistemological use of modals.

Therefore, we can safely conclude that the locus of the illocutionary force is v ’s tense/modal feature²⁸. The core function of tense/modal-features is the determination of the type of illocutionary force.

Generally, functional heads are under the influence of labeling algorithm, or more generally, the requirement of endocentricity. They are detected as heads by Minimal Search, a general principle of efficient computation. This is also the case for $v_{[T]}$. Once $v_{[T]}$ becomes a distinct lexical item, sentencehood is defined with the aid of Minimal Search, as well as other functional “heads.” This is what we call “sentencehood”:

- (147) *The Condition of Sentencehood:*
 An SO is recognized as a sentence if a tense/modal feature is in its Minimal Search domain. (= (20))

The reflection above also helps us to determine how to treat interjection. Tokieda regards interjection and similar nonsentential expressions as a sentence when uttered with illocutionary force:

- (148) [...] the word *kaji* (‘fire’) listed in the dictionary, as long as it only gives us the image of fire, is just a *word*; however, it is certain that when one cries out ‘kaji’ seeing a fire just in front of him/her, the expression carries not only the image but also his/her judgment, emotion, etc. about the image. In other words, we can say that fire, a thing, is integrated in some way in the speaking subject. Without such integration of subject and object, the linguistic expression ‘kaji’ cannot be used. This is the reason why the word ‘kaji’ is recognized as a sentence. (Tokieda 1941:346-347)

As we saw in section 2.2, what one takes to be a sentence a little differs from person to person, owing to non-linguistic factors such as education, though the core intuition of sentencehood is essentially the same. Interjection is actually a kind of “peripheral” phenomena, contrary to the core phenomena we discussed above. Still, interjection has something linguistic that makes it worth being called a sentence. As Tokieda observes, it is illocutionary force that distinguishes them from mere words. Thus, interjection is a “fossil” of animal’s chattering.

²⁸Tense/modal feature is very similar to what Ross (1970) calls “performative feature,” though we do not assume the rule of performative deletion.

5 Conclusion

Sentencehood is a core syntactic concept and emerged as a result of the social nature of language. The sentencehood theory can explain various phenomena including SAI, EPP, ECP, and related parameters. It also enables the elimination of unnecessary assumptions in Minimalism such as Labeling Algorithm, counter-cyclic A-movement, and “nominalizer” *n*. These outcomes, in turn, confirm the sentence to be a formal object which largely depends on syntax.

One consequence of this approach is that the most fruitful way for Minimalism to proceed is to construct a theory of UG adequate in both Cartesian and Humean sense. The study of generative grammar nowadays leans toward the Cartesian way, that is, the “philosophical” way of thinking that reduces the class of concepts in the theory as much as possible. Consequently, as “philosophical” studies proceed, many important concepts have been forgotten without being explained. “Sentence” is one such concept.

It is true that syntax abounds with unnecessary concepts, and the “philosophical” thinking is surely important for generative grammar to make advances. However, it is also true that some concepts excluded from syntax as unnecessary are deeply rooted in our intuition. Unless we investigate what lies behind it, we cannot accomplish the mission of syntax, namely, the cognitive science of linguistic intuition. Such an approach also has rich empirical outcomes, as we saw in section 3.

Sentencehood theory also informs us that we should not overlook the fact that language is not just knowledge but an act. SAI, EPP, and ECP have been regarded as core syntactic phenomena that can only be explained in terms of *competence*; such intuition held by syntacticians is correct, but once we want to give a deeper “principled explanation” that is biologically adequate to those phenomena, we cannot do without delving into the interactions between *competence* and *performance*. In fact, no other satisfactory explanation of sentencehood is conceivable.

That might seem to contradict the argument that the sentencehood largely depends on “competence,” but it does not. Evidently, separating competence from performance is crucial for scientific study of core syntax. However, what we call competence is a subsystem of the Faculty of Language in the broad sense (FLB, Hauser et al. (2002)), which interacts with other adjoining subsystems. We should seriously contemplate how core syntax is affected or even *transformed* by *performance* systems.

References

- Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural language and linguistic theory* 16:491–539.
- Austin, John L. 1962. *How to do things with words*. Cambridge, Mass: Harvard University Press.
- Baker, C. L. 1968. Indirect questions in English. Doctoral Dissertation, University of Illinois.
- Baker, C. L. 1991. The syntax of English *not*: the limits of core grammar. *Linguistic Inquiry* 22:387–429.
- Benveniste, Émile. 1966. *Problèmes de linguistique générale*, volume 1. Paris: Gallimard. [trans. Mary Elizabeth Meek, *Problems in general linguistics*, Coral Gables, FL: University of Miami Press, 1971].
- Bever, Thomas. 2009. Remarks on the individual basis for linguistic structures. In *Of minds and language*, ed. Massimo Piattelli-Palmarini, Juan Uriagereka, and Pello Salaburu. Oxford: Oxford University Press.
- Birner, Betty J. 1992. The discourse function of inversion in English. Doctoral Dissertation, Northwestern University.
- Borsley, Robert D., Maria-Luisa Rivero, and Janig Stephens. 1996. Long head movement in Breton. In *The syntax of the Celtic languages: a comparative perspective*, ed. Robert D. Borsley and Ian Roberts, 53–74. Cambridge: Cambridge University Press.
- Bresnan, Jane. 1994. Locative inversion and the architecture of universal grammar. *Language* 70:72–131.
- Burzio, Luigi. 1986. *Italian syntax: a government-binding approach*. Dordrecht: Reidel.
- Carstens, Vicki, Norbert Hornstein, and Daniel Seely. 2016. Head-head relations in *Problems of Projection*. *The Linguistic Review* 33:67–86.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Pkenum Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1994. Bare phrase structure. *MIT Occasional Papers in Linguistics* 5.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Robert Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 291–321. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.
- Chomsky, Noam. 2015. Problems of projection: extensions. In *Structures, strategies and beyond: studies in honor of Adriana Belletti*, ed. Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3–16. Amsterdam and Philadelphia: John Benjamins.
- Chomsky, Noam. 2019. UCLA lectures. manuscript, MIT/University of Arizona.
- Chomsky, Noam. 2021. Minimalism: where we are now, and where can we hope to go. *Gengo Kenkyu* 160.
- Chomsky, Noam, Ángel J. Gallego, and Dennis Ott. 2019. Generative grammar and the faculty of language : Insights, questions, and challenges. *Catalan journal of linguistics* Special Issue:229–261.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. Oxford: Oxford University Press.
- Collins, Chris. 1991. Why and How come. *MIT Working Papers in Linguistics* 15:31–45.
- Dennett, Daniel. 2017. *From bacteria to Bach and back: the evolution of minds*. New York: W. W. Norton and Company.
- Donati, Caterina. 2006. On *Wh*-head movement. In *Wh-movement: moving on*, ed. Lisa Lai-Shen Cheng and Norbert Corver, 21–46. Cambridge, MA: MIT Press.
- Epstein, Samuel, Hisatsugu Kitahara, and T. Daniel Seely. 2020. Unifying labeling under minimal search in “single-” and “multiple-specifier” configurations. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona* 22:1–11.
- Fukui, Naoki, and Hiroki Narita. 2017. Merge and (a)symmetry. In *Merge in the mind/brain*, ed. Naoki Fukui, 35–74. London: Routledge.

- Fukui, Naoki, and Margaret Speas. 1986. Specifiers and projection. *MIT Working Papers in Linguistics* 8:128–172.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.
- Hamblin, C. L. 1958. Questions. *Australasian journal of philosophy* 36:159–168.
- Hauser, Mark D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–1579.
- Henry, Alison. 1995. *Belfast English and standard English: dialect variation and parameter setting*. Oxford: Oxford University Press.
- Hjelmslev, Louis. 1928. *Principes de grammaire générale*. København: Andr. Fred. Høst.
- Holmberg, Anders. 2016. *The syntax of yes and no*. Oxford: Oxford University Press.
- Hooper, Joan B., and Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465–497.
- Hornstein, Norbert, and Juan Uriagereka. 2002. Reprojections. In *Derivation and explanation in the minimalist program*, ed. Samuel Epstein and T. Daniel Seely, 107–132. Malden, MA: Blackwell.
- Kayne, Richard. 1989. Notes on English agreement. *CIEFL Bulletin* 1:41–67.
- Kayne, Richard. 2008. Antisymmetry and the lexicon. *Linguistic Variation Yearbook* 8:1–31.
- Kirby, Simon, and Morten H. Christiansen. 2003. *From language learning to language evolution*, 272–294. Oxford: Oxford University Press.
- Kitahara, Hisatsugu. 2020. MERGE and minimal search. talk at Sophia University.
- Kuroda, S.-Y. 1969. Remarks on the notion of subject with reference to words like *also*, *even*, and *only*, part 1. *Annual Bulletin (Research Institute of Logopedics and Phonetics, University of Tokyo)* 3:111–129.
- Lacerda, Renato. 2020. Middle-field syntax and information structure in Brazilian Portuguese. Doctoral Dissertation, University of Connecticut.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.
- Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235–290.
- Lema, José, and Maria-Luisa Rivero. 1991. Types of verbal movement in Old Spanish: modals, futures, and perfects. *Probus* 3:237–278.
- Lema, José, and Maria-Luisa Rivero. 1992. Inverted conjugations and V-second effects in Romance. In *Theoretical analyses in linguistics*, ed. Christiane Laeufer and Terrell A. Morgan, 311–328. Amsterdam: John Benjamins.
- Mahajan, Anoop. 2003. Word order and (remnant) VP movement. In *Word order and scrambling*, 217–237. Malden, MA: Blackwell.
- Marantz, Alec. 2007. *Phases and words*, 191–222. Seoul: Dong-In Publishing Co.
- Maruyama, Keizaburo. 1981. *Saussure-no-sisoo*. Tokyo: Iwanami Shoten.
- Matushansky, Ora. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37:69–109.
- McCloskey, James. 1979. *Transformational syntax and model theoretic semantics*. Dordrecht: Reidel.
- McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the Minimalist Program*, ed. Samuel D. Epstein and T. Daniel Seely, 184–226. Malden, MA: Blackwell.
- Michaelis, Laura A., and Knud Lambrecht. 1996. Toward a construction-based theory of language function: the case of nominal extraposition. *Language* 72:215–247.
- Miller, George A., and Noam Chomsky. 1963. Finitary models of language users. In *Handbook of mathematical psychology*, ed. R. D. Luce, R. Bush, and E. Galanter, volume 2, 419–492. Wiley.
- Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, Mass: MIT Press.
- Müller, Gereon. 2004. Verb-second as vP-first. *The journal of comparative Germanic linguistics* 7:179–234.
- Murakami, Madoka. 2013. Verb movement: the contrast between English and Italian. *Studies in Linguistics* 5:117–143.
- Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax*. Amsterdam: John Benjamins.
- Narita, Hiroki, and Naoki Fukui. 2022. *Symmetrizing syntax: merge, minimality, and equilibria*. New York: Routledge.
- Narita, Hiroki, Hironobu Kasai, Takaomi Kato, Mihoko Zushi, and Naoki Fukui. 2017. 0-search and 0-merge. In *Merge in the mind/brain*, ed. Naoki Fukui, 127–154. New York: Routledge.
- Ransom, Evelyn N. 1986. *Complementation: its meaning and forms*. Amsterdam: John Benjamins.
- Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

- Rivero, Maria-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic review* 8:319–351.
- Rivero, Maria-Luisa. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural language and linguistic theory* 12:63–120.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1997. The fine structure of left periphery. In *Elements of grammar: Handbook in generative syntax*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- Rizzi, Luigi, and Uri Shlonsky. 2007. Strategies of subject extraction. In *Interfaces + recursion = language?: Chomsky's minimalism and the view from semantics*, ed. Uli Sauerland and Hans-Martin Gärtner, 115–160. Berlin and New York: Mouton de Gruyter.
- Roberts, Ian. 1994. Two types of head movement. In *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 281–337. Cambridge: Cambridge University Press.
- Roberts, Ian. 2010. *Agreement and head movement*. Cambridge, MA: MIT Press.
- Rochemont, Michael S., and Peter W. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Ross, John R. 1970. On declarative sentences. In *Readings in English transformational grammar*, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 222–272. Washington D. C.: Georgetown University Press.
- Rothstein, Susan. 1983. The syntactic forms of predication. Doctoral Dissertation, MIT.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot. [trans. Wade Baskin, *Course in general linguistics*, New York: Philosophical Library, 1959].
- Sobin, Nicholas. 1987. The variable status of Comp-trace phenomena. *Natural Language and Linguistic Theory* 5:33–60.
- Tokieda, Motoki. 1941. *Kokugogaku-genron*. Tokyo: Iwanami Shoten.
- Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral Dissertation, MIT.
- Wray, Alison. 1998. Protolanguage as a holistic system for social interaction. *Language and Communication* 18:47–67.
- Yamada, Yoshio. 1936. *Nihon-bunpoogaku-gairon*. Tokyo: Hoobunkan.
- Zanuttini, Raffaella, and Paul Portner. 2003. Exclamative clauses: At the syntax-semantics interface. *Language* 79:39–81.