

## Voice Only: No Passive, Causative and Applicative Heads<sup>1</sup>

Chigchi Bai

*School of Foreign Languages, IMUFE*

Email: chigchi.hohhot@qq.com

21-03-2024

### Abstract

The core proposal to drill in this paper is that there is only one type of syntactic head endowed with the ability to introduce arguments in the inventory of functional heads in UG, that is, Voice. In justifying this proposal, I pursue a Split VoiceP approach to articulating Voice and the voice domain, thereby syntactically deriving voice alternations through the single engine, Voice. Based on Legate's (2003) observation of reconstruction effects in English passives and accusative-to-dative raising in Japanese causative-of-passive sentences, it is argued that subjects of sentences including both passives and actives/causatives are introduced beforehand by a voice head to be assigned subjecthood. The proposed analysis not only accounts for relevant facts and properties that previous approaches fail to capture but also unifies and simplifies syntactic operations deriving voice constructions. The previously postulated passive (agent-less Voice), causative and applicative heads are therefore subsumed under Voice. As such, introduction of arguments (external, internal or applied), being unconstrained, comes down to Free Merge (Chomsky 2013, 2015).

**Keywords** Voice, passive, causative, applicative, argument, subject, merge

### 1. Introduction

Voice as a syntactic head characterized by previous studies falls into two types: Voice that introduces arguments (notated as "+D Voice" in this paper) and Voice that does not (notated as "-D Voice"). This paper presents empirical and theoretical supports for an analysis wherein Voice can only be +D Voice and there are a variety of +D voice heads comprising a Split VoiceP. That is, the voice domain splits into more than one Voice projection and the head of each projection is bijective with a particular type of argument, as shown below.

(1) ... [<sub>VoiceP3</sub> DP2 [<sub>VoiceP2</sub> DP2 [<sub>VoiceP1</sub> DP1 [<sub>VP</sub> V ...

Simply stated, Voice1 refers to a transitivizing head, roughly corresponding to the light verb *v*, and Voice2 refers to a causativizing head, roughly corresponding to what has been notated as *v*/Voice<sub>[cause]</sub> or Cause in the literature. What makes Voice2 distinct from

---

<sup>1</sup> This work was supported by NSSFC [No. 21XYY018].

the previously postulated head is that it introduces not only a causer argument but also the subject of passive clauses, that is, an internal argument of the verb, as elaborated in §3-4. Voice3 refers to a head that introduces a high applied argument in gapless indirect passives, roughly corresponding to an applicative head, often notated as High Appl. What makes Voice3 distinct from High Appl is that it also introduces the subject of a passive phrase formed on the basis of a causative verb, which occupies Voice2, as elaborated in §3-4. As such, all types of argument introducers are subsumed under Voice.

On the surface, the model in (1) does not differ from those that are constituted by a succession of any of voice, causative, passive, and applicative heads, and therefore is subject to Collins' (2024: 104) Terminological Assumption.

## (2) Terminological Assumption

Any theory of VoiceP must play a crucial role in accounts of voice phenomena (e.g., passive, inversion, middle). Equivalently, if a projection XP plays no role in accounting for voice phenomena, then it should not be called VoiceP.

To guarantee that the Split VoiceP approach does not fail Terminological Assumption, we must justify the following statements.

- (3) Voice plays a role in syntactically identifying a construction as an alternant of the voice proper, i.e., the passive.
- (4) The subject of a passive phrase lands in a non-nominative position or stops over there before reaching the nominative position.

Any approach that distinguishes Voice as a syntactic head from causative and applicative heads fails to justify (3) and (4) because it has to stipulate that Voice may or may not introduce arguments — either +D Voice or –D Voice. I use “–D Voice” as a cover term for a variety of terms including “agent-less Voice”, “Spec-less Voice”, “Voice with unsaturated Spec”, “Voice with [–D]”, “expletive Voice”, “Voice<sub>[PASSIVE]</sub>”, “PASS(ive)” and their equivalents,<sup>2</sup> if any, and use “–D Voice theory” to refer to a theory that postulates a –D voice head.

It is obvious that the motivation behind postulating –D Voice is to reconcile a proposed Voice theory with the “agent-suppression” property of passives. Frankly, the Split VoiceP analysis must be able to explain the agent-suppression property of passives, without postulating –D Voice and without failing Terminological Assumption. Unlike a –D Voice theory, under which the subject of passives has nothing to do with a voice head and the passive semantics syntactically,<sup>3</sup> the Split VoiceP approach holds that the subject of passives (and actives as well) is introduced by (an instance of) +D Voice at some stage of the derivation and passive semantics is produced configurationally where the patient, the passivized subject, occurs higher than the suppressed agent in the

---

<sup>2</sup> All these labels have appeared in the literature (Embick 2004b; Schäfer 2008b; Alexiadou et al. 2015; Bruening 2012; Harley 2013; Legate et al. 2020).

<sup>3</sup> The understood “passive phrase” in this paper is characterized as a phrase with a structure in which there is a suppressed agent, overt or implicit, that occurs lower than a promoted patient/theme, nominative or non-nominative. See §2.1 and footnote 6 for more detailed discussion.

syntactic hierarchy. Proceeding in this direction, we first need to determine whether the subject of passives can reconstruct into a position that is neither NOM (Spec of T) nor ACC (base position).<sup>4</sup> As observed by Legate (2003), in (5) the *wh*-phrase must reconstruct below *every man* in order for *he* to be bound and above *her/she* for the construction to obey Condition C. That (5a) is grammatical while (5b) is not strongly suggests that the bracketed phrase stops over an intermediate position.

- (5) a. [At which of the parties that he<sub>i</sub> invited Mary<sub>j</sub> to] was every man<sub>i</sub> √  
introduced to her<sub>j</sub> \* ?  
b. \*[At which of the parties that he<sub>i</sub> invited Mary<sub>i</sub> to] was she<sub>j</sub> \* ?  
introduced to every man<sub>i</sub> \* ?

Legate (2012: 234), stating that passives may not be derived by phase, suggests that the reconstruction effect may be explained by necessitating internal merge (IM) as targeting every new label. This can be interpreted as a condition for structure extension, which I tentatively formulate below and apply it in the implementation of the Split Voice approach in §4-5.

#### (6) The Structure Extension Condition (SEC)

The structure is extended by introducing arguments.

The reconstruction effect in passives allows us to argue that a passivized subject stops over in an intermediate position before reaching the NOM position. In what follows, it is shown that a passivized subject lands in a non-NOM position, which serves as additional support for Legate’s (2003) argumentation on the reconstruction effect.

In causative-of-passive sentences (“causative-passive” in Tsujimura 1996: 258-259) in Japanese (7), the internal argument (IA) *Hanako* can occur with dative (DAT), not nominative (NOM), indicating that it has moved away from its base-generated position, where it would otherwise be assigned an accusative case (ACC), to a position lower than Spec of T, a NOM position. As indicated by the passive morphemes (PSs) *-are* in (7) and (8), what *Hanako* underwent was passivization. Aoyagi (2021: 99), following Saito (1982: 92), suggests that causative-of-passive sentences are derived from a passive phrase. The passive head to whose Spec IA moves is assumed to be High Appl in Aoyagi (2021: 100).

- (7) Ziroo-ga Hanako-o/ni Taroo-ni sikar-are-sase-ta.  
Ziroo-NOM Hanako-ACC/DAT Taroo-DAT scold-PS-CS-PST  
‘Ziroo made Hanako be scolded by Taroo.’ (Tsujimura 1996: 259)

- (8) Hanako-ga Taroo-ni sikar-are-ta.  
Hanako-NOM Taroo-DAT scold-PS-PST

---

<sup>4</sup> The following abbreviations are used in this paper: 1SG (first-person singular), 3SG (third-person singular), ACC (accusative case/element), AF (affix), AOR (aorist), CS (causative/causative affix), DAT (dative case/element), EA (external argument), EM (external merge), IA (internal argument), IM (internal merge), INS (instrumental case/element), LEX (lexical), LOC (locative), NEG (negative), NOM (nominative case/element), PS (passive/passive affix), PST (past tense), RX (reflexive suffix), Sbj (potential subject), SBJ (surface subject), SYN (syntactic), TOP (topic).

‘Hanako was scolded by Taroo.’

(Tsujiura 1996: 258)

Leaving aside the details of this head, what suffices for the current purpose is that IA can move to a non-NOM position, instantiating accusative-to-dative raising (A-to-D raising),<sup>5</sup> which is embedded under a causative head, spelled out by *-sase*. The raised object, IA, cannot be base-generated in its surface position because its surface position is not thematic.

A-to-D raising, however, cannot represent a sentence since a NOM argument is yet to merge. How a DP is introduced and becomes the NOM argument depends on the head that selects the phrase constructed by A-to-D raising. That head may either be another argument-introducing head or be a tense-aspect related one. In the first case, the derivation extends in the voice domain, leading to a causative (of passive) structure (7), while in the latter, it proceeds in the tense-aspect domain, remaining a passive structure (8). Importantly, passivization is already completed by promoting *Hanako*, as shown in (9), regardless of where it ends up in the surface structure, given that promotion of IA and suppression of EA make up the core property of passivization.

(9) A-to-D raising as passivization:

[TP [XP Hanako<sub>i</sub> [VoiceP Taroo [VP *t<sub>i</sub>* sikar ] ] -are ] ]

A-to-D raising further confirms that passivization involving NOM is not a one-step movement, but rather a successive-cyclic one, as suggested in Legate (2003, 2005). (4) is thus justified by A-to-D raising. The successive-cyclic nature of passivization then requires that there be a non-T head that reintroduces (internally merges) IA. This then necessitates a Voice-over-Voice structure of the voice domain, assuming that an argument-introducing head is Voice. However, to say this much does not suffice to justify (3) for guaranteeing that the Split VoiceP approach does not fail Terminological Assumption in (2). It must be explained that the voice head that reintroduces IA in passivization and the head that introduces EA (Kratzerian Voice) are of the same substance. That is, it must be verified that passives and transitives/actives share the same syntactic property that can contribute to identifying both of them as being derived by a single engine, i.e. Voice. Before examining the syntactic properties of the related constructions, let us have a look at how Voice was characterized and how passivization was derived in the literature in §2.

This paper is structured as follows. Section 2 examines the mainstream approach and the –D Voice approach to passives, reasoning out why they fail to capture voice constructions in a unified way. Section 3 is devoted to explicating the common property of the passive as the voice proper and the causative as a peripheral voice construction based on syntax-morphology mirroring of causatives and passives in Japanese and Mongolian. Section 4 shows how clauses are built under the Split VoiceP model, by applying it to certain constructions. Section 5 compares the Split VoiceP approach with other approaches, thereby demonstrating the merits of the former and eliminating –D Voice (Passive), Applicative and Cause as syntactic heads. Section 6 concludes this

---

<sup>5</sup> “A-to-D raising” of course does not necessarily mean that ACC and DAT are actually assigned; it means that a DP undergoes raising from a position where ACC is assigned normally to another where DAT is assigned.

paper.

## 2. The failure of the mainstream approach and the –D Voice approach to passives

This section shows how the one-step movement approach (Chomsky 1981) and –D Voice approach fail to capture certain fundamental facts about passives and causatives. We first overview the basics of the two approaches in §2.1 and §2.2 and specify why they are a failure in accounting for passivization in §2.3.

### 2.1. Basics of the mainstream approach

Chomsky’s (1981) classic treatment of passives and many followers including Jaeggli (1986) take it that passive morphology and Case-filter play crucial roles in deriving what we call “passives”. That is, passive morphology absorbs the ability of the verb to assign an appropriate case to the object, which therefore moves to Spec of TP, where it receives a nominative case. However, having faced many challenges, the case-driven passivization was replaced by the EPP-triggered one in some sense. Following the Chomskyan tradition, there have been numerous definitions of “passive” in terms of syntax. Among them, the following appears to be the defining property of passive that most, if not all, linguists agree with: There is a suppressed agent, overt or implicit, that occurs lower than a promoted patient that occupies the nominative position, instantiating a configuration as below. The representation in (10a) radically differs from that in (10b), which is presented for actives.

(10) a. Passive configuration under the mainstream approach:<sup>6</sup>

If DP instantiates  $DP(\lambda y (P y, x, t_y))$ , where  $y$ ,  $y$  being nominative, is lower than  $x$ ,  $x$  being non-nominative, in the thematic hierarchy with respect to the predicate P, then P is passive and DP is subject of P.

b. Active (transitive) configuration under the mainstream approach:

If DP instantiates  $DP(\lambda x (P x, y))$ , where  $x$ ,  $x$  being nominative or case-less, is the highest argument in the thematic hierarchy with respect to the predicate P, then P is active and DP is subject of P.

This contrast is illustrated in (11) informally.

(11) a. Active:  $[_{TP} X_{[nom, sbj]} [_{VP=domain\ of\ subject-predicate\ relation} X_{[k-less, sbj]} \dots Y_{[acc, obj]} ]]$

---

<sup>6</sup> In light of A-to-D raising, I restate the passive configuration as follows, the condition “ $y$  being nominative” is removed.

(i) If DP instantiates  $DP(\lambda y (P y, x, t_y))$ , where  $y$  is lower than  $x$ ,  $x$  being non-nominative, in the thematic hierarchy with respect to the predicate P, then P is passive and DP is subject of P.

“Passive semantics” then refers to the meaning interpreted for such a configuration. On a –D Voice view, it would be the case that what solely contributes to producing the passive meaning is the agent-suppressing property of the –D voice head, or a dedicated voice-specifying feature like [passive/non-act] on it. That is, syntax is not truly autonomous in producing the passive meaning on a –D Voice view, which would not specify a configuration such as “(P  $x$ ,  $y$ )”, appealing only to non-nominative  $y$ , for deriving passives.

b. Passive: [TP=domain of subject-predicate relation  $Y$ [nom, subj] [VP  $X$ [-nom, adjunct] ...  $Y$ [k-less, obj] ]]

Crucially, while in transitives, the surface subject is already a subject internal to VP, instantiating the VP-internal subject, in passives it is not, diminishing the VP-internal subject. That is, the subjecthood of the subject of passives is obtained outside the predicate VP, with TP being the subjecthood domain, while it is obtained inside the predicate, with VP being the subjecthood domain. What they share is the case-less status inside VP. Passivization, object-to-subject raising, and the subject-to-subject raising are thus unified under the case-driven movement approach. However, under the labeling approach developed in Chomsky (2013, 2015), which dispensed with case-driven movement, the two types of raising are not subject to unification. Subject-to-subject raising is triggered by deriving a singleton set such as  $\{t_{\text{subj}}, \text{vP}\}$ , while object-to-subject raising is triggered by the requirement of sharing features or satisfying EPP. Thus, there are always asymmetries between passivization and subject-to-subject raising, which is certainly unfavorable as far as unification and simplification of syntactic operations under the Minimalist spirit are concerned.

## 2.2. Basics of the –D Voice approach, compared with the +D Voice approach

Now we briefly overview the core proposals in the literature that are devoted to articulating Voice as a syntactic head and the voice domain. The discussion is focused on the questions as to whether there exists a functional head in syntax that may or may not introduce an argument. That is, we are going to answer whether a head referred to as –D Voice or its equivalents exists in the inventory of functional heads.

Any syntactician engaged in articulating Voice would first refer to Kratzer (1996), who characterizes Voice as a syntactic head that introduces an external argument, that is, +D Voice. Kratzer’s (1996) Voice, roughly the same as Johnson’s (1991)  $\mu$  and Chomsky’s (1995)  $v$ , features the ability to introduce EA. Since passives do not realize an overt agent, quite many studies such as Schäfer (2008b), Bruening (2012), Harley (2013), Alexiadou et al. (2015), and Legate et al. (2020) have postulated –D Voice in various flavors in addition to Kratzer’s (1996) Voice. However, the –D Voice theory is explicitly criticized by Collins (2024), who defends the position that external arguments, whether overt or implicit, are syntactically projected in passives by demonstrating that *by*-phrases in passives are subject to Argument Criterion, which he proposes in parallel with Chomsky’s (1981: 36; 1986: 97) Theta-Criterion.

### (12) Argument Criterion

- a. Each argument is introduced by a single argument-introducing head.
  - b. Each argument-introducing head introduces a single argument.
- (Collins 2024: 8)

One effect of Argument Criterion is that if a head introduces an argument, it must always do because it forces a bijection (one-to-one relation) between an argument-introducing head and an argument (Collins 2024: 9). Given Argument Criterion, it is

predicted that the subject of passives,<sup>7</sup> being an argument, is introduced by an argument-introducing head and that there does not exist –D Voice.

The –D Voice theory holds that suppression of EA by –D Voice contributes to producing passive semantics. That is, passive semantics is assigned by a dedicated agent-less head such as –D Voice, which is endowed with a voice-specifying (passive–designating) feature such as [PASSIVE], [NON-ACTIVE], and so on.<sup>8</sup>

Authors such as Pylkkänen (2008), following the mainstream (one-step movement) approach, differentiate passives and causatives, without appealing to dedicated voice specifying features. They seem to treat passives as the voice proper and treat causatives and applicatives as peripheral voice constructions. Pylkkänen (2008) holds that causatives and passives are distinct, with the former headed by a causative head, Cause, and the latter by T (in the case of regular passives) or by an Applicative head (in the case of indirect passives).

(13) Direct passive:

Heya-ga dorobou-ni-yotte aras-are-ta.  
room-NOM thief-by destroy-PS-PST  
'The room was destroyed (by the thief).' (Pylkkänen 2008: 64)

(14) Indirect passive:

Taroo-ga coochi-ni nak-are-ta.  
Taro-NOM coach-DAT cry-PS-PST  
'Taro's coach cried on him.' (Pylkkänen 2008: 70)  
'Lit. Taro was cried by his coach.'

Pylkkänen (2008) takes Voice to be an EA-introducing head in the Kratzerian sense. A high applied argument is introduced by High Appl, which is a non-voice head in her sense.<sup>9</sup> For her, Cause differs from both Voice and High Appl in that Cause does not introduce a causer argument; what introduces a causer argument is then Voice. However, these are not the essential differences among what are labeled “Voice”, “High Appl” and “Cause”. According to Harley (2013: 53), passives are headed by Voice with [PASSIVE], that is –D, whereas causatives are headed by  $v^0$ . This means that, strictly speaking, passives and causatives are not interconnected voice constructions according to Harley's (2013) view. However, to take a broader typological look at valency alternations, the causative and passive are equally among the members of voice constructions, as discussed by Haspelmath (2022) and many others, and therefore it is predicted that they are derived by the same head Voice in syntax.

In summary, Voice as a syntactic head has been characterized in two ways by previous studies, as summarized below.<sup>10</sup>

<sup>7</sup> Collins (2024), however, not concerned with the subject of passives, is not explicit about how Argument Criterion obtains where the subject of passives is concerned.

<sup>8</sup> Unfortunately, such features, unlike [CASE], [ $\phi$ ] and [FINITE/TENSE], which are all primitive, are not well-motivated; they are at best dedicated.

<sup>9</sup> Pylkkänen (2008) does not take a head introducing an applied argument to be Voice because she, with many others, strictly follows the Chomskyan convention that the subject of passives does not stop over Spec of a head lower than T.

<sup>10</sup> See also Oseki (2017: 13) for a detailed summary.

- (15) a. Voice is an EA-introducing head, that is, +D Voice (Kratzer 1996; Pylkkänen 2008).  
b. Voice may either be +D Voice (in actives) or be –D Voice (in passives) (Schäfer 2008b; Bruening 2012; Harley 2013; Alexiadou et al. 2015; Legate et al. 2020).

### 2.3. The failure of the two approaches

The mainstream approach and the –D voice approach both face the same limitations, one concerning the interplay between voice and Voice,<sup>11</sup> another concerning the VP-internal subject as a conceptual standard.

Firstly, they fail to capture the passive and the causative as interconnected voice constructions in a unified way. In non-theoretical studies such as Haspelmath (2022) and many others, too, the passive and the causative have both been regarded as voice alternants as they are. However, theoretical syntacticians, especially generative linguists seem to have ignored to some extent the interconnection between the passive and the causative as voice alternants. They have mostly treated the two constructions as unconnected constructions that are derived independently. For many of them, the causative is not a voice alternant but the passive is. The major reason for why the passive and the causative have been treated as such should be that causative sentences always install a new argument while passive sentences suppress an argument,<sup>12</sup> which is an apparent difference between them. However, this difference does not qualify as a reason for separating them as unconnected constructions.<sup>13</sup> I demonstrate that they are connected not only pre-theoretically but also theory-internally. As highlighted in the following sections, their interconnection lies in the fact that the passive and the causative as an active voice alternant share the property related to subject-installing. The causative installs a new argument, the causer, as a (potential) subject, and the passive installs an old argument for the second time as a (potential) subject. In direct passives, the old argument as such is an internal argument, in gapped indirect passives such as passive-of-causative sentences, it is an external argument,<sup>14</sup> and in gapless indirect passives such as adversity passives, it is an applied argument.<sup>15</sup> All such arguments are potential subjects upon being installed and become surface nominative subjects. Therefore, the passive and the causative do not differ in that both are installers of (potential) subjects. Their difference lies in the fact that the causative does not suppress a different argument, for example, the agent, by installing a potential subject, while the passive does. Importantly, suppression of an argument should by no means be a reason for regarding the passive and the causative as unconnected voice constructions.

Returning to the common property of the passive and the causative, the job of

---

<sup>11</sup> The terms “voice” and “Voice” are used differently. The former is used in a more general sense when I speak of voice phenomena, voice constructions, voice alternation and voice domain, etc. The latter is used to refer to a syntactic head.

<sup>12</sup> But consider adversity passives in Japanese and many other languages, exemplified in (34) and (35) in section 4, in which a new argument is installed despite the characterization of them as “passives”.

<sup>13</sup> Recent studies such as Cheng (2023) claim that the passive and the causative are connected structurally.

<sup>14</sup> See (26).

<sup>15</sup> See (14).



installing an argument as a potential subject is done by a syntactic head, which functions invariably in the passive and the causative. That syntactic head is a (potential-)subject-introducer, not simply an argument introducer. Arguments do not differ according to the voice type as far as their status as potential subjects is concerned. This is more remarkable given that new arguments are introduced in both the causative and the passive (gapless adversity passive) and both internal and external arguments are introduced for the second time in passives (A-to-D and agent-promoting passives). The syntactic head that is an introducer of such arguments as potential subjects is called Voice.<sup>16</sup> In this sense, all subjects should be Voice-internal subjects (or predicate-internal subjects).<sup>17</sup>

Both the mainstream approach and the –D voice approach not only fail to account for the interconnection between the passive and the causative as described but also sometimes reject to admit their interconnection. That is, they ignore the interplay between voice phenomena in a pre-theoretical context and Voice as a syntactic head that derives voice alternants in a theoretical context.

Secondly, they leave open the question of why the subject is assumed to originate as a potential subject (VP-internal subject) in transitives while it is not in passives,<sup>18</sup> as sketched in §2.1. That is, they diverge from the conceptual standard — the nominative grammatical subject originates as a potential subject internal to VP (in a broader sense) invariably in actives and passives. Therefore, the crux of the matter in dealing with the understudied common properties shared by passives and actives concerning the potential subjecthood is to abandon both the one-step movement tradition and the –D Voice spirit, and to find an alternative way to account for the issues raised in the above discussion.

Some remarks on what “potential subject” means are in order. As already clear in discussion in §2.1, “potential subject” refers to the VP-internal subject traditionally. However, it has a slightly different referent in the Split VoiceP analysis. It refers to a non-nominative DP that enters into a subject-predicate relation with the verb phrase absolved from tense-aspect properties. That DP becomes the grammatical subject (nominative) if there is no other DP higher than it. If there is a higher one, that DP remains non-nominative, failing to become the grammatical subject. It is expected that the grammatical subject originating as a potential subject is universal. A potential subject may be thematic as in transitives and unergatives or non-thematic as in adversity passives. The subject of unaccusatives is base-generated in object position, where it is assigned an internal role, but is merged for the second time (internal merge) to enter

---

<sup>16</sup> Notice that the label “Voice” itself does not matter but what it refers to does.

<sup>17</sup> Internal arguments in their base-generated position are not potential subjects, although they are internal to VoiceP or to the predicate. They are potential subjects only when they are reintroduced by a syntactic head as described.

<sup>18</sup> “Potential subject” here is characterized as a non-nominative DP that enters into a subject-predicate relation with the verb phrase absolved from tense-aspect properties. That DP becomes the grammatical subject (nominative) if there is no other DP higher than it. If there is a higher one, that DP remains non-nominative, failing to become the grammatical subject. It is expected that the grammatical subject originating as a potential subject is universal. A potential subject may be thematic as in transitives and unergatives or non-thematic as in adversity passives. The subject of unaccusatives is base-generated in object position, where it is assigned an internal role, but is merged for the second time (internal merge) to enter into a subject-predicate relation with the verb, with subjecthood assigned to it. In this sense, whether a potential subject is thematic or not and whether it has the external role or an internal one are not relevant.

into a subject-predicate relation with the verb, with subjecthood assigned to it. In this sense, whether a potential subject is thematic or not and whether it has the external role or an internal one do not matter for the licensing of subjecthood. Importantly, following Kratzer (1996), EA is severed from the verb and sits in Spec of VoiceP. Kratzer's (1996) theory is in fact telling us that subjecthood is licensed outside the verb before a proper case is licensed to the relevant argument. With this said, why this is so only in actives (transitives) if we treat passives differently with regards to the licensing of subjecthood. This necessitates a treatment of the subject of passives as a potential subject internal to the predicate but external to the verb,<sup>19</sup> thereby assuring that it enters into a subject-predicate relation with the verb, that it, it is assigned subjecthood, complying with the conceptual standard — the predicate-internal subject, which is called VP-internal subject traditionally.

Notice that Spec of TP is not a position in which to assign subjecthood if the VP-internal hypothesis is correct. This said, the subject-predicate relation must be established before TP is projected and this should be the same for both actives and passives. Quite strikingly, the subjecthood of subjects has not got a lot of attention in the discussion of syntax of passives, with most of the research concerned with the motivation of passivization, role-linking, and structural organization of the passive clause.

### 3. Split VoiceP: Layered affixation of voice-suffixes as evidence

I have pointed out certain limitations of the mainstream (one-step movement) approach and the –D Voice approach. I also presented a line of reasoning in favor of the +D Voice approach in §2.2, which is implemented in §4 and §5. In further defense of this approach, I proceed to present certain supportive facts for the claim that Voice as a syntactic head introduces both EA and IA and derives both passives and causatives as interconnected voice constructions.

Specifically, the syntax-morphology mirroring in passives and causatives in Japanese and Mongolian indicates that the voice domain splits into at least three voice projections, notated as “VoiceP1”, “VoiceP2” and “VoiceP3”, each headed by a distinct instance of Voice that introduces a particular type of argument. Given that affixation as a morphological operation on the surface is realization of a particular syntactic operation, in accordance with Mirror Principle (Baker 1985), certain voice suffixes spell out the voice heads, comprising layered affixation.

(16) [VoiceP3 [VoiceP2 [VoiceP1 [...] -lex1 ] -syn2 ] -syn3 ]

Voice1, corresponds to Kratzerian Voice, which introduces EA, thereby replenishing the valency of the verb. Voice1, being a (semi-)lexical head, is spelled out by a transitivizing morpheme or lexical causative one (notated as “lex1”). Voice2 selects VoiceP1 and VoiceP2 is in turn selected by Voice3. Voice2 and Voice3 are purely

---

<sup>19</sup> The predicate and the verb are not the same notions. The predicate is any phrase that is formed on the verb, which is core of the predicate, and is absolved from tense-aspect properties.

functional heads, spelled out by fully-fledged suffixes notated as syn2 and syn3, respectively.

We proceed to exemplify (16). The layered structure [VoiceP2 [VoiceP1 [...]]], mirroring the string -lex1-syn2, has been widely discussed in the literature including Pylkkänen (2008), Harley (2008, 2013), Wurmbrand and Shimamura (2017), Nie (2020) and many others. The following sentences exemplify -lex1-syn2. In the passive-of-lexical-causative sentences (17) and (19), lex1 is a causative morpheme CS and syn2 is a passive morpheme PS. In causative-of-lexical-causative sentences (18) and (20), lex1 and syn2 are both CS.

- (17) Itgegčid bolon tede-ne ariun sudaruud gal-d šat-aa-gd-v. [Mongolian]  
 followers and their holy books-NOM fire-DAT burn-CS-PS-PST  
 ‘The followers and their holy books were burned in the fire.’  
 (*Alma-in Nom: Alma-in Huu Asan: Ch.4*)
- (18) Bi d üü-geer hubčas-aa hat-aa-lg-san. [Mongolian]  
 1st-NOM younger sister-INS cloth-ACC-RX dry-CS-CS-PST  
 ‘I had my younger sister dried my clothes.’ (Kullman and Tserenpil 2015: 117)
- (19) Ookii biru-ga seihi-ni yotte tat-e-rare-ta. [Japanese]  
 big building-NOM government-by stand-CS-PS-PST  
 ‘A big building was built by the government.’  
 (adapted from Tsujimura 2014: 269)
- (20) Taroo-ga Ziroo-ni sensei-o nak-as-ase-ta. [Japanese]  
 Taro-NOM Ziro-DAT teacher-ACC cry-CS-CS-PST  
 ‘Taro made Ziro make the teacher cry.’ (Pylkkänen 2008: 122)

As seen clearly, lex1 (-aa in Mongolian and -as in Japanese) is the exponent of a lexical-causativizing head, which has sometimes been assumed to be identical to a transitivity head. What is spelled out by syn2 (-gd and -lg in Mongolian and -rare and -ase in Japanese) is either a higher causativizing or a passivizing head.<sup>20</sup>

However, implementation of [VoiceP3 [VoiceP2 [...]]], which is mirrored by -syn2-syn3, has not been explicated. Evidence that the morphemes represented by syn2 and syn3 are syntactic productive affixes follows from the following facts. First, they can be fully-fledged, unlike lex1, which is more compact and sometimes lacks a consonant. Second, they always occur outside transitivity morphemes. Third, they are not compatible with idiosyncratic meanings.<sup>21</sup> The -syn2-syn3 string may be PS<sub>syn2</sub>-CS<sub>syn3</sub> (in Japanese) or CS<sub>syn2</sub>-PS<sub>syn3</sub> (in Japanese and Mongolian). First, we discuss PS<sub>syn2</sub>-CS<sub>syn3</sub>.

PS<sub>syn2</sub>-CS<sub>syn3</sub> instantiates A-to-D raising, which represents passivization, as we observed in §1. Given that passivization is a syntactic operation and lexical-

<sup>20</sup> For discussion of the lexical-syntactic distinction of Japanese causative suffixes, see Shibatani (1976), Jacobsen (1992), Kuroda (1993), Miyagawa (1984, 1998) and Harley (2008), Pylkkänen (2008) among many others. Discussion on Mongolian has been rare, but see Janhunen (2012: 149) and Bai (2024).

<sup>21</sup> This can be syntactically tested out by Harley’s (2008: 3ff) diagnostics.

causativization feeds passivization but not vice versa, PS<sub>syn2</sub> must be attached higher than CS<sub>lex1</sub> and the attachment of CS<sub>syn3</sub> is even higher than that of PS<sub>syn2</sub>. PS<sub>syn2</sub> introduces the passivized IA, e.g., *Hanako* in (22) and CS<sub>syn3</sub> introduces a causer, e.g., *Ziroo* in (22), respectively.

(21) *Hanako-ga Taroo-ni sikar-are-ta.* [Japanese]  
 Hanako-NOM Taro-DAT scold-PS-PST  
 ‘Hanako was scolded by Taro.’ (= 8)

(22) *Ziroo-ga Hanako-o/ni Taroo-ni sikar-are-sase-ta.* [Japanese]  
 Ziro-NOM Hanako-ACC/DAT Taro-DAT scold-PS-CS-PST  
 ‘Ziro made Hanako be scolded by Taro.’ (= 7)

[ Ziroo-ga [ Hanako-o/ni [ Taroo-ni [ Hanako sikar ] -are ] -sase ] ...

Importantly, CS *-sase* in A-to-D raising is different from that in regular transitive-based causatives in that the height of its affixing position varies. In the regular transitive-based causative (23), *-sase*, selecting a transitive phrase, is merged immediately above the transitivizing head (or lexical causative head), while in the causative-of-passive sentence (22), *-sase*, selecting a passive phrase created by A-to-D raising, is merged higher than the transitive-selecting head.

(23) *Ziroo-ga Taroo-ni Hanako-o sikar-ase-ta.* [Japanese]  
 Ziro-NOM Taro-DAT Hanako-ACC scold-CS-PST  
 ‘Ziro made Taro scold Hanako.’

[ Ziroo-ga [ Taroo-ni [ Hanako-o sikar ] -ase ] ...

Notice that in (22), the causee is IA *Hanako* and in (23), it is EA *Taroo*. Since affixation of CS *-sase* follows that of PS *-are* in (22), CS is merged higher in causative-of-passive sentences than in regular causatives. CS in regular causatives and PS in causative-of-passive sentences are affixed at the same height, signaling that causativization and passivization can be derived by the same voice head.

In (24), promotion of the subject *Dorž* is bound with affixation of PS *-gd*. That is, passivization as a syntactic operation is mirrored by affixation of PS as a morphological operation. In (25), which is also a causative-of-passive sentence, the first-person subject *bi* (dropped in this sentence) is base-generated as a causer. The introduction of the causer takes place hand in hand with affixation of CS *-uul*, as a Mirror Principle effect. Notice that in (25), before the causer is introduced, IA *Dorž* has undergone passivization (ACC-to-ACC raising)<sup>22</sup> and precedes EA *aav*. What is important is that in (25), affixation of *-gd* is bound with introduction of *Dorž* and affixation of *-uul*, a causative

<sup>22</sup> The Mongolian counterpart of the Japanese A-to-D raising is accusative-to-accusative raising. However, ACC-to-ACC is also available in Japanese as an alternative to A-to-D raising, which Mongolian lacks.

morpheme,<sup>23</sup> bound with introduction of the causer (dropped in this sentence).

- (24) Dorž aav-d-aa tani-gd-san. [Mongolian]  
 Dorž-NOM father-DAT-RX recognize-PS-PST  
 ‘Dorž was recognized by his father.’

[ Dorž [ aav-D-aa [ Dorž tani ] -gd ] ...

- (25) Dorž-ig aav-d-ni tani-gd-uul-h-gui-in [Mongolian]  
 Dorž-ACC father-DAT-RX recognize-PS-CS-INF-NEG-GEN  
 tuld sahal naa-san.  
 for beard attach-PST  
 ‘In order not to make Dorž recognized by his father, I attached beard to his face.’  
 (Umetani 2006: 95)

[ 1sg [ Dorž-ig [ aav-d-ni [ Dorž tani ] -gd ] -uul ] ...

Because affixation of PS *-gd* precedes that of CS *-uul*, promotion of IA *Dorž* precedes introduction of the causer, the dropped first-person subject, as a Mirror Principle effect. According to the traditional generative view, it would be the case that promotion of *Dorž* is not bound with and is quite later than affixation of *-gd* because patient-promotion is assumed to have nothing to do with a passive marker on the verb. From the perspective of Mirror Principle, however, a passive suffix, at least in agglutinative languages such as Mongolian and Japanese, is the very exponent of the head that induces passivization.<sup>24</sup> The head that induces passivization cannot be a locus of tense, T, although the patient may ultimately occur in Spec of TP in the surface structure.

In sum, in Mongolian and Japanese causative-of-passive sentences, the passive morphemes *-gd* and *-rare* instantiate syn2, bound with introduction of IA, and the causative morphemes *-uul* and *-sase* instantiate syn3, bound with introduction of the causer.

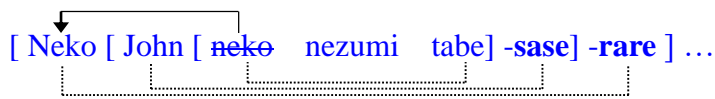
We proceed to discuss Japanese passive-of-causative sentences,<sup>25</sup> in which the order of CS and PS is reversed, with CS being syn2 and PS being syn3. Causative-of-passive sentences are indirect passives in the sense that the subject is not base-generated as IA but rather as EA. What has been puzzling in causative-of-passive sentences is PS<sub>syn3</sub>, which was often undifferentiated from PS<sub>syn2</sub> in direct passive sentences by previous studies. Morphology alone is not helpful in distinguishing them. Instead, we need to look at their syntactic properties instead. Compare the following sentences.

<sup>23</sup> Mongolian causative suffixes include *-uul*, *-lg*, *-aa*, and their allophonic variants.

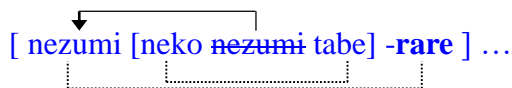
<sup>24</sup> A passive morpheme in agglutinative languages such as Japanese *-rare* and Mongolian *-gd* functions more like the passive head (Voice, strictly speaking) *be* in English than the agent-suppressing suffix *-ed*, although it may be claimed that *-rare* and *-gd* also function as agent-suppressors.

<sup>25</sup> Passive-of-causative sentences had various names including “causative passive” in Tsujimura (1996), who used the same label for our causative-of-passive (22, 25), “indirect causative” in Harley (2013) and “causative-passive” in Agoyagi (2021).

- (26) Neko-ga John-ni nezumi-o tabe-sase-rare-ta. [Japanese]  
 cat-NOM John-DAT mouse-ACC eat-CS-PS-PST  
 ‘The cat was made to eat the mouse by John.’ (Aoyagi 2021: 88)



- (27) Nezumi-ga neko-ni tabe-rare-ta. [Japanese]  
 mouse-NOM cat-DAT eat-PS-PST  
 ‘The mouse was eaten by the cat.’ (Aoyagi 2021: 87)



In both sentences, *nezumi* is the patient. When *nezumi* is promoted to SBJ, PS *-rare* is attached to the root verb. In contrast, when *nezumi* stays in-situ, functioning as an object, *-rare*, in fact, is not attached to the root, as seen in (26). Therefore, in this case *-rare* is not associated with IA *nezumi*, but with EA *neko*. Now it is clear that in (26) affixation of *-rare* takes place hand in hand with reintroduction of *neko* as the subject, as represented by the dotted line. In other words, in (27), PS *-rare* spells out the head that promotes IA *nezumi*, while in (26), it spells out the head that promotes EA *neko*. Importantly, in (26) affixation of *-rare* does not take place before *neko* is promoted because with both *neko* and *nezumi* being in-situ, passivization has not taken place. Once promotion of *neko* takes place, it represents passivization, affixation of *-rare* also takes place. That is, (external) merger of *neko* in its base position precedes affixation of *-rare* in passive-of-causative sentences. This means that syntactic operations such as introduction of agents (*neko*), introduction of causers (*John*) and promotion of agents (*neko*) are mirrored by morphological operations, i.e. spellings of the agent-introducing head, of the causer-introducing head, and of the agent-promoting head, namely, root-realization (*tabe*), affixation of CS (*-sase*), and affixation of PS (*-rare*), respectively.

What is particularly notable is that the height of merger of the passivized argument varies between (26) and (27) and therefore the height of affixation of *-rare* varies accordingly. Affixation of *-rare* in *-sase-rare* (26) is higher than that of root-selecting *-rare* (27). That is, passivization in (26) and that in (27) are *not operations of the same level*. But, instead, causativization in (26) and passivization in (27) are *operations of the same level* in the sense that they take place at the same height, immediately after the completion of transitivity.

In summary, syn2, syn3 and lex1 must be distinguished. Lex1 is the exponent of a semi-lexical and semi-functional head, which introduces EA. Notice that lex1 cannot be PS. Syn2 and syn3, the exponents of purely-functional heads, can either be CS or be PS. Layered affixation of lex, syn2 and syn3, all being voice suffixes, allows us to conclude that the corresponding syntactic operations give rise to a layered structure, which is represented as Split VoiceP.

Notably, as evidenced by causative-of-passive sentences (PS-CS, *-rare-sase* and *-gd-uul*), the passivizing head is able to introduce an argument, given that promotion of

IA from ACC position and suppression of EA make up the core property of the passive. This is crucially important because it yields no difference between the passive, the voice proper, and the causative with regard to their ability to introduce arguments. It is then very natural to conclude that a causativizing/transitivizing head, much like a passivizing head, is Voice. Thus, (3), repeated below as (28), is justified, thereby verifying that the Split VoiceP approach does not fail Terminological Assumption.

(28) Voice plays a role in syntactically identifying a construction as an alternant of the voice proper, i.e. the passive.

#### 4. Applying Split VoiceP

Two points can be summarized from the discussion in the preceding section. First, VoiceP is a split projection similar to CP, which splits into separate projections such as FocP and TopP.<sup>26</sup> Second, the sentence subject, notated as SBJ, of sentences including both passives and causatives/transitives is (re)introduced beforehand by one of the split voice heads to get assigned subjecthood. That is, SBJ originates as a potential subject, notated as Sbj, which is assigned subjecthood in position lower than the NOM position. From these, it follows that clauses, which represent a subject-predicate configuration, are built by introducing arguments as Sbj and promoting a last-introduced one to SBJ, the NOM subject, with others, if any, suppressed or demoted to non-NOM elements. In this sense, VoiceP, rather than TP, is the very projection representing a subject-predicate relation invariably in passives and actives.

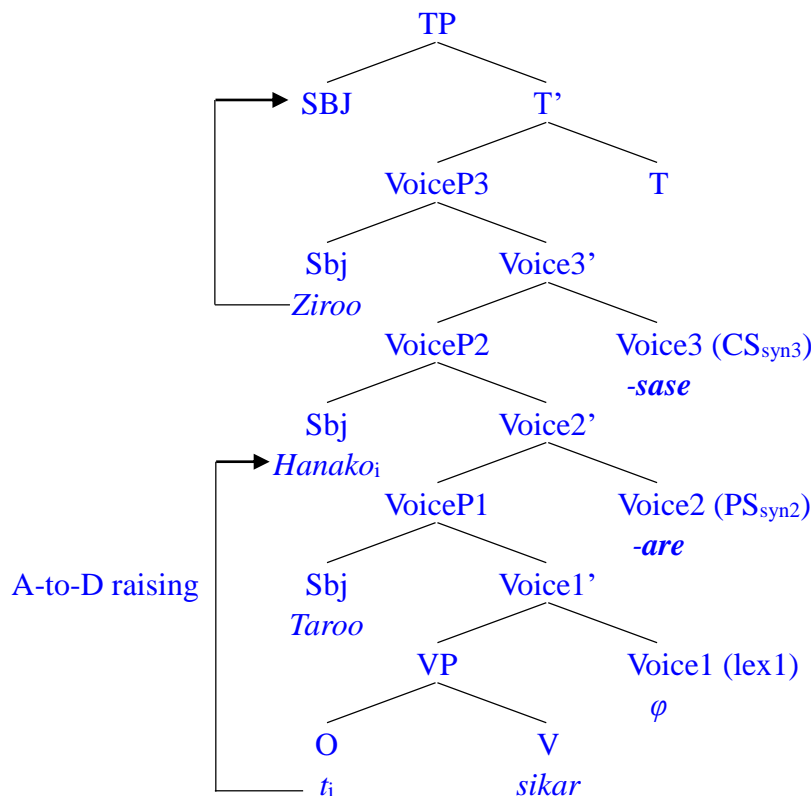
##### 4.1 Applying to causative-of-passive and passive-of-causative sentences

To represent the Split VoiceP structure of causative-of-passive sentences with a tree diagram:

---

<sup>26</sup> The Split VoiceP analysis being pursued here is quite similar to Sigurðsson and Wood's (2020) analysis named "stacked VoiceP analysis", which is proposed for capturing Icelandic indirect causatives, and Nie's (2020) Voice-over-Voice analysis, which is proposed for capturing morphological causatives. Sigurðsson and Wood (2020) and Nie (2020) argue that two VoicePs are "stacked" to form a voice-over-voice structure in which the causer is introduced by the higher voice head and the causee, overt or implicit, is introduced by the lower one.

(29) Structure of causative-of-passive sentence (22):



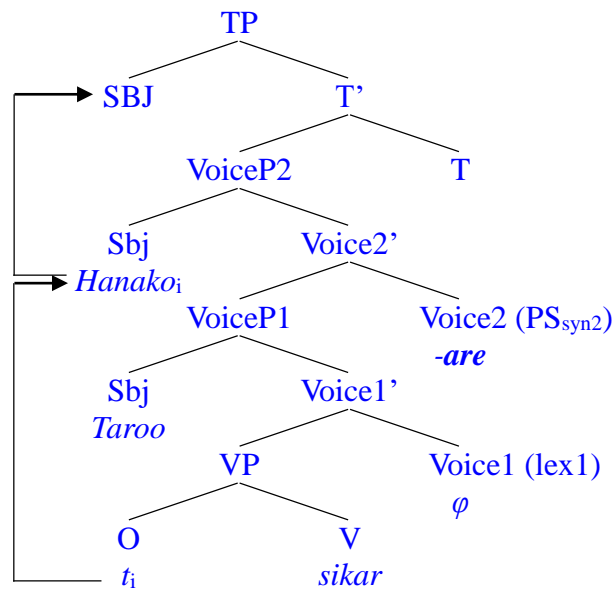
Three arguments, *Taroo* (EA), *Hanako* (IA), and *Ziroo* (causer) are (re)introduced as Sbj by Voice1, Voice2, and Voice3, respectively. A-to-D raising of *Hanako* represents passivization. Importantly, the last-merged Sbj *Ziroo*, being the highest argument, is promoted to SBJ in Spec of TP, where it is assigned NOM. *Taroo* and *Hanako*, which are also Sbj, end up with non-NOM since they are not the highest Sbj in the voice domain. The case to assign to a non-last-merged Sbj depends on the case marking system of a given language. In Japanese, DAT is assigned while in Mongolian, ACC is assigned.<sup>27</sup> DAT is assigned to a demoted (or suppressed) EA in both Japanese and Mongolian.

However, if VoiceP2 does not extend to VoiceP3, *Hanako* will be the last-merged Sbj, which is promoted to SBJ, with the sentence ending up as a passive-of-transitive/lexical causative structure, as shown below. The same holds true of *Taroo* if VoiceP1 is not selected by Voice2, with the sentence ending up as a transitive/lexical causative structure.

<sup>27</sup> ACC is also assigned alternatively in Japanese. The differential case-marking in this case has to do with the semantics of causatives such as affectedness, volition, animacy, and so on. The same is true of non-passivized causees. In Japanese, again DAT and ACC alternate for non-passivized causees while in Mongolian, INS and ACC alternate. Which of the two alternants to use depends on the parametric variety of causative semantics (Dixon 2000).

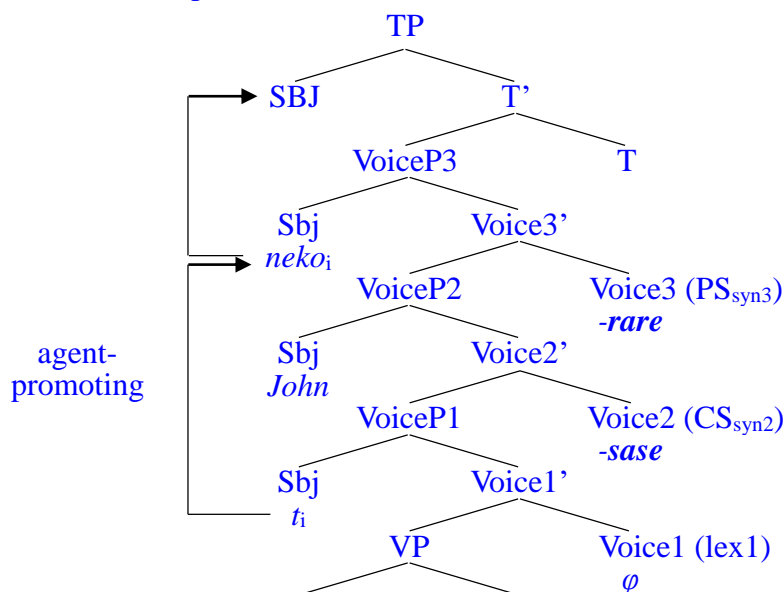


(30) Structure of passive sentence (21):



Notice that Voice is not endowed with dedicated features such as [PASSIVE], [ACTIVE] and [CAUSE], contra the previous studies, which have often postulated these features for capturing the distinction between the voice constructions. Thus, none of the split heads Voice1, Voice2, and Voice3 is predetermined as passive Voice or active Voice.<sup>28</sup> Causative and passive semantics is determined configurationally. Therefore, Voice2 and Voice3 may either be passivizing heads or be causativizing heads; they are spelled out either by CS or by PS. When Voice2 introduces a new DP as Sbj via EM, causative semantics is produced, and when VoiceP2 is selected by Voice3 that introduces a DP as Sbj via IM, passive-of-causative semantics is produced.

(31) Structure of passive-of-causative sentence (26):



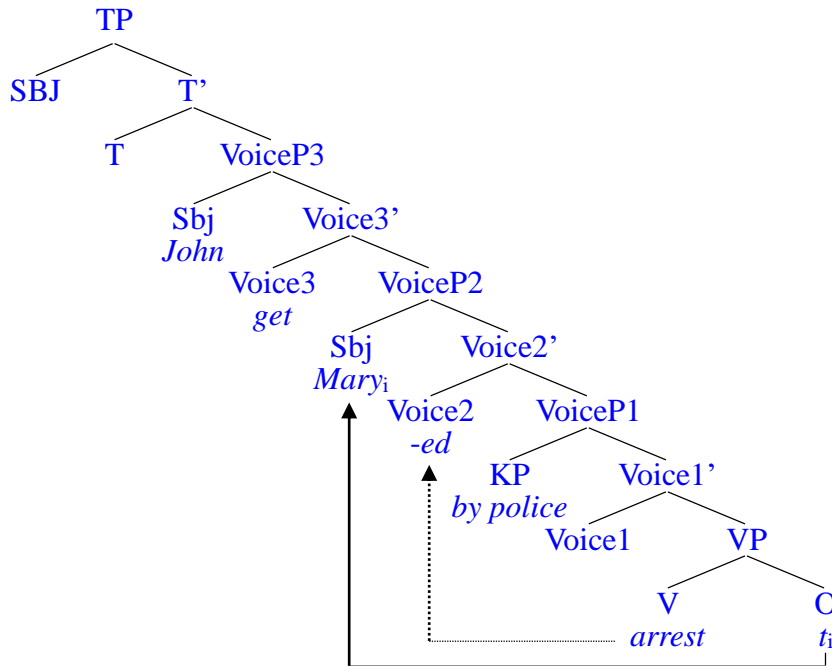
<sup>28</sup> Sigurðsson and Wood (2020), who also propose a layered VoiceP analysis, concludes that the embedded VoiceP, with a highly underspecified argument in its Spec, is not exactly an active or a passive VoiceP. Since they are not concerned a passive that embeds VoiceP, Sigurðsson and Wood (2020) seem to characterize the higher VoiceP as an active VoiceP.

O                      V  
*nezumi*                      *tabe*

## 4.2 Applying to *get*-passives and *be*-passives

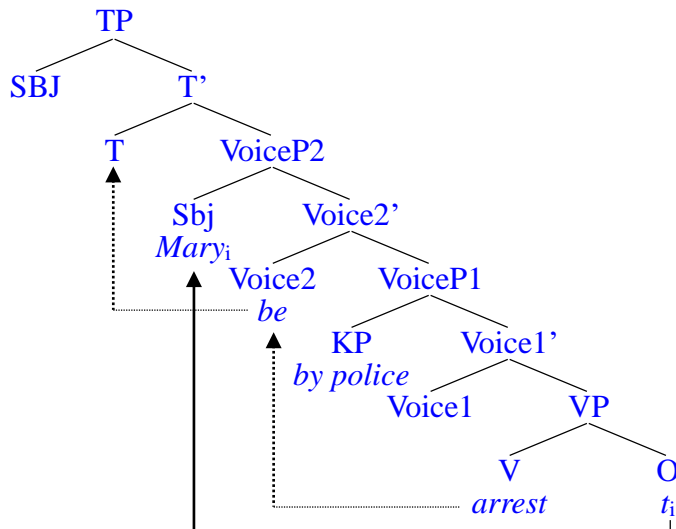
The Split VoiceP analysis can also apply to English *get*-passives, which are a different type of causative-of-passive sentence, where *get* spells out Voice3 and the morpheme *-ed* Voice2. *By*-phrases constitute KP, base-generated as an (implicit) argument in Spec of the EA-introducing head (Collins 2024), VoiceP1. More details are given in §5.4.

(32) Structure of *get*-passive (e.g., *John got Mary arrested by police.*):



The same analysis applies to *have*-causatives, e.g., *I just had my hair cut today.* Unlike *get*-passives, *be*-passives do not involve projection of VoiceP3, with *be* base-generated in Voice2 and undergoing head-movement to T.

(33) Structure of *be*-passive (e.g., *Mary was arrested by police.*):



### 4.3 Applying to adversity passives

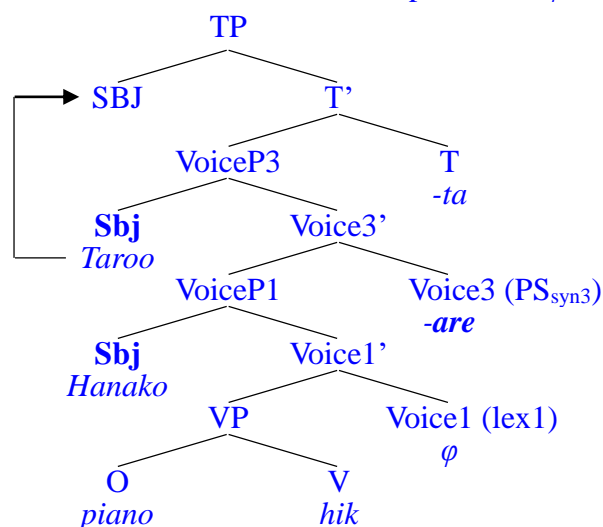
Passivization in adversity/gapless passives, whether with transitive input (34) or with intransitive input (35) mirrors affixation of the single suffix PS<sub>syn3</sub>. It cannot be PS<sub>syn2</sub> because affixation of PS<sub>syn2</sub> mirrors raising of IA, i.e. direct passivization. Morphologically, PS *-rare* is adjacent to the verb (*hik* and *nak*, for example) but syntactically, it is not. That is, while PS<sub>syn2</sub> spells out a vP-selecting head, then PS<sub>syn3</sub> spells out a head that selects the phrase of the vP-selecting head.

- (34) Taroo-ga Hanako-ni piano-o asa-made hik-are-ta.  
 Taro-NOM Hanako-DAT piano-ACC until morning play-PS-PST  
 ‘Taro was affected by Hanako’s playing the piano until the morning.’  
 ‘Lit. Taro was played the piano by Hanako until the morning.’  
 (Tsuji-mura 1996: 238)

- (35) Taroo-ga coochi-ni nak-are-ta.  
 Taro-NOM coach-DAT cry-PS-PST  
 ‘Taro’s coach cried on him.’ (Pylkkänen 2008: 70)  
 ‘Lit. Taro was cried by his coach.’

- (36) Structure of adversity passive (34):

[TP Taroo<sub>i</sub>-ga [VoiceP3 t<sub>i</sub> [VoiceP1 Hanako-ni [VP piano hik]  $\varnothing$ ] -are] -ta]



### 4.4 Subjecthood Condition and derivation of voice alternation

As noted in §2.3 and in the beginning of this section (§4), the sentence subject SBJ originates as a potential subject Sbj, assigned subjecthood. This property is formulated as “Subjecthood Condition”.

- (37) Subjecthood Condition

A nominative subject is merged beforehand to be assigned subjecthood.

In connection with Subjecthood Condition, I propose the following denotation for

VoiceP, following Kratzer (1996).

(38)  $\lambda e [DP(x)(e) \ \& \ V(y)(e)]$

(38), however, excludes the case in which Voice introduces IA. I then propose (39) for VoiceP involving introduction of IA.

(39)  $\lambda e [DP(y)(e) \ \& \ V(y)(e)]$

The unspecification of the role for DP comes as a reflection of Free Merge, where nothing constraints merger of a DP. The role of DP in (38) and (39) does not affect the nature of Voice. This means that the so-called external argument and internal argument are undistinguished upon being introduced by Voice. Strictly speaking, there do not exist external arguments under a Voice theory, as noted by Kratzer (1996: 131ff), with “external argument/EA” used for the purpose of notation. As regards internal arguments, they are not to be done away with because they are base-generated truly in VP. However, their status as internal arguments is not sensed by Voice. Voice is not designated to specify what role it introduces.

In light of Subjecthood Condition, I rewrite (39) as (40), where DP is not identified as the same role as *y* or not but as a potential object. That is, (40) is the denotation of a subject-predicate relation. In this sense, Voice is not simply an argument introducer, but rather an introducer of a subject.

(40)  $\lambda e [DP(sbj)(e) \ \& \ V(y)(e)]$

As the builder of a subject-predicate relation, Voice extracts any DP out of VP that want to become a subject and reinstalls it into the machinery denoted by (40).

To put the point another way, the property of EA as a potential subject is “deprived of” by what precedes it, namely the promoted IA and therefore remains a non-subject element. At this stage of derivation, passivization is already completed. That is, the promoted IA receives a passive interpretation and the predicate is passive-marked.<sup>29</sup> In this sense, EA is suppressed (or demoted) not because that Voice does not introduce EA as assumed by the –D Voice approach but because EA is deprived of its subjecthood. Therefore, EA is now not a predicate-internal (VP-internal) subject. Subjecthood transmitted to it, the promoted IA is now a predicate-internal subject, that is, a Voice-internal subject. If no other DP is introduced higher than the promoted IA, then subjecthood sits in IA, which is therefore chosen for SBJ and moves to Spec of TP. If a different DP is introduced, for example, as in causative-of-passive sentences, subjecthood further shifts upward and that DP becomes SBJ. In this sense, clause formation is creation of the subject-predicate relation at different steps. Stated another way, clause formation is structure extension by introducing arguments as potential subjects.

---

<sup>29</sup> That is, the relevant head (Voice2) is inserted a passive morphology.

In what follows, we examine how Subjecthood Condition serves to extend the structure out of VP. Assume that VP consists of V and IA (DP1). Given Subjecthood Condition, for a clause (subject-predicate structure) to be derived, Voice1, selecting VP, introduces some DP as Sbj, either via EM or via IM. If EM applies, Sbj is EA and a two-place predicate is derived, as in (42). If IM applies, Sbj is IA and a one-place predicate, as in (43).

(41) Basic structure:

[VoiceP1 Sbj [VP DP1 ]]

(42) External merge of DP2 as Sbj:

[VoiceP1 DP2 [VP DP1 ]] (transitive; lexical causative)

(43) Internal merge of DP1 as Sbj:

[VoiceP1 DP1<sub>i</sub> [VP t<sub>i</sub> ]] (unaccusative; (dispositional) middle; fake passive)

There are three reasons to assume that IA is base-generated as the object and promoted to Spec of VoiceP1 as Sbj. First, unaccusatives are derived by (weak) phase and their subject is realized as SBJ in a successive-cyclic way, stopping over the edge of a (weak) phase head before reaching Spec of T (Legate 2003). If a head selecting the root is a phase head, then Voice1 is a (weak) phase head, Spec of which is an intermediate landing site for IA. Second, under Free Merge (Chomsky 2015), it is expected that an argument-introducing head does not specify whether the argument it introduces is EA or IA. That is, nothing prevents IA from merging into Spec of Voice1. Third, Voice1, being a semi-lexical head, is able to replenish the lexical semantics of the verb by (re)introducing an argument. Some more details are given below.

The lexical semantic type of verb is not restricted to transitive, unergative and unaccusative. There are quite many types of non-unaccusative intransitive verb whose subjects originate as IA. For example, Lekakou (2005) and many others have observed a core property shared by middles across languages. Middles promote IA to the subject via movement (or, IM) and EA unprojected syntactically.<sup>30</sup> It is therefore reasonable to say that middle verbs are non-unaccusative intransitive verbs.<sup>31</sup>

For another case, we take a look at Mongolian “fake” passives like *sana-gd*, which is morphologically passive but semantically active, denoting a spontaneous situation. *Sana-gd*, though morphologically passive, does not mean “to be felt/thought/missed” but “to seem”. Importantly, the subject *nutag* is interpreted as a theme (more than as a patient) and assigned NOM, much like the case of unaccusatives.

(44) Nad            nutag-mini            sana-gd-laa.            [Mongolian]  
       1st-DAT    homeland-NOM-RX    miss-PS-PST  
       ‘I got homesick.’

<sup>30</sup> This is evidenced by the fact that (dispositional) middles in English do not allow modification by agentive adverbials, control into purpose clauses, and a *by*-phrase (Alexiadou 2014: 21).

<sup>31</sup> Given that middles are either active-, passive- or reflexive-marked according to the morphosyntactic parameters of the particular languages (Alexiadou 2014: 21ff), the head that (re)introduces an argument as the subject of middles is not designated for specifying the argument type and the voice type.

‘Lit. To/By me, my homeland is missed.’

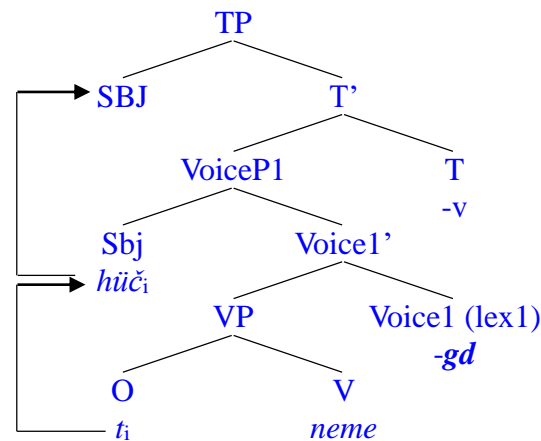
There are many verbs like *sana-gd* in Mongolian. They have been lexicalized as intransitives, where morpheme *-gd* fails to function as a genuine passive marker. The dative element, e.g., *nad* does not have the property of EA. Since SBJ, e.g., *nutag* is IA, it is base-generated as O within VP. However, since raising of *nutag* does not represent passivization, *-gd* does not spell out Voice2 or Voice3, but spells out Voice1. That is, affixation of *-gd* takes place hand in hand with introduction of IA *nutag* as Sbj in Spec of Voice1. The semantics produced in this case is neither causative/transitive nor passive one, but rather semantics associated with spontaneity.<sup>32</sup>

For another example, *neme-gd* means “increase (spontaneously)”, not “be added”. The dative argument is by no means interpreted as an agent, but as a recipient/beneficiary-like argument (Bai 2023).<sup>33</sup>

- (45) Nad      hüč                      neme-gd-v.                      [Mongolian]  
 1st-DAT strength-NOM add-PS-PST  
 ‘My strength increased.’  
 ‘Lit. To me, strength was added.’

On the mainstream view, IA in passives and unaccusatives is promoted from O to SBJ at one step, without skipping over Spec of an intermediate head. However, the syntactic and semantic properties of fake passives point to the fact that promotion of IA is not a one-step movement. IA must stop over Spec of Voice1 before reaching NOM position.

- (46) Structure of fake passive sentence (45):



It is not difficult to see from the above discussion that fake passives and lexical causatives, both of which are formed in VoiceP1, represent voice constructions that alternate at the lowest voice layer in the voice domain. Fake passives involve IA-

<sup>32</sup> Such fake passives do not differ from genuine passives regarding the fact that IA is promoted. They, however, differ in that what EA undergoes is demotion in fake passives while it is suppression in genuine passives. Demotion is downgrading of the theta role of a given nominal, which may be accompanied with downgrading of case, while suppression is downgrading of the case of a nominal, without downgrading of its theta role. To put it using Haspelmath’s (2022) terminologies, fake passives are a role-removing construction as opposed to causatives, which are a role-installing one.

<sup>33</sup> Bai (2023) observes that *neme-gd* was used with an agentive phrase in Middle Mongolian and underwent a diachronic change from an agent-taking verb to a recipient/beneficiary-taking one.

promotion while lexical causatives involve EA-installation.<sup>34</sup> This makes them akin to genuine passives and syntactic causatives, which involve IA-promotion and EA-installation, respectively. Genuine passives and syntactic causatives represent constructions that alternate at a higher voice layer in the voice domain.

Subjecthood condition applies equally in extending the structure to VoiceP2 and VoiceP3, thereby deriving the various voice constructions diagrammed in §4.1-4.3, which are thus interconnected via Voice.

## 5. Eliminating Passive (–D Voice), Applicative and Causative Heads

This section compares the Split VoiceP approach to other approaches, pointing out the limitations of the latter. By doing so, it is further suggested that –D Voice/Passive, Cause (causative head), and (high) Applicative as syntactic heads should be eliminated.<sup>35</sup> My attention is narrowed to the configurational nature, rather than the aspectual nature such as existential binding, of passives and others. Four subsections are allotted to discuss the issue.

Before continuing, let us look at the correspondence between the terminological labels used in this paper and those used in the literature.

Table 1. Comparisons and correspondences among labels for argument-introducers

	indirect passive	syntactic causative	canonical passive	lexical causative	transitive
this paper	Voice3	Voice2		Voice1	
the mainstream				v	
Collins (2005)			<i>Voice</i>	v	
Pylkkänen (2008)	HighAppl	Voice		Voice	
Kratzer (1996)				Voice	
Bruening (2012)			Passive+ <i>Voice</i>	Voice	
Alexiadou et al. (2015)			<i>Voice</i>	Voice <sub>causer</sub>	Voice <sub>agent/holder</sub>
Harley (2013)		Voice		v	Voice

The italic *Voice* refers to a voice head that is assumed to not introduce EA by the relevant author. Not all authors are explicit about every argument-introducer. For some authors, *Voice*, introducing higher or lower causer arguments, is a recursive head. The head introducing the lower causer and the head introducing EA may either be bundled or split.<sup>36</sup>

### 5.1 Split VoiceP vs. differential VoiceP (Legate et al. 2020)

The question of why genuine passives of passive are not allowed by UG has been one of the tough questions for theories dealing with passives.

<sup>34</sup> Notice that fake passives do not qualify what is named “passive” given that passivization is characterized as promoting IA to a position *higher than* where EA is merged.

<sup>35</sup> See Nie (2020) for further reasons to eliminate a causative head.

<sup>36</sup> Either option is attested according to Pylkkänen (2008), while for Harley (2012), the two heads must be split. In this paper, Voice1 represents a bundled head, leaving aside the details of its further splitting property.

(47) \*Nuzumi-ga John-ni neko-ni tabe-rare-rare-ta.  
mouse-NOM John-DAT cat-DAT eat-PS-PS-PST

(48) \*John was was killed. (Ljutikova and Tatevosov 2015: 48)

The mainstream approach, in which passives are derived from an active base by one-step movement of the patient, cannot rule out passivization of passives. To explain the absence of passives of passive, Legate et al. (2020) propose the differential VoiceP analysis whereby actives and passives are subtypes of Voice and built independently using different lexical items. That is, passives involve alternative syntactic structure building. Legate et al. (2020) argue that the difference between passives and actives lies in the difference in satisfying the external role introduced in Spec of Voice. In actives, the external role is satisfied by a DP. In contrast, in passives, it is satisfied either by a *by*-phrase, which is right-adjoined to VoiceP, or by existential binding on Voice. That is, Voice<sub>PASS</sub> introduces the external role but does not introduce an external role bearer, the agent DP, while Voice<sub>ACT</sub> introduces the agent. Therefore, strictly speaking, according to Legate et al. (2020), passivization is not movement, that is, it is not an operation induced by a syntactic rule. Specifically, suppression of the agent makes up the core property of passivization, regardless of promotion of the patient to a higher position.<sup>37</sup>

If passives of passive were attested, under Legate et al. (2020) analysis, the agent would be expected to be suppressed on the first round of passivization and subsequently the second round of passivization would also require an agent, the most prominent argument in the syntactic hierarchy, to be suppressed. However, there are no other agents available for suppression. The patient is the most prominent argument after the completion of the first round of passivization but suppression of it does not count as passivization,<sup>38</sup> thereby ruling out a passive-of-passive.

Legate et al.'s (2020) differential VoiceP analysis is akin to the assumption that Voice may be either +D or –D. It is therefore a featurally based one, with passives and actives differing in the selection of +D or –D. Legate et al. (2020) assume that Voice<sub>PASS</sub> introduces the external role but does not syntactically project EA in its Spec. This is to say that an implicit argument is not necessarily present in syntax, contra Collins (2005, 2024). It therefore seems that Legate et al.'s (2020) model of building passives using lexical items different than those used in building actives may only capture a subtype of passives that differ from actives in their lexical base, remaining subconstruction-specific, but neither passive-specific nor universal.

Indeed, the facts, especially those from Sanskrit, reported by Legate et al. (2020), point to the fact that the apparent passive suffix may not function as a genuine passive marker but as a non-active or unaccusative marker, as noted by Legate et al. (2020: 800ff), much like the fake passive in Mongolian, as discussed in §4.4. Leaving aside the historical development of such suffixes, Legate et al. (2020: 800ff) do not make

<sup>37</sup> I use the verb “suppress” to denote what Legate et al. (2020) denotes using “demote”, for the reason given in note 20.

<sup>38</sup> In Legate et al.'s (2020: 803, 806) words, passivization (suppression) of the agent cannot make the theme available for a second round of passivization.



explicit the properties of their lexicalization into non-passive morphemes. The Sanskrit examples of the purported passive-of-passive (characterized as unaccusative by Legate et al. 2020: 802) and the Turkish examples (characterized as impersonal-of-passive by Legate et al. 2020: 784) indicate that the passive morphology has been lexicalized as part of the verb stem, with no indication of passive meaning.<sup>39</sup>

- (49) a. *muc-ya-te* [Sanskrit]  
 escape-PS<sup>40</sup>-3SG  
 ‘gets free, escapes’ or ‘is freed’  
 b. *pac-ya-te* [Sanskrit]  
 cook-PS-3SG  
 ‘ripens, softens, cooks’ or ‘is cooked’ (Legate et al. 2020: 800)

- (50) Harp-te IMP<sub>i</sub> [PRO<sub>i</sub> kimse tarafından vur-**ul-un**-mak] iste-**n**-mez.[Turkish]  
 war-LOC [PRO anyone by shoot-PS-PS-inf want-PS-NEG-AOR  
 ‘In war, one does not want to be shot by anyone.’  
 (Özkaragöz 1986: 77, cited from Legate et al. 2020: 772)

All such examples show that the action denoted by the verb affixed by the relevant PS lacks an external cause or agentivity, indicating that the morphologically passive verb is not semantically and syntactically passive. The lack of an external cause or agentivity means that EA is unprojected in syntax, unlike the case of passives, e.g., *The mutton was cooked (by John)*, in which EA is projected in syntax as an implicit argument (Collins 2005, 2024). This means that the non-active PS (as in Sanskrit) and the impersonal PS (as in Turkish), much like the fake PS in Mongolian, are not genuine passive markers but rather lexical-property-related morphemes.<sup>41</sup> If it spelled out a head introducing the external role, PS would be a genuine passive marker and the verb would be a passive-of-transitive rather than a non-active or impersonal.

Unfortunately, Legate et al.’s (2020) analysis fails to differentiate between the “fake-passive” property and the implicit-argument property of PS. For example, the Sanskrit verb *pac-ya-te* in (49b), which means either “ripen” or “be cooked”, will have a single syntactic structure according to Legate et al.’s (2020) analysis. Therefore, Legate et al.’s (2020) claim to derive passivization based on different lexical items from transitives remains too strong. Notice that in English/German-type languages, passives are formed on the basis of a fully-fledged transitive structure that already contains EA. e.g., *cook* (Alexiadou et al. 2015: 123), reflecting the original insights of Chomsky (1957, 1981). Accordingly, an analysis in favor of treating passivization as a rule-based

<sup>39</sup> Notice that *iste-n* ‘want-PS’ in (50) is translated into “want” rather than “be wanted” and arguably the same is true of the embedded PS *-ul* in *vur-ul-un* ‘shoot-PS-PS’, with the embedding PS *-un* as a genuine passive marker. In this is so, (50) should be analyzed as an example of passive-of-impersonal, not impersonal-of-passive.

<sup>40</sup> Legate et al. (2020) labeled this morpheme “non-active”.

<sup>41</sup> Under the Split VoiceP approach, they spell out Voice1, Spec of which is open for the patient, but not for the agent. Notice that in transitives and lexical causatives, Voice1 introduces the agent. Recall our earlier discussion that Voice does not specify the type of arguments and therefore there does not exist a dedicated agent-introducing head in the strict sense. Kratzer’s (1996) agent-introducing Voice is revised as Sbj-introducing Voice under the Split VoiceP approach in this paper. See §4 for details.

operation is preferred to Legate et al.’s (2020) analysis, whereby passivization is treated as a result of suppression of the agent, which is attributed to a feature-based distinction of passives and actives. Strictly speaking, such an analysis, separating promotion of the patient from the core property of passives, treats passivization as a lexical operation rather than a syntactic operation. Thus, the absence of passives of passive still remains an open question.

The proposed Split VoiceP analysis, by contrast, accounts for the absence of passives of passive and the Sanskrit and Turkish facts reviewed above (along with the Mongolian facts) more straightforwardly. Notice that the Split VoiceP property of the voice domain requires the structure be extended by introducing arguments, that is, by realizing arguments in separate projections. This is formulated as SEC in (6), repeated below.

(51) **The Structure Extension Condition**

The structure is extended by introducing arguments.

Moving the patient out of Spec of Voice and realizing it in Spec of another Voice right above the first Voice would yield a structure violating SEC because movement of the patient without skipping over a different DP would not target a new label, given that separate voice heads are split out of the single engine, Voice.

(52) \*<sub>[VoiceP3 DP1<sub>i</sub> [VoiceP2 t<sub>i</sub> ...</sub>

This is exactly what Kratzer’s (1996: 132) Realization Principle predicts.

(53) **The Realization Principle**

Arguments of a head must be realized within the projection of that head.

SEC and RP require a structure such as (54), in which DP2 interrupts between the two occurrences of DP1.<sup>42</sup>

(54) <sub>[VoiceP4 DP1<sub>i</sub> [VoiceP3 DP2 [VoiceP2 t<sub>i</sub> ...</sub>

The newly installed DP2 serves as a causer, which functions as an external-role bearer. DP2 is then accessible for suppression on a second round of passivization. When suppression of DP2 takes place hand in hand with further promotion of DP1, a second round of passivization takes place. However, the yielded structure does not qualify as passive-of-passive, but rather as passive-of-causative-of-passive.<sup>43</sup>

If we apply the differential VoiceP analysis (Legate et al. 2020) and the mainstream analysis (Chomsky 1981) to passives, many facts related with the morphology-syntax mirroring properties of alternating voice constructions across agglutinative languages such as Mongolian and Japanese (§3-4) in addition to the absence of passives of passive

<sup>42</sup> This is to say that IM does not apply successively. What precedes or follows IM must be EM.

<sup>43</sup> If passive-of-causative-of-passive is attested, it will be something like the hypothesized sentence *The fox was made (by the hunter) to be eaten by the wolf*. Indeed, passive-of-causative-of-passive is attested in Karachay-Balkar, a Turkish language (Lyutikova and Tatevosov 2015), where causativization is called for to mediate between two applications of the passive. The Split VoiceP property does not constrain the recursiveness of the causative-passive tie, whose occurring in unlimited rounds should be possible in principle. However, further recursion would not be attested due to a ‘morphological bottle-neck’ (Kiparsky 2013: 24).

and the ambiguous property of certain passive morphemes will remain unexplained.

## 5.2 Split VoiceP vs. Split vP (Harley 2013)

Harley (2013) proposed a split vP analysis of causative-passive interactions. According to Harley's (2013) analysis, the syntactic inflectional domain consists of a single Voice<sub>[PASSIVE]</sub> projection, whereas the lexical domain, vP, is so rich that it is split into four distinct projections, including a verbalizing head *v* (lower), an EA-introducing head Voice (lower), a causative head *v* (higher) and an AA-introducing head Appl, when necessary.

(55) [VoiceP [AppIP [vP [VoiceP [vP [RP ... ]]]]]]  
}  
split vP

What is crucial for this analysis is the distinction between the lower Voice, a –D head, and the higher vP, which introduces EA and assigns causative semantics. There, however, at least three problems with this analysis. First, the structure in (55) fails to capture a succession of two syntactic heads, as mirrored by distinct morphemes. Applying (55) to the passive-of-causative sentences (Harley's 2013 indirect passive) in (56) and (57), the syntactic CS *-sase* would be distributed into the lexical domain, spelling out the lower Voice, which is taken to be a lexical head in Harley (2013), as shown in (58) and (59).

(56) Hiroko-ga pizza-o tabe-sase-rare-ta. [Japanese]  
 Hiroko-NOM pizza-ACC eat-CS-PS-PST  
 'Hiroko was made to eat the pizza.' (= 26)

(57) Ken-ga (Naomi-ni) John-o tazune-sase-rare-ta. [Japanese]  
 Ken-NOM Naomi-DAT John-ACC visit-CS-PS-PST  
 'Ken was caused (by Naomi) to visit John.' (Aoyagi 2021: 99)

(58) [TP [VoiceP [vP φ [VoiceP Hiroko [vP [RP Pizza ]]]]]]  
 (59) [TP [VoiceP [vP Naomi [VoiceP Ken [vP [RP John ]]]]]]  
}  
Lexical thematic domain

Under the Split VoiceP analysis, as opposed to Harley's split vP, what introduces EA (*Hiroko* and *Ken*) is Voice1 and what introduces the causer (*Naomi*) is Voice2. What makes the Split vP and the Split VoiceP analyses radically different is that in the split vP analysis, the higher Voice, being designed as Passive, is –D, producing passive semantics, whereas in the Split VoiceP analysis, the higher Voice, i.e. Voice3 is +D, to whose Spec EA is promoted, producing passive semantics.<sup>44</sup>

<sup>44</sup> What the split vP and Split VoiceP structures have in common is the derivation mirrored by the ordering of the *-sase-rare* string.

- (60) [TP [VoiceP3 [VoiceP2 φ [VoiceP1 Hiroko [VP [RP Pizza ]]]]]
- (61) [TP [VoiceP3 [VoiceP2 Naomi [VoiceP1 Ken [VP [RP Pizza ]]]]]
- Lexical thematic domain

The second problem with Harley (2013) split vP analysis is that it fails to correctly predict the mirroring between morphology and semantics associated with the *-sase-rare* string. It incorrectly associates the lower syntactic morpheme *-sase* to lexical, not syntactic, causative semantics.

Last, Harley (2013), incorrectly concludes that the causer (what would otherwise be in Spec of the causative *-sase*) is truly syntactically absent. Harley (2013) takes *-sase* (the head of the higher vP) to be the exponent of a –D head. Unfortunately, as evidenced by (57), the causative *-sase* (at least in passive-of-causative sentences) can occur with the causer (*Naomi*); therefore, it is the exponent of a +D head, not a –D head.

Importantly, these problems all disappear if we apply the proposed Split VoiceP analysis to passive-of-causative sentences such as (56) and (57). Therefore, the split projection is VoiceP, but not vP.

### 5.3 Split VoiceP vs. ApplicativeP (Pylkkänen 2008)

In Pylkkänen’s (2008) applicative theory, no –D head is explicitly postulated. Voice is taken to be the same as Kratzer’s (1996) Voice, corresponding to Voice1 here. Pylkkänen (2008) follows the mainstream in assuming that promotion of an argument to Spec of T marks the completion of passivization. The applicative theory, however, is not preferred where my data and Collins’ (2024) reasoning are concerned.

To add a fact concerning the flaw of the applicative framework, it is somewhat messy, to my understanding, about the ordering and height of the argument-introducing heads. According to Pylkkänen (2008: 74), the passive morpheme of canonical passives is higher than that of indirect adversity passives. But this is compatible with the fact that an applied argument in adversity passives, which is introduced by High Appl, is merged later than all arguments introduced by other heads. This, violating Mirror Principle, remains a puzzle, as also mentioned by Pylkkänen (2008) in her note 18 on page 142.

More importantly, with the applicative framework, passive-of-causative sentences are ruled out. As we saw earlier, passive-of-causative sentences must involve movement of the agent from the EA position to a higher position. In the applicative framework, the EA position is Spec of VoiceP, but Voice is already higher than High Appl. Even if we move EA from Spec of VoiceP directly to Spec of T, a correct word/morpheme order will not be derived. Unfortunately, there is no discussion on passive-of-causative sentences throughout the monograph. The closest structure I can find in the monograph is the following, which is labeled “Passivized causative” by Pylkkänen.

- (62) \*Taroo-ga Hanako-ni-yotte musuko-ni sin-ase-rare-ta. [Japanese]  
 Taroo-NOM Hanako-by son-DAT die-CS-PS-PST  
 ‘The son was caused to die by Hanako.’ (Pylkkänen 2008: 93)

In this sense, *sin-ase* is ambiguous between lexical (*let-to-die* reading) and syntactic causatives (*make-die* reading). Therefore, it is not clear how the applicative framework is applied to this sentence. Another difficulty in applying this framework to passive-of-causative sentences such as (57) and (62) is that when the causer, the DAT element (*Naomi-ni* in (57)), is present, it finds no position where it can merge.

It is then suggested here that (High) Applicative heads, failing Terminological Assumption, should be eliminated. Notice that High Appl and Voice are similar elements in that an applied argument that High Appl introduces and an external argument that Voice introduces are both external to the verb, as also noted by Pylkkänen (2008: 138-139).

#### 5.4 Split VoiceP vs. projecting complement (Wood and Marantz 2017)

Wood and Marantz (2017) argue that there is a type of argument introducer, notated as “*i\**”, whose syntactic properties are derived from its structural environment: it projects the syntactic category of its complement and it assigns the thematic role implied by its complement (2017: 269). That is, Voice, Applicative and little *v* (corresponding to Voice1 here) among the list they offer are distinct “uses” of the same syntactic head. Wood and Marantz’s (2017) main concern is the unification of prepositions and little *v* with respect to their argument-introducing property. They divided the unified argument introducer as a syntactic head into five distinct types.

- (63)a. voice introduces the external arguments of vPs (often agents)
- b. low appl introduces an argument related to a DP
- c. little *p* introduces the external arguments of PPs (figures)
- d. prepositions introduce non-core arguments in a manner syntactically distinct from high appl
- e. high appl introduces a non-core argument

Obviously, Wood and Marantz (2017) do not treat the head that reintroduces IA in passives as an argument-introducer arguably because they do not take passivization as a successive-cyclic movement. They explicitly state that the argument introducer, their *i\**, can be expletive (Wood and Marantz 2017: 269), that is, –D. It remains unclear in their work whether this –D claim is related to passivization, however. Importantly, IA is not included in their list of arguments. Accordingly, the IA-reintroducing head (Voice2 and/or Voice3 in this paper) would not be an argument-introducer on their view. This arbitrariness of a syntactic head characterized as an argument-introducer appears to be a limitation of this analysis. In contrast, the Split VoiceP analysis can capture the fact that IA, the subject of passives, is reintroduced by an argument-introducer.

#### 5.5 Split VoiceP vs. Smuggling (Collins 2005, 2024)

Admittedly, the smuggling approach (Collins 2005, 2024) plays an important role in deriving the word order in passives of English-type languages. I benefit in developing

the Split VoiceP approach from some of Collins' (2024) assumptions and insights including Argument Criterion and the KP account of *by*-phrase, which were advocated in favor of the smuggling approach. However, because the smuggling approach fails to account for the absence of passives of passive, I do not adopt it. An additional reason for this is related with Collins' (2024) characterization of Voice.<sup>45</sup> Collins (2024) refutes Kratzer (1996) together with those who advocate –D Voice, because he thinks that Kratzer's (1996) Voice fails Terminological Assumption. Collins' (2024) refutation of Kratzer (1996) comes in defense of Collins' (2005) smuggling analysis that forces the assumption that Voice is present as a “smuggler”, attracting VP/PartP as a whole into its Spec, with EA merged in Spec of v. Unfortunately, Collins' (2024) separation of Voice, which is in fact treated as –D Voice by him, from its EA-introducing ability is not appropriate. One reason for this is that if EA were introduced by v, a head lower than Voice, the subject of passives, IA, would have to also be reintroduced by a higher v head (recall the discussion in §1), forcing a split vP analysis. However, reintroducing IA through a v head would be incompatible with the fact that the passive is the voice proper, the defining property of which must be related to a voice head, rather than to v. Moreover, if EA were introduced by v, EA would not be an argument that is severed from the verb because Voice, in Kratzer's (1996) theory, is the lowest head that can sever EA from the verb.

The smuggling approach seems to be motivated for the sake of tackling the problem raised by the mismatching between the argumenthood of the *by*-phrase and its adjunct-like positioning. This problem can be resolved neatly by assuming that a *by*-phrase as KP merges in Spec of Voice1 and successive head movement takes place, as shown informally in (61).<sup>46</sup> In (61), the cluster of the auxiliary and the participle is the result of a successive head movement, which is an in-syntax or a post-syntax operation. The exact position of the participial affix, *-ed*, may also be the head of a participial projection. Alternatively, it may occupy Voice2, the head of the passive voice projection.

(64) [TP [VoiceP2 this book<sub>i</sub> BE [VoiceP1 KP (by John) -ED [VP read t<sub>i</sub> ]]]

IA is internally merged in Spec of Voice2, followed by its movement to Spec of T.<sup>47</sup>

## 6 Conclusion

This paper has discussed how Voice should be characterized in the Minimalist syntax and how the voice domain should be articulated. I have argued for the claim that arguments are introduced as potential subjects by Voice, which is the only argument-introducer in the universal inventory of functional elements that derives voice constructions including passives, causative, applicatives, and middles, echoing

<sup>45</sup> See also Lima Júnior and Augusto (2015) for discussion of drawbacks of the smuggling approach.

<sup>46</sup> See also (32).

<sup>47</sup> An alternative solution is that KP merges in the right Spec of Voice1 like an adjunct does. If this is the case, the following needs to be rationalized, an issue out of scope for the purposes of this paper.

(i) In English type languages, EA is merged in the left Spec of a relevant head when it is a DP and it is merged in the right Spec when it is a KP

Kratzer's (1996: 120) statement that Voice is truly at the heart of a theory of voice. Passive and causative semantics is assigned by the Voice-over-Voice configuration of Split VoiceP. No dedicated heads such as Passive, Voice<sub>[passive]</sub>, Voice<sub>[active]</sub>, Cause, and (High) Applicative are necessary, nor are dedicated voice-specifying features.<sup>48</sup> From this, it follows that argument introducing can come down to Free Merge; (Re)merger of an argument, external or internal, is unconstrained; no more than is necessary in syntactically accounting for voice phenomena.

The working principles for the Split VoiceP analysis include Argument Criterion, formulated by Collins (2024), the Structure Extension Condition, formulated in §1, and Subjecthood Condition, formulated in §4.4. All these require that Voice, if it is designed as a truly argument-introducer as it is, it must introduce some argument, overt or implicit. Therefore, echoing again Kratzer's (1996: 132) statement that all we seem to need is a very general principle for the syntactic realization of *all* arguments, I take the Split VoiceP approach as being developed to be that very general principle.

The empirical motivation and evidence for the Split VoiceP analysis came primarily from Japanese and Mongolian. Unification of voice constructions and simplification of syntactic operations for deriving them without appealing to non-primitive features comprise the theoretical motivation for the analysis. The extension of this analysis, if it is proven to be on the right track, to other languages and the collection of additional empirical supports from them remain to be undertaken in separate papers. I have not focused myself on relevant issues such as phase, labeling, and case-assignment in connection with VoiceP. The details of the internal organization of the lexical thematic domain (VoiceP1/vP) and a number of periphrastic voice constructions were also not discussed. Separate papers will address these issues.

## References

- Aoyagi, Hiroshi. 2021. On the causative and passive morphology in Japanese and Korean. *Open Linguistics* 7: 87-110.
- Alexiadou, Artemis. 2014. Active, middle, and passive: the morpho-syntax of Voice, *Catalan Journal of Linguistics* 13: 19-40.
- Alexiadou, Artemis, Elena Anagnostopoulou and Florian Schäfer. 2015. *External Arguments in Transitivity Alternations: A Layering Approach*. Oxford: Oxford University Press.
- Alexiadou, Artemis, Elena Anagnostopoulou, Florian Schäfer. 2018. Passive. In *Syntactic Structures after 60 Years*. eds. N. Hornstein et al., 403-425. Berlin: Mouton de Gruyter.
- Bai, Chigchi. 2023. A diachronic change in Mongolian passives viewed from "nemegde-". *Mongolian Language* 540: 27-31.
- Bai, Chigchi. 2024. Morphology, syntax, and semantics of Mongolian causatives: An interface analysis. To appear in *Acta Mongolica* 22.

---

<sup>48</sup> For previous discussion in favor of eliminating such features from argument-introducing heads, see Wood and Marantz (2017).

- Baker, Marc. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373-415.
- Bruening, Benjamin. 2012. By-phrases in passives and nominal. *Syntax* 16: 1-41.
- Cheng, Gong. 2023. How are causatives and passives related? Paper presented at *The 10th International Conference on Formal Linguistics* held at Tsinghua University, Beijing.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague, The Netherlands: Mouton de Gruyter.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York, NY: Praeger.
- Chomsky, Noam. 1995. *The Minimalist Program*. 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 Honour of Adriana Belletti*, eds. Elisa Di Domenico, Cornelia Hamann and Simona Matteini, 3-16. Amsterdam: John Benjamins.
- Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81-120.
- Collins, Chris. 2024. *Principles of Argument Structure: A Merge-Based Approach*. Cambridge, MA.: MIT Press.
- Dixon, R.M.W. 2000. A typology of causatives: Form, syntax and meaning. In *Changing Valency: Case Studies in Transitivity*. eds. Dixon, R.M.W., & Alexandra Y. Aikhenvald, 30-83. Cambridge, MA.: The MIT Press.
- Embick, David. 2004b. On the structure of resultative participles in English. *Linguistic Inquiry* 35: 355-92.
- Harley, Heidi. 2008. On the causative construction. In *Handbook of Japanese Linguistics*. eds. Miyagawa, S. and Saito, M., 20-53. Oxford: Oxford University Press.
- Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125: 34-57.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In *Causatives and Transitivity*. eds. B. Comrie and M. Polinsky, 87-120. Amsterdam: John Benjamins.
- Haspelmath, Martin. 2022. Valency and voice constructions. [lingbuzz/005941](https://lingbuzz/005941)
- Jaeggli, Osvaldo. 1986. Passive. *Linguistic Inquiry* 17(4): 587-622.
- Jacobsen, Wesley. 1992. *The transitive structure of events in Japanese*. Tokyo: Kurosio.
- Janhunen, Juha A. 2012. *Mongolian*. Amsterdam/Philadelphia: John Benjamins.
- Johnson, Kyle. 1991. Object Positions. *Natural Language and Linguistic Theory* 9: 577-636.
- Kiparsky, Paul. 2013. Towards a null theory of the passive. *Lingua* 125: 7-33.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*. eds. J. Rooryck and L. Zaring, 109-137. Dordrecht: Kluwer.
- Kuroda, Shigeyuki. 1993. Lexical and productive causatives in Japanese: An



- examination of the theory of paradigmatic structure. *Journal of Japanese Linguistics* 15: 1-81.
- Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34: 506-515.
- Legate, Julie Anne. 2005. Phases and Cyclic Agreement. In *Perspectives on Phases*. eds. M. McGinnis and N. Richards, 147-156. Cambridge, MA.: MITWPL.
- Legate, Julie A. 2012. The Size of Phases. In *Phases: Developing the Framework*, eds. Ángel J. Gallego, 233-250. Boston, MA.: Mouton de Gruyter.
- Legate Julie A. Faruk Akkuş, Milena Šereikaitė, Don Ringe. 2020. On passives of passives. *Language* 96: 771-818.
- Lekakou, Maria. 2005. *In the Middle, Somehow Elevated*. UCL, doctoral dissertation.
- Lima Júnior and Augusto. 2015. Is ‘smuggling’ really necessary? The most recent analyses of passive sentences reconsidered in terms of phasehood and cyclic movement. *ReVEL, edição especial, n.9*: 62-92.
- Lyutikova, Ekaterina and Sergei, Tatevosov. 2015. Recursive passivization: a causative coercion account. In *Affix Ordering Across Languages and Frameworks*. eds. Manova, Stela, 17-52. Oxford: Oxford University Press.
- Miyagawa, Shigeru. 1984. Blocking and Japanese causatives. *Lingua* 64: 177-207.
- Miyagawa, Shigeru. 1998. (S)ase as an elsewhere causative and the syntactic nature of words. *Journal of Japanese Linguistics* 16: 67-110.
- Nie, Yining. 2020. Morphological causatives are voice over voice. *Word Structure* 13: 102-126.
- Oseki, Yohei. 2017. Voice morphology in Japanese argument structures. *New York University Qualifying Paper*.
- Özkaragöz, İnci. 1986. Monoclausal double passives in Turkish. *Studies in Turkish Linguistics*. eds. Dan Isaac Slobin and Karl Zimmer, 77-91. Amsterdam: John Benjamins.
- Pylkkänen, Liina. 2008. *Introducing Arguments*. Cambridge, MA.: MIT Press.
- Saito, Mamoru. 1982. *Case Marking in Japanese*. Unpublished manuscript. Cambridge, MA.: MIT Press.
- Schäfer, Florian. 2008b. *The Syntax of (Anti-)Causatives: External Arguments in Change-of-state Contexts*. Amsterdam: John Benjamins.
- Shibatani, Masayoshi. 1976. The grammar of causative constructions: A conspectus. In *Syntax and Semantics: The Grammar of Causative Constructions*. eds. Masayoshi Shibatani, 1-39. London: Academic Press.
- Sigurðsson, Einar Freyr and Jim Wood. 2020. On the implicit argument of Icelandic indirect causatives. *Linguistic Inquiry* 52, 579-625.
- Tsujimura, Natsuko. 1996. *An Introduction to Japanese Linguistics*. Malden, MA.: Blackwell.
- Tsujimura, Natsuko. 2014. *An Introduction to Japanese Linguistics (Third Edition)*. Malden, MA.: Wiley-Blackwell.
- Umetani, Hiroyuki. 2006. Serial use of the passive and causative suffixes in Mongolian. In *Gengo kenkyuu no shyatei*. eds. Kato Shigehiro and Yoshida Hiromi, 83-102. Tokyo: Sanbi Insatu.

- Wood, Jim and Alec Marantz 2017. The interpretation of external arguments. In *The Verbal Domain*. eds. RD'Alessandro, I Franco, Á Gallego, 255-278. Oxford: Oxford University Press.
- Wurmbrand, Susi and Koji Shimamura. 2017. The features of the voice domain: Actives, passives, and restructuring. In *The Verbal Domain*. eds. Roberta D'Alessandro, Irene Franco and Ángel J. Gallego, 179-204. Oxford: Oxford University Press.