A Simpler Syntax of Voice: Preliminaries based on Japanese Data

Chigchi Bai
School of Foreign Languages, IMUFE
chigchi.hohhot@qq.com
January 08, 2025

Abstract
This paper proposes to simplify the clause building mode by introducing arguments as potential subjects (sbj) at various heights through Voice heads, which constitute Split VoiceP. The gist of the Split VoiceP analysis to propose is that Voice is the only argument-introducer (specifically, sbj-introducer) in the inventory of functional heads available by UG and that the lexicon does not specify a particular type of Voice head. Being simply a sbj-introducer, Voice has nothing to do with the thematic role of DP it introduces as sbj. Any of so-called external, internal, or (high) applied arguments can be introduced by Voice, under Free Merge. Multiple arguments can be introduced as different sbjs by different Voice heads. A last-merged sbj is promoted to the nominative subject (SBJ), with others remaining non-nominative elements. The variation of the grammatical voice arises from different types and heights of merge in the Voice doiman, not from voice-specifying features on Voice heads in the lexicon, and thus they are unified and interconnected by Voice. Clauses are thus built by introducing arguments as sbjs. Clause building is ultimately minimizable to the simplest syntactic operation Merge.

Keywords: Voice, passive, causative, argument, subject, merge, Japanese

1. Introduction

Simplifying clause building mode has been one of the primary goals of MP. Clauses are built by means of merge, which is induced by functional heads. At least three core functional categories are invented by UG, namely, *v*, T, and C (cf. Chomsky 2001, Richards 2007). They, taking care of the lexical/thematic domain, the temporal/inflectional domain and the force/information domain, respectively, make up the "spine" of a sentence. Each of *v*P, TP and CP splits into various sub-projections, giving rise to a cartography (Rizzi 1997; Cinque 2006), in a broader sense.

(1) Universal mode: CP > TP > vP

(2) TopP-FocP > TP-AspP > vP-VP

It has mostly been assumed that the subject-predicate relation is realized in the middle domain, TP, where agreement is also realized. However, ever since the Predicate-

internal subject hypothesis (PISH),[1] it has been agreed that subjects are already subjects in the thematic domain, vP, but are restricted to active sentences (including transitive and unergatives).

Unfortunately, few studies have explicitly suspected the asymmetry between active subjects and passive subjects imposed by the classic PISH.

(3) Why does PISH hold in actives but not in (what we have called) passives?

One would answer this question by saying "because passive verbs are (inherently) intransitives, with no external argument (EA), just like unaccusatives". This answer, which is certainly not satisfactory, leads to another question as follows.

(4) Why is only EA taken as an inherently subject?

Note that EA is not necessarily a sentence subject and a sentence subject is not necessarily EA. The answers that immediately come up in mind are as follows.

(5) Answer 1: EA is a top-ranked argument of V.

(6) Answer 2: EA is outside of V/VP.

Answer 1, essentially semantic/thematic-based, states that EA is assigned the highest theta-role by V, and hence its subjecthood. This answer, however, is conceptually questionable. It fails to account for why not all EAs, being assigned the highest theta-role universally, surface as sentence subjects (SBJ), for example, in canonical passives. Subjecthood is not a semantic primitive. Answer 2, essentially configuration-based, does not appeal to the argument structure of the verb. This would entail that any DP outside VP may surface as SBJ, since semantic relations are not at play here. However, a DP outside VP may or may not surface as SBJ, for example, in syntactic causatives, as exemplified below.

(7) Ziroo-ga Taroo-ni Hanako-o sikar-ase-ta.[2]
Ziro-NOM Taro–DAT Hanako-ACC scold-CS-PST
'Ziro made Taro scold Hanako.'

Taroo, EA of the verb, is base-generated outside VP. However, it is not SBJ. Thus, the following entailemnts do not obtain.

[1] Traditionally called VP-internal subject hypothesis.

[2] The following abbreviations are used in this paper: 1SG (first-person singular), 3SG (third-person singular), ACC (accusative case/element), AF (affix), CS (causative/causative affix), DAT (dative case/element), EA (external argument), EM (external merge), IA (internal argument), IM (internal merge), INF (infinitive), INS (instrumental case/element), LEX (lexical), NEG (negative), NOM (nominative case/element), PS (passive/passive affix), PSS (possessive suffix), PST (past tense), sbj (potential subject), SBJ (surface subject), SYN (syntactic), TOP (topic).

(8) “DP is EA” entails “DP is SBJ”.

(9) “DP is SBJ” entails “DP is EA”.

Accordingly, EA is not a premise for deriving SBJ and not in all cases SBJ results from EA. Where does PISH play a role then? The answer is obtained by observing that the classic PISH is not about SBJ, but about EA, or that it is, at best, about SBJ of active transitive clauses. The classic PISH says nothing about passives, middles, and unaccusatives, in which SBJ is not EA. Thus, the classic PISH remains construction-specific, not being global.

If the classic PISH is not about SBJ, is there any universal with respect to subjecthood, constraining on what can be SBJ? This question leads to another question: What is a subject? There are too many definitions of “subject”, but no consensus seems available. What remains a universal is the subject-predicate relation. But, what is “subject” when we speak of the subject-predicate relation?

Given that the classic PISH fails, a thematic-based definition of subjects remains invalid. That is, the “agent=SBJ” account does not hold. This results in invoking the configuration-based definition of subjects. The premise for this is that the so-called EA, which is a potential subject (sbj), not necessarily SBJ, is outside VP. Thus, it is expected that configuration-based subjects, not constrained by the argument structure of V, hold in passives, middles and unaccusatives as well as in active transitives. Consequently, it is the syntactic position of a DP that qualifies it as sbj. Therefore, subjecthood (the property of being a potential subject) arises in the configuration such as [$_{XP}$ DP [$_{VP}$]], regardless of where DP is base-generated and what role it bears.

Let us spell out the asymmetry between active and passive subjects, which are imposed by the classic PISH, in another way. The classic PISH entails the following.

(10) For transitives, subjecthood arises in Spec-VoiceP/vP, where sbj promotes to SBJ; that is, sbj→SBJ. *v* assigns subjecthood of a DP or a subject-predicate relation involved in a clause.

(11) For passives, subjecthood arises in Spec-TP, where a non-sbj (e.g., an object) element promotes to SBJ; that is, obj→SBJ. T assigns subjecthood of a DP or a subject-predicate relation involved in a clause.

How can there be such an asymmetry? If we want to maintain any underlying constraint on subjecthood to be global, we must unify the sbj-assigning head as either T or v/Voice, but not both. Two analyses are considered.

(12) Alternative 1: All subject-predicate relations are assigned by T.

(13) Alternative 2: All subject-predicate relations are assigned by *v*.

For alternative 1, PISH would be eliminated entirely and EA would turn out to be a

non-sbj. For generativists, this is unfavorable. For Alternative 2, passive subjects along with active subjects are predicate-internal, occupying a position where they acquire subjecthood. Alternative 2 wins for the following reasons. As shown by the causative-of-passive sentence in Japanese, the embedded subject *Hanako* has moved from its VP-internal position to a position higher than EA but lower than SBJ. Importantly, the promotion of the embedded subject already instantiates passivization. Firstly, the promotion comes with the passive morpheme *-are*, which spells out the head that introduces the embedded subject into its Spec via internal merge (IM); secondly, the promotion produces a passive meaning. This means that the embedded subject, as a passive subject, occupies a position outside VP. The same holds true in Mongolian (15).

(14) Causative-of-passive in Japanese:
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)

(15) Causative-of-passive in Mongolian:
Dorž-ig aav-d-ni tani-gd-uul-h-gui-in
Dorž-ACC father-DAT-PSS 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)

That position is certainly Spec of a functional head above VP. Let us call it Voice, complying with Kratzer (1996). Proceeding in this direction, we need to elaborate on Voice and VoiceP.

This paper is organized as follows. Section 2 overviews relevant approaches to Voice as a syntactic head, pointing out the drawbacks of the various –D Voice proposals. Section 3 presents Japanese data supporting the claim that the voice domain consists of at least three separate Voice heads, which make up Split VoiceP. Section 4 argues for Subjechood Condition, which states that nominative subjects must be assigned subjecthood beforehand within the predicate, and derives various voice constructions on the basis of the Split VoiceP structure and Subjecthood Condition. Section 5 shows how the lowest Voice head plays a role in deriving unaccusatives, middles, and fake passives. Section 6 concludes this paper.

2. Voice as a syntactic head in the literature

We have two concerns here: To examine how Voice is characterized as syntactic head in the current generative theory and to explicate the difficulties that the –D Voice approach faces. Voice as a syntactic head has been characterized in two ways (+D Voice

and –D Voice) by previous studies, as summarized below.[3]

(16) Core properties of two types of Voice

+D Voice only	Both (differential Voice)		–D Voice only
	+D Voices	–D Voice	
Introduce EA	Introduce EA	Introduce EA-role/Variable	Introduce nothing
Kratzer 1996	Pylkkänen 2008; Schäfer 2008b; Bruening 2012; Harley 2013; Alexiadou et. al. 2015; Legate et al. 2020		Collins 2024

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) μ 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 (or defective/expletive Voice) in various flavors in addition to Kratzer's (1996) Voice.

In what follows, we compare specific proposals about Voice available in the literature. Let us first use the following examples to illustrate primary voice alternants.

(17) John-wa Mary-ni uta-o ut-ase-ta. (syntactic caustive)
John-TOP Mary-DAT song-ACC sing-CS-PST
'John made Mary sing a song.'

(18) John-wa keisatu-ni tukam-are-ta. (syntactic passive)
John-TOP police-DAT seize-PS-PST
'John was seized by the police.'

(19) John-wa kuruma-o kowa-si-ta. (lexical caustive)
John-TOP car-ACC break-CS-PST
'John broke the car.'

(20) Keisatu-wa John-o tukam-da. (transitive)
police-TOP John-ACC seize-PST
'The police seized John.'

As has been assumed, lexical causatives and transitives are derived via the same level of head, often notated as v or (Kratzerian) Voice, and syntactic causatives are derived by a higher head, which is sometimes identified as Voice. For syntactic passives, views differ, some hold that passives are derived via Voice at the same level of active/causative Voice, and others hold that they are derived via Voice, which has no active variant. The

[3] See also Oseki (2017: 13) for a detailed summary.

following table summarizes the hierarchical position of the various voice heads.

(21) Comparisons and correspondences among labels for argument-introducers[4]

	High level		**Low level**	
	syntactic causative	syntactic passive	lexical causative	transitive
the mainstream			v	
Kratzer (1996)			Voice[+D]	
Collins (2005)		Voice	v	
Pylkkänen (2008)	Voice[+D]	Voice[–D]	Voice[+D]	
Bruening (2012)		Voice[–D]	Voice[+D]	
Alexiadou et al. (2015)		Voice[–D]	$\text{Voice}_{\text{causer}}$[+D]	$\text{Voice}_{\text{agent/holder}}$[+D]
Harley (2013)	Voice[+D]	Voice[–D]	v	Voice[–D]
Legate et al. (2020)		Voice[–D]	Voice[+D]	

This table suffices to show the overuse of Voice, an undesirable result, which was led to by the postulation of –D Voice. The –D Voice theory holds the following views. The active-passive distinction is attributed to the difference in selectional ability between heads, which have ultimately been assumed to be distinctive features of Voice. If a Voice head has the ability to select for a DP as its argument, then it is often said to be Voice[+D] or Voice[act], and if a voice head lacks this ability, then it is said to be Voice[–D], Voice[passive/non-act]. There have, therefore, often been assumed to exist two opposite variants of Voice, one for actives (including transitives/casuatives) and another for passives, middles, and sometimes unaccusatives. In such an approach, UG provides different types of Voice heads in the lexicon and therefore various voice types such as active and passive are determined by Voice heads. In other words, semantics and/or interpretations of voice alternants (passive meaning, active meaning, etc.) are predetermined in the lexicon.[5] This allows one to say that, for example, a sentence is a passive sentence because it is derived via Voice with a passive feature. This would amount to saying that a sentence is a passive sentence because its voice is passive, a sentence is a middle sentence because its voice is middle, and a sentence is an active sentence because its voice is active. However, one, whether generativist or not, would not be impressed by this (see Haider 2024a: 14). An important consequence of the –D Voice theory is that the subject of passives has to do with neither –D Voice, the passive head postulated by it, nor the passive semantics syntactically.[6]

[4] 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. Either option is attested according to Pylkkänen (2008), while for Harley (2013), the two heads must be split. In this paper, Voice1 represents a bundled head, leaving aside the details of its further splitting property.

[5] Alexiadou et al. (2018: 419) states, "The lexicon can provide two variants of the semi-functional Voice-head, which differ in their syntax and semantics.

[6] In this paper, a DP is said to be a passive subject if it satisfies the following.
(i) If DP instantiates DP(λy (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-

It seems reasonable to attribute a head's ability of introducing arguments to its selectional feature for DP. However, it does not seem reasonable to assume that such a feature is arbitrary. Bruening (2012), one of the most well-worked out analyses, claimed that Voice has [S:N], a DP-selecting feature, which may or may not be checked off by a DP. If [S:N] is checked off by a DP, then Voice is an active Voice. If [S:N] is not checked off by a DP, then it, Bruening (2012) argues, is checked off by a head selecting Voice with [S:N]. That head is Passive, which has an existential binding property, contributing passive semantics, according to Bruening (2012).

Likewise, Pylkkänen (2008), Alexiadou et al. (2015), and Legate et al. (2020) argued that Voice introduces an external argument variable or role but lacks a DP as an external argument in syntax. Harley (2013: 37ff) also assumes differential Voice, without elaborating on related features. According to her, external arguments are introduced by Voice. She however, implies that Voice may introduce nothing but is spelled out by a passive morpheme. This is to say that Voice she assumes is a differential Voice.

The gist of these analyses is that there is some head like the so-called Passive or passive Voice that existentially closes off the external argument. The most important reason for the authors to postulate –D Voice is that they all believe that EA is not syntactically present in passives, while the verb has an EA interpretation.

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.

(22) 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 so 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 there does not exist –D Voice as a variant of Voice, alternative to +D Voice; if Voice is a +D Voice, it is always so.[7]

All the specific analyses for –D Voice cannot escape the fate facing "the overuse problem" for Voice. If you identify a head, namely, H, with argument-introducing ability as "Voice", why do you identify another without that ability also as "Voice"? If Voice is H with [S:N] that is checked off by a DP argument that it introduces, then H

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.

[7] 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.

should not turn out to be a different one on another occasion. If there is really H that has [S:N] checked off by means of being selected by another head, then H should not be Voice. The –D Voice approach would make the voice type and voice semantics of a sentence predetermined by a particular type of Voice in the lexicon. In human languages, numerous voice constructions such as passive, middle, reflexive, and causative, among many are available. Under the –D theory, these voice types are all specified in the lexicon, which contains corresponding types of Voice heads. How can this be so? If UG had really invented Voice with a passive/active-specifying property/ability/feature in the inventory of functional heads, then UG would remain quite "expedient".

Collins (2024), however, 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-vP.

For Collins (2024), VoiceP is a projection that appears with passives to account for word order and the licensing of the participial morphology (2024: 82); VoiceP plays no role in the projection (external Merge) of arguments (2024: 103); the external argument is not externally merged in Spec-VoiceP (2024: 141). Therefore, Collins (2024) thinks that Voice is subject to Terminological Assumption.

(23) 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.

However, Collins' (2024) conclusion remains too strong. If we identify the head that introduces EA as *v*, which selects the lexical core VP, then it would be unnecessary to postulate a Voice head for other reasons. If Voice existed just for voice morphemes, then we would have to abandon the morphology-syntax mirroring. Voice, as will be evident, does play a crucial role in accounting for voice phenomena; it does not fail Terminological Assumption.

3. Voice and Split VoiceP (preliminaries)

In this paper, I argue that Voice is the very argument introducer and presents argumentations for this. Some of the characteristics of Voice that will emerge as consequences of the discussion are given below.

(24) a. Clauses are built through successive mergers of arguments as potential subjects (sbj, opposed to SBJ, the surface subject) in Spec of independent Voice heads;
b. Voice is a sbj-introducing head; sbj can be EA, IA, or an applied argument;
c. The Voice domain consists of (at least) three distinct Voice projections.

3.1. Introduction

This section presents empirical and theoretical support 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.

(25)… [$_{\text{VoiceP3}}$ DP3 [$_{\text{VoiceP2}}$ DP2 [$_{\text{VoiceP1}}$ DP1 [$_{\text{VP}}$ V]]]] …

Split VoiceP contains at least three VoicePs. Simply stated, Voice1 refers to a transitivizing head, roughly corresponding to Kratzerian Voice and Chomskyan *v*. Voice2 is another instance of Voice that selects VoiceP1 and Voice3 selects VoiceP2. VoiceP1, VoiceP2 and VoiceP3 make up a split Voice projection, as with the case of CP, which splits into distinct projections such as TopP and FocP. Voice1, Voice2 and Voice3 do not differ in nature. They differ only in their height. Importantly, none of the three is specified as a particular type of Voice. Neither is identified as Passive or Active in the lexicon. They themselves do not produce passive and active semantics. Nor do they contribute to voice semantics in any way. They are simply argument-introducers, not being sensitive to the type of arguments. Either the so-called EA or IA may be introduced by any by Voice1, Voice2 and Voice3.

On the surface, the model in (25) does not play a role in accounting for voice phenomenon and is subject to Terminological Assumption. Therefore, to guarantee that the Split VoiceP approach does not fail Terminological Assumption, we must justify the following statements.

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

(27) The subject of a passive phrase lands in a non-nominative position or stops over there before reaching the nominative position.

By justifying (26), the Split VoiceP approache must be able to identify the so-called passive and a different voice construction as voice alternants interconnected by Voice. By justifying (27), it must also be able to identify the passive subjects as DPs introduced by Voice. By justifying (26) and (27), we will be able to show that Voice does play a role in accounting for voice phenomena and deriving voice variation.

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, 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 that passive semantics is produced configurationally

where the patient,[8] the passivized subject, occurs higher than the suppressed agent in the syntactic hierarchy.

3.2. Passivization is as a short-step movement

Uner the mainstream tradition, passivization is a long-step momement. By "long-step momement", I mean that IA moves at a single step to SBJ, without stopping over an intermediate position, such as Spec-vP/VoiceP. In this section, I, elaborating on the derivation of causative-of-passive sentences in Japanese, show that passivization is a short-step movement, where IA stops over a position lower than SBJ but higher than EA or ends up there.

In causative-of-passive sentences ("causative-passive" in Tsujimura 1996: 258-259) in Japanese (28), 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 morpheme (PS) -*are* in (28) and (29), 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).

(28) 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)

(29) Hanako-ga Taroo-ni sikar-are-ta.
Hanako-NOM Taroo-DAT scold-PS-PST
'Hanako was scolded by Taroo.' (Tsujimura 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),[9] 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,

[8] 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.

[9] "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.

the derivation extends in the voice domain, leading to a causative (of passive) structure (28), while in the latter, it proceeds in the tense-aspect domain, remaining a passive structure (29). Importantly, passivization is already completed by promoting *Hanako*, as shown in (30), regardless of where it ends up in the surface structure, given that the promotion of IA and the suppression of EA make up the core property of passivization.

(30) A-to–D raising as passivization:
[$_{TP}$ [$_{XP}$ Hanako$_i$ [$_{VoiceP}$ Taroo [$_{VP}$ t_i sikar]] -are]]

A-to-D raising further confirms that passivization involving NOM is not a one-step movement, but rather a successive-cyclic one. (27) is thus justified by A-to-D raising. The successive-cyclicity 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, saying this much does not suffice to justify (26) for guaranteeing that the Split VoiceP approach does not fail Terminological Assumption in (2). It must be explicated 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.

3.3. Passivizing head is an argument-introducer

This section presents arguments for the claim that a passivizing head also introduces arguments, as with the case of a transitivizing head, thereby showing that passive subjects and active subjects are both introduced by Voice, with the view of justifying (26). "Passivizing head" here is used to refer to a functional head that introduces IA via IM into its Spec, thereby deriving a passive structure. This does not mean that there is a head that is designated as a passive voice head. The layered affixation of voice suffixes serves as empirical support for my argumentation. The first step for this move is to show that a canonical passive suffix and a canonical causative suffix can spell out the same level of Voice so that we can prove that IA and a causer can be introduced by the very same head. From this, it will follow that passives and causatives as interconnected voice constructions.

The syntax-morphology mirroring in passives and causatives in Japanese indicates that the voice domain splits into at least three voice projections, notated as "VoiceP1", "VoiceP2" and "VoiceP3". 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.

(31) [$_{VoiceP3}$ [$_{VoiceP2}$ [$_{VoiceP1}$ […] -lex1] -syn2] -syn3]

As outlined in section 1, 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 functional heads, spelled out by fully-fledged suffixes notated as syn2 and syn3, respectively. What we argue for in what follows is then that Voice2 is either a (syntactic) passivizing head or a causativizing head (but not designated as such in the lexicon); that is, syn2 is either a passive suffix or a causative suffix.

Let us first exemplify the layered structure [$_{\text{VoiceP2}}$ [$_{\text{VoiceP1}}$ […]]], which mirrors the string -lex1-syn2. Previous studies such as Pylkkänen (2008), Harley (2008, 2013), Wurmbrand and Shimamura (2017), Nie (2020) among others have discussed -lex1-syn2 in one way or another. The following sentences exemplify -lex1-syn2. In the passive-of-lexical-causative sentence (32), lex1 is a causative morpheme CS and syn2 is a passive morpheme PS. In causative-of-lexical-causative sentences (33), lex1 and syn2 are both CS.

(32) Ookii biru-ga seihu-ni yotte tat-e-rare-ta.
big building-NOM government-by stand-CS-PS-PST
'A big building was built by the government.' (adapted from Tsujimura 2014: 269)

(33) Taroo-ga Ziroo-ni sensei-o nak-as-ase-ta.
Taro-NOM Ziro–DAT teacher-ACC cry-CS-CS-PST
'Taro made Ziro make the teacher cry.' (Pylkkänen 2008: 122)

As seen clearly, lex1 is the exponent of a lexical-causativizing head, which has sometimes been assumed to be identical to a transitivizing head. What is spelled out by syn2 is either a higher causativizing head or a passivizing head.[10]

However, the implementation of [$_{\text{VoiceP3}}$ [$_{\text{VoiceP2}}$ […]]], which is mirrored by -syn2-syn3, has not been explicated in the literature. The following facts serve as evidence that the morphemes represented by syn2 and syn3 are syntactic productive affixes. First, they can be fully-fledged, unlike lex1, which is more compact and sometimes lacks a consonant. Second, they always occur outside transitivizing morphemes. Third, they are not compatible with idiosyncratic meanings.[11] The -syn2-syn3 string may be PS_{syn2}-CS_{syn3} or CS_{syn2}-PS_{syn3}. We first discuss PS_{syn2}-CS_{syn3}.

PS_{syn2}-CS_{syn3} instantiates A-to-D raising, which represents passivization, as we observed in §3.2. Given that passivization is a syntactic operation and lexical-causativization feeds passivization but not vice versa, PS_{syn2} must be attached higher than CS_{lex1} and the attachment of CS_{syn3} is even higher than that of PS_{syn2}. PS_{syn2} introduces the passivized IA, e.g., *Hanako* in (34) and CS_{syn3} introduces a causer, e.g., *Ziroo* in (35).

[10] For more details 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.
[11] This can be syntactically tested out by Harley's (2008: 3ff) disgnostics.

(34) Hanako-ga Taroo-ni sikar-are-ta.
Hanako-NOM Taro-DAT scold-PS-PST
'Hanako was scolded by Taro.' (= 29)

(35) Ziroo-ga Hanako-o/ni Taroo-ni sikar-are-sase-ta.
Ziro-NOM Hanako-ACC/DAT Taro-DAT scold-PS-CS-PST
'Ziro made Hanako be scolded by Taro.' (= 28)

[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 (36), *-sase*, selecting a transitive phrase, is merged immediately above the transitivizing head (or lexical causative head), while in the causative-of-passive sentence (35), *-sase*, selecting a passive phrase created by A-to-D raising, is merged higher than the transitive-selecting head.

(36) 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] ...

Note that in (34), the causee is IA *Hanako* and in (35), it is EA *Taroo*. Since the affixation of CS *-sase* follows that of PS *-are* in (34), 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.

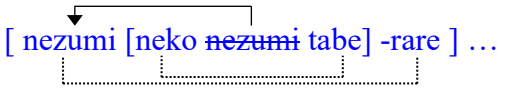
In sum, in causative-of-passive sentences, the passive morpheme *-rare* instantiates syn2, bound with the introduction of IA, and the causative morpheme *-sase* instantiates syn3, bound with the introduction of the causer.

We proceed to discuss passive-of-causative sentences,[12] 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_{syn3}, which was often undifferentiated from PS_{syn2} in direct passive sentences in previous studies. Morphology alone is not helpful in distinguishing them. Instead, we need to look at their syntactic properties instead. Compare the following sentences.

[12] Passive-of-causative sentences had various names including "causative passive" in Tsujimura (1996), "indirect causative" in Harley (2013) and "causative-passive" in Agoyagi (2021).

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

(38) Nezumi-ga neko-ni tabe-rare-ta.
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 (37). Therefore, in this case *-rare* is not associated with IA *nezumi*, but with EA *neko*. Now it is clear that in (37) 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 (38), PS *-rare* spells out the head that promotes IA *nezumi*, while in (37), it spells out the head that promotes EA *neko*. Importantly, in (37) 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 the promotion of *neko* takes place, it represents passivization and the affixation of *-rare* also takes place. That is, (external) merger of *neko* in its base position precedes the affixation of *-rare* in passive-of-causative sentences. This means that syntactic operations such as the 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 (37) and (38) and therefore the height of affixation of -*rare* varies accordingly. The affixation of *-rare* in *-sase-rare* (37) is higher than that of root-selecting *-rare* (38). That is, passivization in (37) and that in (38) are *not operations of the same level.* However, instead, causativization in (37) and passivization in (38) are *operations of the same level* in the sense that they take place at the same height, immediately after the completion of transitivizing.

In sum, syn2, syn3 and lex1 must be distinguished. Lex1 is the exponent of a semi-lexical and semi-functional head, which introduces EA. Note 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, (26), repeated below as (39), is justified, thereby verifying that the Split VoiceP approach does not fail Terminological Assumption.

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

3.4. Formulation of Voice as an argument introducer

As seen from the discussions in the preceding sections, PS and CS are exponents of Voice heads, which are bivalent functional heads introducing an argument as a sbj and a clausal complement. Voice is thus an argument-introducer.

Wood and Marantz (2017) argue that there is a type of argument introducer, notated as "*i**", as an independent head has a selectional feature, [S:D], which is satisfied by selecting a DP as an argument. That is, the syntactic properties of *i** are derived from its structural environment: it determines the syntactic category of its complement, for example, vP, and 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. 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 short-step 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. Additionally, it would become more complicated if we decompose the transitive phrase into 'a more complex' structure containing an additional head such as *i** mediating between a DP and a verbal head.

Unlike Wood and Marantz (2017), I argue that Voice, being a truly autonomous syntactic head, is the very argument-introducer with [S:D]. [S:D] on Voice is satisfied by introducing a DP, which therefore acquires subjecthood but not necessarily bears the EA role. In this respect, the present analysis is similar to that of Bruening (2012), who assumes [S:N] for Voice. However, the present analysis differs from Bruening (2012) in that it holds that [S:D] is always valued by a DP, not by a head (his Pass) selecting VoiceP. Note that [S:D] here is construed as an ability to select for a DP; it can also be construed as a selectional property that characterizes Voice. However, [S:D] here is not

a voice-specifying feature. It is not sensitive to whether a DP is an external or internal argument. Any DP, when selected by Voice, may check off [S:D]. This comes as a natural result of Free Merge. [S:D] is not arbitrary. It is always encoded by Voice. Voice is thus always +D Voice.

4. Subjecthood Condition and derivation of voice variation

A sentence subject originates as a potential subject; that is, what becomes sentence subject is assigned subjecthood before it becomes a sentence subject. This property is formulated as "Subjecthood Condition".

(34) Subjecthood Condition
Voice introduces a DP and assigns subjecthood to it. That DP is thus a potential subject, sbj. A sbj becomes the sentence subject SBJ if it is the highest sbj.

In connection with Subjecthood Condition, we propose the following denotation for VoiceP, following Kratzer (1996).

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

(35), however, excludes the case in which Voice introduces IA. We then propose (36) for VoiceP involving introduction of IA.

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

The unspecification of the role for DP is a reflex of Free Merge, where nothing constraints merge of a DP. The role of DP in (35) and (36) does not affect the nature of Voice. This means that the so-called EA and IA are undistinguished upon being introduced by Voice. Strictly speaking, there does not exist EA under a Voice theory, as noted by Kratzer (1996: 131ff), with "external argument/EA" used just for the purpose of notation. As regards IA, they are not to be done away with because they are base-generated truly in VP. However, their status as IA is not sensed by Voice. Voice is not designated to specify the role of the argument it introduces.

In light of Subjecthood Condition, we rewrite (35) and (36) as (37), where DP is not identified as the same role as *y* or not but rather as a potential object. That is, (37) 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.

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

As a builder of subject-predicate relation, Voice extracts any DP out of VP that wants to become a subject and reinstalls it into the machinery denoted by (37).

In other words, the status 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.[13] In this sense, EA is "suppressed" in passives not because Voice does not introduce EA as assumed by the –D Voice approach but because EA is deprived of its subjecthood. In other words, the case is that EA is suppressed because IA is promoted, but not that IA is promoted because EA is suppressed. Deprived of its subjecthood, EA becomes a non-predicate-internal subject. When subjecthood transmitted to the promoted IA, IA is now a predicate-internal subject (specifically, 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-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 building is structure extension by introducing arguments as potential subjects.

In what follows, we elaborate on 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 (38). If IM applies, sbj is IA and a one-place predicate, as in (39).

(38) Basic structure:
[$_{\text{VoiceP1}}$ sbj [$_{\text{VP}}$ DP1]]

(39) External merge of DP2 as sbj:
[$_{\text{VoiceP1}}$ DP2 [$_{\text{VP}}$ DP1]] (transitive; lexical causative)

(40) Internal merge of DP1 as sbj:
[$_{\text{VoiceP1}}$ DP1$_i$ [$_{\text{VP}}$ t_i]] (unaccusative; (dispositional) middle; fake passive)

Subjecthood condition applies equally in extending the structure to VoiceP2, thereby deriving the various voice constructions, which are thus interconnected via Voice.

Voice2 selects VoiceP1 and introduces a DP as sbj2. There are two candidates for sbj2. If EM applies, DP3 is merged as sbj2, yielding a syntactic causative structure (a causative of transitive structure). Thus, causative semantics arises upon creating the configuration in which DP3, a new DP, is merged (see also Nie 2020 and Wood and Sigurðsson 2021). Alternatively, DP1 is merged as sbj2 via IM, which yields a structure in which the patient is located above the agent. This is what we call "passive". Thus, passive semantics arises upon creating the configuration in which DP1 is merged above DP2. A third option is to merge DP2 as sbj2 via IM. However, this option cannot be

[13] That is, the relevant head (Voice2) is inserted a passive morphology.

instantiated because it does not extend the structure.

(41) Basic structure1 of VoiceP2:
[$_{VoiceP2}$ sbj2 [$_{VoiceP1}$ DP2(=sbj1) [$_{VP}$ DP1]]]

(42) External merge of DP3 as sbj2:
[$_{VoiceP2}$ DP3(=sbj2) [$_{VoiceP1}$ DP2(=sbj1) [$_{VP}$ DP1]]] (syntactic causative)

Ziroo-ga Taroo-ni Hanako-o sikar-ase-ta.
Ziro-NOM Taro-DAT Hanako-ACC scold-CS-PST
'Ziro made Taro scold Hanako.'

(43) Internal merge of DP1 as sbj2:
[$_{VoiceP2}$ DP1(=sbj2) [$_{VoiceP1}$ DP2(=sbj1) [$_{VP}$ t_{DP1}]]] (syntactic passive)

Nezumi-ga neko-ni tabe-rare-ta.
mouse-NOM cat-DAT eat-PS-PST
'The mouse was eaten by the cat.' (Aoyagi 2021: 87)

Ookii biru-ga seihu-ni yotte tat-e-rare-ta.
big building-NOM government-by stand-CS-PS-PST
'A big building was built by the government.'
(adapted from Tsujimura 2014: 269)

(44) Internal merge of DP2 as sbj2:
*[$_{VoiceP2}$ DP2(=sbj2) [$_{VoiceP1}$ t_{DP2}(=sbj1) [$_{VP}$ DP1]]]

When the derivation of VoiceP2 is based on (40), which is a one-place predicate, only one option can be instantiated. That is, DP2 is merged as sbj2 via EM. IM of DP1 does not apply because it fails structure extention.

(45) Basic structure2 of VoiceP2:
[$_{VoiceP2}$ sbj2 [$_{VoiceP1}$ DP1(=sbj1) [$_{VP}$ t_{DP1}]]]

(46) External merge of DP2 as sbj2:
[$_{VoiceP2}$ DP2(=sbj2) [$_{VoiceP1}$ DP1(=sbj1) [$_{VP}$ t_{DP1}]]] (causative of lexical causative)

Reiko-ga Hanako-ni yoofuku-o aw-ase-sase-ta.
Reiko-NOM Hanako-DAT clothing-ACC meet-CS-CS-PST
'Reiko made Hanako match her clothing.' (Miyagawa 1998: 88)

(47) Internal merge of DP2 as sbj2:
*[$_{VoiceP2}$ DP1(=sbj2) [$_{VoiceP1}$ t_{DP1}(=sbj1) [$_{VP}$ t_{DP1}]]]

The structure further extends to VoiceP3, with two basic structures. On the basis of the

first one, sbj3 is merged with a three-place predicate (syntactic causative). Now there are four DPs that can be considered. Firstly, DP4 is merged as sbj3. This would yield a structure in which the syntactic causative is causativized. However, this structure is not instantiated in many languages, arguably due to a morphological bottleneck effect. IM applies for DP2, not for DP3 and DP1. DP2 is merged as sbj3, yielding a passive-of-causative structure, in which the agent is promoted to SBJ. IM does not apply for DP3 due to the failure of structure extension. IM of DP1 as sbj3 also fails because it cannot move over two DPs (DP2 and DP3) at a single step. If IM of DP1 over DP2 and DP3 was successful, it would yield a passive-of-passive structure. However, genuine passives-of-passive do not exist.

(48) Basic structure1 of VoiceP3:
[$_{\text{VoiceP3}}$ sbj3 [$_{\text{VoiceP2}}$ DP3(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ DP1]]]]

(49) External merge of DP4 as sbj3:
*[$_{\text{VoiceP3}}$ DP4(=sbj3) [$_{\text{VoiceP2}}$ DP3(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ DP1]]]]
(causative of syntactic causative)

(50) Internal merge of DP3 as sbj3:
*[$_{\text{VoiceP3}}$ DP3(=sbj3) [$_{\text{VoiceP2}}$ t_{DP3}(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ DP1]]]]

(51) Internal merge of DP2 as sbj3:
[$_{\text{VoiceP3}}$ DP2(=sbj3) [$_{\text{VoiceP2}}$ DP3(=sbj2) [$_{\text{VoiceP1}}$ t_{DP2}(=sbj1) [$_{\text{VP}}$ DP1]]]] (passive of casuative)

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

(52) Internal merge of DP1 as sbj3:
*[$_{\text{VoiceP3}}$ DP1(=sbj3) [$_{\text{VoiceP2}}$ DP3(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ t_{DP1}]]]]

When structure further extends to VoiceP3 on the basis of the second basic structure, which is a syntactic passive, only one of the three options can be instantiated. That is, DP3 is merged as sbj3 via EM, which yields a causative-of-passive structure. IM of DP2 fails because it would give rise to the same outcome as (39), in which DP2 precedes DP1. IM of DP1 also fails because there would be no structure extension.

(53) Basic structure2 of VoiceP3:
[$_{\text{VoiceP3}}$ sbj3 [$_{\text{VoiceP2}}$ DP1(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ t_{DP1}]]]]

(54) External merge of DP3 as sbj3:
[$_{\text{VoiceP3}}$ DP3(=sbj3) [$_{\text{VoiceP2}}$ DP1(=sbj2) [$_{\text{VoiceP1}}$ DP2(=sbj1) [$_{\text{VP}}$ t_{DP1}]]]] (causative of passive)

Ziroo-ga Hanako-o/ni Taroo-ni sikar-are-sase-ta.
Ziro-NOM Hanako-ACC/DAT Taro-DAT scold-PS-CS-PST
'Ziro made Hanako be scolded by Taro.' (= 15/28/35)

(55) Interternal merge of DP2 as sbj3:
*[$_{VoiceP3}$ DP2(=sbj3) [$_{VoiceP2}$ DP1(=sbj2) [$_{VoiceP1}$ t_{DP2}(=sbj1) [$_{VP}$ t_{DP1}]]]]

(56) Interternal merge of DP1 as sbj3:
*[$_{VoiceP3}$ DP1(=sbj3) [$_{VoiceP2}$ t_{DP1}(=sbj2) [$_{VoiceP1}$ DP2(=sbj1) [$_{VP}$ t_{DP1}]]]]

In any step of derivation, a last-merged sbj, which is the highest one, is chosen for promotion to SBJ, with others, if any, remaining non-nominative. A non-last-marged sbj fails promotion since it is blocked by a last-merged sbj, instantiating a Relativized Minimality effect. Relativized Minimality does not hold within VoiceP. In one language, a non-last-merged sbj may surface with dative, accusative, instrumental, or some other case, while in another language, it may result in PP. In this sense, agentive PPs like *by*-phrases are indeed on par with non-nominative DPs, both being KPs, as discussed by Collins (2024). Thus, voice alternations are derived by demoting or suppressing non-last-merged sbjs, not by placing voice-specifying features on Voice heads.

5. Elaboration on derivation of one-place predicates

One-place predicates include unaccusatives, (dispositional) middles and fake passives. As seen in section 4, these voice alternants are derived by promoting DP1 to Spec-VoiceP1, where DP2 would otherwise be merged as an external argument or a causer. There are certain reasons to assume that IA is base-generated as the object and promoted to Spec-VoiceP1 as sbj.

Uunaccusatives are derived by phase and their subject is realized as SBJ in a successive-cyclic way, stopping over the edge of a phase head before reaching Spec of T (Legate 2003; Richards 2007). If a head selecting the root is a phase head, then Voice1 is a phase head, Spec of which is an intermediate landing site for IA, where it is sbj and ready for promotion to SBJ. Regarding middles, Lekakou (2005) observed a core property shared by them across languages: IA is promoted to the subject position via IM, with EA absent from syntax. This is evidenced by the fact that (dispositional) middles in English do not allow modification by agentive adverbials, can control into purpose clauses, and do not tolerate the licensing of the *by*-phrase (Alexiadou 2014: 21). Morphologically, middles are either active-, passive- or reflexive-marked according to the morphosyntactic parameters of the particular languages (Alexiadou 2014: 21ff). This is to say that the middle subject is introduced by a head unspecific with respect to voice type.

With Spec of VoiceP filled by IA, there is no position open to EA. In this sense, EA is just syntactically absent from the beginning but not suppressed or demoted. The

would-be agentive interpretation of EA in middles arises from the ability of the verb to participate in a transitive-middle alternation. Where no alternation is available, the would-be agentive interpretation of EA would not follow, which is the case of unaccusatives (of the *arrive* type). In this sense, the middle itself is intransitive, just like unaccusatives. *Break* type of unaccusatives participate in labile alternations and therefore are also interpreted as two-place predicates in transitives, whereas *arrive* type of unaccusatives do not. Both types of unaccusatives differ from middles in that they carry change-of-state/location meaning, whereas middles carry a stative generic meaning. The change-of-state/location meaning of unaccusatives arises from their special structuring of roots and/or inner aspectual organization, whereas the generic meaning of middles arises from being existentially bound. When both take place, the verb may be both middle and unaccusative.

(58) This kind of vases break easily. (Alexiadou 2014: 36)

For another case, we 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.

(59) Nad nutag-min sana-gd-laa. [Mongolian]
1st-DAT homeland-NOM-PSS miss-PS-PST
'I got homesick.'
'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* to Spec-VoiceP1 does not represent passivization, *-gd* does not spell out Voice2, but spells out Voice1. That is, the affixation of *-gd* takes place hand in hand with the introduction of IA *nutag* as sbj in Spec of Voice1. The semantics produced in this case is neither causative/transitive nor passive semantics, but rather semantics associated with spontaneity.[14]

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).[15]

[14] 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.

[15] Bai (2023) observes that *neme-gd* was used with an agentive phrase in Middle Mongolian and

(60) Nad hüč neme-gd-sen. [Mongolian]
1st-DAT strength-NOM add-PS-PST
'My strength increased.'
'Lit. To me, strength was added.'

In 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 indicate that the promotion of IA is not a one-step movement. IA must stop over Spec of Voice1 before reaching the NOM position.

It is thus not difficult to see 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-promotion while lexical causatives involve EA-installation.[16] Both genuine and fake passives involve IA-promotion, but they differ regarding the intermediate site for IA. Genuine passives and syntactic causatives represent constructions that alternate at a higher voice layer in the voice domain.

It is not surprising that SBJ originating as IA stops over Spec-VoiceP1 under Free Merge (Chomsky 2015); nothing prevents from IA merging into Spec-VoiceP1. Voice as an argument-introducing head does not specify whether the argument it introduces is EA or IA. Voice1, being a semi-lexical head, is able to replenish the lexical semantics of the verb by (re)introducing an argument.

Some remarks regarding phase are in order. VoiceP1, which roughly corresponds to *v*P, is a phase, if we adhere to the classic phase theory (Chomsky 2001, 2008). However, it is evident that VoiceP2 and VoiceP3 behave identically to VoiceP1 and should also be taken as phases. Unfortunately, the voice domain will not contain multiple phases in the same clause. For the moment, let us assume that the highest VoiceP constitutes a phase. This would mean that when, for example, Voice2 selects VoiceP1, which is already a phase, the phasehood shifts to VoiceP2. However, this is not supposed to be the case because VoiceP1 as a phase would not allow DP1 (IA) to move to Spec-VoiceP2 in passives. With this said, the phasehood in the voice domain is not fixed but flexible if we are to maintain to use the phase theory to capture locality in syntax. That is, phasehood is determined contextually (Boškovic 2014; Gallego 2017). Alternatively, phase does not exist in VoiceP, with the locality effects ascribed to Relativized Minimality, as done by Halpert and Zeijlstra (2024). A further elaboration on this issue is left out, with our focus on the articulation of VoiceP.

6. Conclusion

underwent a diachronic change from an agent-taking verb to a recipient/beneficiary-taking one.

[16] Note 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.

This paper has argued that subject-predicate relations, the hallmark of clauses, are established through the syntactic head Voice in variably in actives and passives. The projection of the predicate must have a DP in its edge so that it is predicated by the predicate. Predicate-internal subjects are thus purely structural, rather than thematic or semantic.[17] Subjecthood reduces to a purely structural phenomenon (Poole 2016: 3). If grammatical subjects, which are formal relations, appear to be constrained by semantic, pragmatic, or prosodic factors, this is a misapprehension (Haider 2024b: 10). Voice variation arises in narrow syntax, rather than in the lexicon. Specifically, subjechood is distributed on multiple edge positions, where the merge type (external or internal) and merge height of DPs as sbjs derive various voice patterns. Voice, being the subject-introducing functional head available by UG, is thus truly at the heart of a theory of voice (Kratzer 1996: 120), whereby the grammatical voice results from Voice.

(to continue)

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.

Baker, Marc. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373-415.

Boškovic, Željko. 2014. Now I'm a phase, now I'm not a phase: On the varriability of phases with extraction and ellipsis. *Linguistic Inquiry* 45 (1): 27-89.

Bruening, Benjamin. 2012. *By*-phrases in passives and nominal. *Syntax* 16: 1-41.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris.

Chomsky, Noam. 1986. *Knowledgeof Language: Its Nature, Origin, and Use*. New York, NY: Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, eds.,

[17] This is on par with Poole's (2016) observation that subjecthood reduces to a purely structural phenomenon.

Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2008. On Phases. In *Foundational Issues in Linguistic Theory*, eds., Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta, 136-166. 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.
Cinque, Guglielmo. 2006. *Restructuring and Functional Heads: The Cartography of Syntactic Structures*. Oxford: Oxford University Press.
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.
Gallego, Ángel. 2017. Remark on the EPP in labeling theory: evidence from Romance. *Syntax* 20: 384-399.
Haider, Hubert. 2024a. ONE GRAMMAR IN 7.164 VARIANTS – Confessions of a veteran syntactician. https://ling.auf.net/lingbuzz/008473.
Haider, Hubert. 2024b. On the cognitive evolution of grammatical subjects. https://ling.auf.net/lingbuzz/007796
Halpert, Claire and Hedde Zeijlstra. 2024. Off Phases: It's All Relative(ized). https://ling.auf.net/lingbuzz/008323.
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. 2022. Valency and voice constructions. lingbuzz/005941
Jacobsen, Wesley. 1992. *The transitive structure of events in Japanese*. Tokyo: Kurosio.
Johnson, Kyle. 1991. Object Positions. *Natural Language and Linguistic Theory* 9: 577-636.
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 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.
Miyagawa, Shigeru. 1984. Blocking and Japanese causatives. *Lingua* 64: 177-207.
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*.
Poole, Ethan. 2016. Deconstructing subjecthood. Unpublished manuscript, University of Massachusetts Amherst. http://ethanpoole.com/papers/poole-2016-subjecthood.pdf.
Pylkkänen, Liina. 2008. *Introducing Arguments*. Cambridge, MA.: MIT Press.
Richards, Mark. 2007. Derving the edge: what is in phase. *Syntax* 14(1): 74-95.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, eds., Haegeman, L., 281-337. Berkeley: Kluwer.
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.
Tsujimura, Natsuko. 1996. *An Introduction to Japanese Linguistics*. Malden, MA.: Blackwell.
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.
Jim, Wood and Einar Freyr Sigurðsson. 2021. On the Interaction of Reflexives and Periphrastic Causatives in Icelandic. *Penn Working Papers in Linguistics* 27(1): 257-266.
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.
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.