The articulated Voice/v layer in Tamil

Sandhya Sundaresan and Thomas McFadden

sandhya.sundaresan@uni-leipzig.de, mcfadden@zas.gwz-berlin.de

Abstract

This paper contributes to the ongoing discussion of the proper analysis of clausal region immediately above the verb, by arguing that we need to recognize not just one or two functional heads, but a full layer of structure corresponding to Kratzer (1996)'s Voice or Chomsky (1995)'s v. This layer should include at least four functional heads arrayed above the root as shown below:

Pass > Mid > Trans > Cause > $\sqrt{}$

The primary evidence for these proposals comes from the Dravidian language Tamil, which is extremely informative due to its highly inflecting, agglutinative nature, and its flexibility in combining together distinct elements traditionally subsumed under the heading of 'voice'. Employing standard Mirror Principle reasoning, we can use the sequences of verbal suffixes that the language supplies to argue not only for an inventory of syntactic heads, but also for a specific hierarchy, due to the rigid ordering restrictions observed.

1 Background

A body of work in the late 1980s and early 1990s uncovered evidence for a syntactic position or positions above the typical location of the main verb in V, but below the representation of tense and agreement in Infl. As the arguments for the position were largely syntactic, and could not easily be correlated with familiar morphological or semantic categories,

the identity of this head (and whether there was just one or several) remained controversial and unclear, alternatively being labeled as Pr(ed), μ, or an additional segment of a layered VP structure (see e.g. Larson, 1988; Pesetsky, 1989; Johnson, 1991; Bowers, 1993).

In a seminal paper, Kratzer (1996) was able to motivate the existence of such a head on semantic rather than just syntactic grounds. She showed that the special status of external arguments as opposed to internal arguments (established e.g. by Williams, 1981; Marantz, 1984) could not be sufficiently accounted for if they were simply specifier and complement of a single syntactic object. Rather, a more radical solution was called for, a partial syntacticization of the Neo-Davidsonian semantics of Parsons (1990), whereby only the traditional internal argument is syntactically an argument of the verb, and the external argument (henceforth EA) is introduced by a distinct (functional) head. Kratzer labels this head Voice, associating it with traditional voice phenomena (i.e. passive vs. middle vs. active) and the assignment of structural accusative case, and suggesting that it is this that is responsible for the various syntactic facts discussed by the previous authors.

Subsequent work adopted Kratzer's basic proposal, sometimes preferring the label v introduced by Chomsky (1995). An array of additional functions were associated with this functional head split off from the verb, including the introduction of causative and eventive semantics (Pylkkänen, 2002), information related to inner aspect or *Aktionsart* (Ramchand,

2008), and the verbalization of category-neutral roots (Marantz, 1997). Very quickly, this led to the obvious question of whether a single head could in fact be responsible for so many distinct syntactic and semantic phenomena.

Work on causative constructions and alternations in particular (see e.g. Pylkkänen, 2002; Alexiadou, Anagnostopoulou, and Schäfer, 2006) uncovered convincing evidence that a single Voice/v head as assumed in earlier work would not do, and must be split into two projections. The lower one — usually labelled v or Caus(e) — is responsible for causative/eventive semantics and for the verbalization of non-verbal structure below. The higher one — usually labelled Voice — is responsible for the introduction of the EA, for agentive semantics and for the traditional 'voice' phenomena. The arguments for this move come primarily from how causative/inchoative alternants and various passive types differ in their semantics and in their ability to combine with particular modifiers and PPs. Harley (2013) bolsters this analysis with clear morphosyntactic evidence from Hiaki. Here an applicative suffix appears morphologically outside the causative suffix, hence by the Mirror Principle (Baker, 1985) the applicative is assumed to be above the causative head hierarchically. Yet the EA asymmetrically c-commands the argument introduced by the applicative, so the EA must be introduced by a head higher than the applicative. Taken together, these two facts imply that the causativizer and the head that introduces the EA must be distinct, the latter must be higher

than the former, and that the applicative comes in between.

These developments in our understanding of the space directly above the verbal root raise a number of questions. Having already gone from a single V node, to a Voice-V pair, to a Voice-Caus-Root triple, how many distinct heads must we recognize in total? To what extent is the inventory of such heads universal or parametrized? Are the relevant heads always present, or are some of them optional, as we might expect e.g. for applicatives? Are the heads rigidly ordered with respect to each other, and if so, how?

2 Contribution of this paper

This paper aims to contribute to this evolving discussion by providing novel evidence from the Dravidian language Tamil, which provides transparent morphological evidence showing that even further articulation is indeed needed in this layer of the clausal structure.

We must distinguish four heads responsible respectively for the syntax and semantics of causativity, transitivity, 'get'-like middles and the passive, all of which can — perhaps surprisingly — co-occur. There is, furthermore, clear mirror-principle evidence for the following rigid ordering of these four functional heads above the predicate root:

(1) Pass > Mid > Trans > Cause > $\sqrt{\ }$

As such, we contend that it is appropriate to talk about an articulated

layer or *domain* of event-modifying material rather than just one or even two heads — much along the lines of the articulated C layer/domain proposed within the cartographic framework (Rizzi, 1997, etc.).

The evidence we will present for this analysis is fairly straightforward, and derives from the fact that Tamil is a highly inflected, agglutinative language. A typical verbal form consists of the verb root followed by a sequence of functional morphemes with a rigid relative ordering. E.g. the form *oḍæ̆kkæpaṭṭadŭ* 'was broken' breaks down as follows:

(2) **oḍæ-** root 'break'
-kkæ- transitive stem
-paṭ- passive marker
-ṭ- past marker
-adŭ 3sg neuter agreement marker

The allowable combinations and orderings of such morphemes give us a fairly transparent window into the underlying hierarchical structure above the verb root, and will allow us to justify the proposal laid out above based on standard Mirror Principle argumentation.

We begin in Section 3 with the marking for transitivity. This presents the greatest analytic challenges on the morpho-phonological front, as it appears on the surface to involve a process of stem modification rather than a clear affix, and sorting it out first will simplify examination of the other markers. In Section 4, we turn to the passive marker, which is relatively straightforward in its behavior on all levels. This sets the scene

for the syntactically and semantically most challenging item, the suffix *ko*⌊ which, we will argue, instantiates a functional head 'Mid(dle)' with an argument-structure resembling that of a 'get'-passive. We lay out the basic facts of its distribution in Section 5, and then go into more detail on its proper analysis as a kind of middle marker in Section 6. Section 7 then examines how the three heads discussed up to this point — Trans, Pass and Mid — fit together, and in particular what happens when all three co-occur on a single verb form, motivating the hierarchy proposed above. Section 8 then embeds our findings for Tamil in the context of the broader work on the make-up of the *v* domain cross-linguistically, in particular the distinction made between Voice and Cause. We argue that the function attributed to just the Voice head in prior work is essentially distributed across a functional sequence consisting of three contiguous, rigidly ordered heads. We further show, on the basis, again, of morphological evidence from Tamil, that a Cause head must nevertheless still be distinguished in addition to these. This yields the full functional sequence above the predicate root, given in (1).

3 Transitivity marking in Tamil

Transitivity distinctions in Tamil are typically marked by a systematic alternation involving the final consonant (cluster) of the verb root or the immediately following suffix. In either position, a simplex consonant in the

intransitive is geminated in the transitive, which in turn triggers phonologically regular devoicing and (potentially) cluster simplification.[1] Thus the unaccusative form of 'break' is *oɖæ-nʤ-* in the past, with a voiced nasal + obstruent cluster as the past marker, as in (3), whereas the transitive form is *oɖæ-čč*, with a voiceless geminate, as in (4):

(3) **Unaccusative**:
Paanæ oɖæ-nʤ-adŭ/*oɖæ-čč-adŭ.
Pot break-ASP.INTR-3NSG/*break-ASP.TR-3NSG
'The pot broke.'

(4) **Transitive:**
Sri paanæ-jæ oɖæ-čč-aan/*oɖæ-nʤ-aan.
Sri pot-ACC break-ASP.TR-3MSG/*break-ASP.INTR-3MSG
'Sri broke the pot.'

Further examples of the alternation are provided in Table 1. Again, the pattern to note is that the transitive variants in the 3rd column have voiceless geminates in places where the intransitive variants have simplex voiced obstruents or nasal + voiced obstruent clusters.

The gemination alternation illustrated above seems to be regulated by the following generalization. When the verb root ends in a geminable consonant, it is geminated and devoiced, with subsequent cluster simplification, as in examples 1-3 in Table 1. When the verb root ends in a non-geminable segment — either a vowel or a consonant that does not

[1]Voicing is not phonologically contrastive (at least outside of borrowed vocabulary), being rather predictable on the basis of gemination and position in the word. Obstruents are voiceless initially and when geminated, voiced intervocalically and following nasals, and do not occur finally. In a cluster consisting of a nasal plus a geminate obstruent, the nasal is deleted.

	Translation	Unaccusative past	Transitive past
1.	SHRINK	surunɡ-in-	surukk-in-
2.	MELT	urug-in-	urukk-in-
3.	RUN	oo˞-in-	oot̪t̪-in-
4.	BREAK	o˞æ-nɡ-	o˞æ-čč-
5.	GROW	va˞ar-nd-	va˞ar-tt-
6.	BURST	ve˞i-nɡ-	ve˞i-čč-

Table 1: Unaccusative and transitive forms

permit gemination[2] — then the immediately following suffix is affected, as in examples 4-6 in Table 1. Given the region of clause structure involved here, this is typically a tense/aspect or infinitive marker.

But what is the syntactic correlate of (what looks like) a morpho-phonological process? In order to get at what's going on here syntactically, we will adopt the general approach of Bye and Svenonius (2012) who argue that patterns which have typically been analyzed in terms of morphological processes (and indeed have been presented as evidence for process-based morpholog) can, in fact, be profitably analyzed as piece-based, but where the pieces involved have underspecified forms that trigger phonological operations. For the specifics of the Tamil transitivity alternation we follow Christdas (1988), who proposes that the transitive variant involves an affix with a highly underspecified phonology, consisting of a consonant slot without specification for place or manner of artic-

[2]So far, it seems that /r/ is ungeminable, but we have not yet found secure examples of verbs with roots ending in certain other consonants like /ɻ/ that participate in the transitivity alternation in order to test how they behave.

ulation. The missing feature values will then be copied autosegmentally from the closest eligible consonant slot, deriving the apparent mobility of the alternation according to the shape of the verb root. The result will in any case be gemination, which triggers the further phonological rules of devoicing and cluster simplification (see Christdas, 1988, for details).

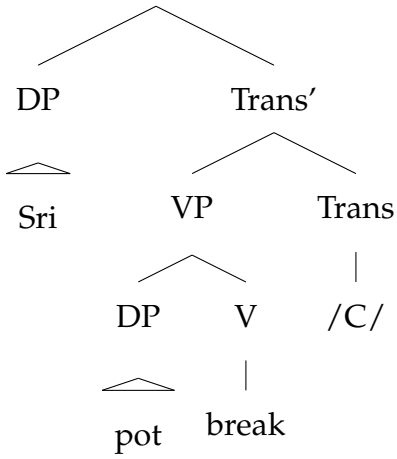
Under this approach, the transitivity alternation is thus the result of fully concatenative morphology, not a morphologically triggered 'process' of gemination. This is crucial because it means that once we get past the complexity introduced by the phonology, we can take the abstract morphology we uncover behind it to transparently reflect syntactic structure. In particular, we will argue that this underspecified consonant slot is the phonological realization of a syntactic head and, furthermore, that this head is situated above the verbal root but below T, since these are the targets for the gemination that it effects. Based on its syntactic and semantic role in tracking transitivity, we then propose that this is the head that combines with the verb phrase and introduces the external argument in its specifier. It is thus essentially the Voice head of (Kratzer, 1996).

However, since we'll shortly be arguing that there are other morphemes in the verbal complex performing other functions associated with the notion of voice, we'll depart from Kratzer by using the label "Trans" for this head in order to avoid confusion.[3]

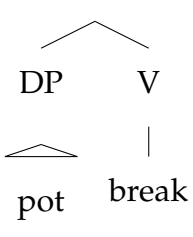
[3]It is important to note that the placeless consonant /C/ is not the only exponent of the Trans head. While it seems to be nearly universal with verbs participating in the

We envision then a transitive verbal structure like that in (5a), familiar again from Kratzer (1996)'s structure for the introduction of an external argument, and an unaccusative structure like that in (5b). We will update this a bit later, with the proposal that there is an additional Cause head between Trans and V. We are also not especially tied to the idea that the unaccusative variant lacks a Trans head entirely. An alternative would be to assume that Trans is always present, but has an inactive variant in unaccusatives that does not introduce an argument and has no phonological realization. At the moment we have no evidence that would decide between the two options, and adopt the one below for concreteness:

(5) a.



b.



causative-inchoative variation, and even among non-alternating verbs, transitives seem, by and large, to show a reflex of this /C/, either on their stem or in their tense formative, while intransitives do not, as we would expect (Schiffman, 1999, p. 75). There are, however, some apparent exceptions, and Schiffman does not reliably distinguish between unergatives and unaccusatives, which we of course predict should behave differently from each other. It seems we must at least allow for the possibility of Trans having marked allomorphs in the context of certain lexical roots, much like English T does.

The unaccusative variant in (5b) is basic, containing less structure and less overt phonological material. The root and following suffix will thus have the ungeminated form by default. In (5a), the presence of the underspecified C node realizing the Trans head will lead to gemination, either of the root or of the tense/aspect morpheme that comes above it. In the case of *oḍæ* 'break', which ends in a vowel, it will be the latter.

4 Passive marking in Tamil

Now consider the passive structure in (6a):

(6) a. (Sri-aal) paanæ oḍæ-**kkæ-pa**ṭ-ṭ-adŭ.
(Sri-INSTR.) pot.NOM break-TR.INF-PASS-PST-3NSG
'The pot was broken (by Sri).'

b. * (Sri-aal) paanæ oḍæ-**jæ-pa**ṭ-ṭ-adŭ.
(Sri-INSTR.) pot.NOM break-INTR.INF-PASS-PST-3NSG

Note that the passive is formed in Tamil by means of the suffix *-paḍ-*.[4] This example is instructive on two points that are relevant for our concerns. First, it shows that the passive suffix has to be added to the transitive alternant of the verb *oḍæ-kkæ-*, with use of the intransitive stem *oḍæ-jæ-*, as in (6b), being ungrammatical.[5] This lends further support to the view (espoused by Embick, 2004, and many others) that passives are

[4]It appears here as *-paṭ-* due to voicing assimilation with the following past tense suffix.

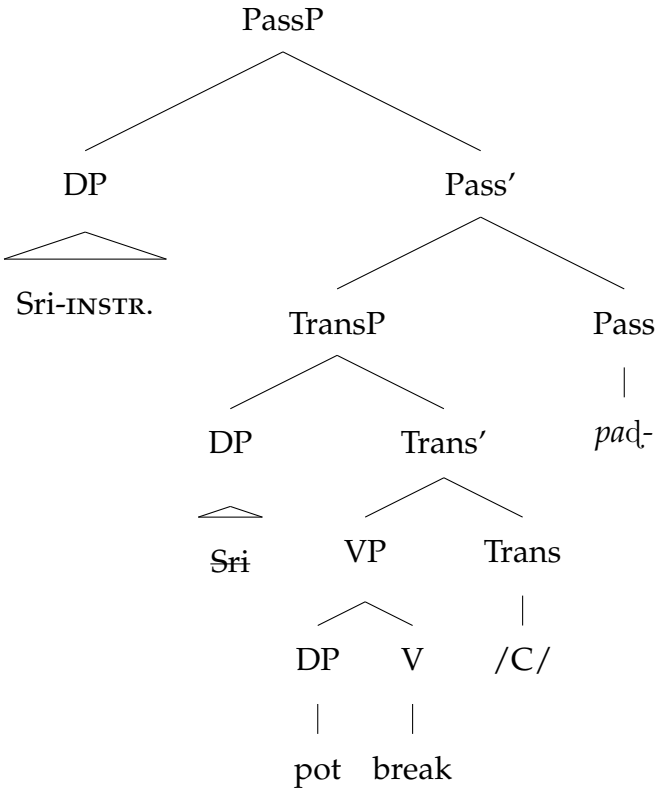
[5]The phonology of the alternation of the so-called 'infinitive' suffix is more complicated than the cases discussed until now, and a discussion goes a beyond what is appropriate here given our primarily syntactic concerns.

more agentive than unaccusatives in a clear sense. The former involve a fundamentally transitive agentive structure, the expression of whose external argument is suppressed, while the latter do not have the structure to host an agentive argument to begin with. Second, (6a) also shows that the passive marker appears **outside** the verbal structure to which the phonological gemination process applies. That is, the passive structure is transparently built on top of the transitive structure, not, so to speak, underneath it, and not in any combination with the intransitive structure.

This is most elegantly captured by proposing, in line with the Mirror Principle, that the head spelled out by the passive marker *-paḍ-* — we'll call it Pass — is above Trans in the syntactic structure. This results in further articulation in the v domain, as illustrated in (7):[6]

[6]We are ignoring for now the structural status of the 'infinitive' suffix realized as *-kkæ-* in example (6a), which shows up in a number of distinct contexts in forming complex verb forms. We have as yet identified no clear semantic or syntactic contribution for it, thus it remains an open question whether it is the realization of a syntactic head between Trans and Pass or is somehow inserted for purely morphological purposes.

(7)



The data from passives thus provide a clear argument against the idea that there is a single head (e.g. Voice) that is responsible for the transitive-unaccusative alternation as well as the active-passive one. Rather, the evidence transparently suggests that these alternations are encoded on two distinct heads, with the head responsible for encoding passivity being higher than that which introduces the external argument.

(8) **Interim conclusion:**

The syntax/semantics of transitivity and Passive is encoded on two separate heads with Pass > Trans.

4.1 Crosslinguistic concerns for Pass

As far as we are aware, the interim conclusion stated in (8) is not actually at odds with the evidence considered by prior work (Alexiadou et al., 2006, among many others) which did posit a single Voice head for both alternations. The focus of much of that work was to motivate a distinction between that head and a lower Caus(e) head responsible for introducing a causal relation between two eventualities. A central insight is that the 'causer' is not the agentive external argument, or indeed any individual but is, rather, a causing event. An agent may be implicated in that causing event, but need not be (indeed, so-called 'anticausatives' can actually be causative), thus it is correct to distinguish Cause from the agent introducing Voice. In this strand of work, Voice was then made the locus both of the transitive-unaccusative alternation and of traditional 'voice' alternations because they all involve the external argument, and because there was evidence suggesting that they should be split up. That is, in the languages considered, morphology associated with (in)transitivity does not co-occur with morphology marking the passive. It was thus simplest to assume a single Voice head with a variety of potential feature specifications, e.g. for agentivity and for whether or not it can or must introduce an argument in its specifier.

The data from Tamil make a crucial contribution here, because they do involve precisely the evidence necessary to show that the simplest assumption (of a single head) will not work here. The morphological re-

alization of the passive can co-occur with, and is clearly distinct from that for the transitive variant of an alternating verb. We must thus distribute these functions over two functional heads. Crucially, though, we foresee no difficulty in recasting Alexiadou et al. (2006)'s analysis of the facts from English, German and Greek, in terms of a single Voice head with an array of distinct feature specifications, into our analysis with both Pass and Trans heads, each allowing fewer featural variants. The lack of morphological articulation of these distinct heads could be dealt with under a spanning analysis like that of Ramchand (2008) where multiple heads can be made to spell out a single morpheme. Whether such a move is **empirically** warranted, however, can only be determined upon further investigation and careful comparison of these phenomena in the different languages – an endeavor beyond the scope of the current paper.

Splitting up the roles of transitivity and passive over two heads raises one important technical question. How do we model the interaction between the presence of the Pass head and the special behavior of the agent introduced by Trans? In the passive in Tamil, as in most languages, the agent is typically suppressed entirely, and when it is expressed, it bears oblique marking in the form of instrumental case, whereas it would be nominative in the active. This question does not present itself in such an urgent form for theories that have a single Voice head responsible for passive marking and introduction of the external marking. The passive variant of Voice must simply be specified, among other things, to not

introduce anything overt in its specifier. For us, this option is not available, since the two functions are not localized in a single head. Instead, we must posit a dependency between the Pass head and the Trans head, which has the desired effect on the requirements that Trans places on its specifier. At the moment we have no evidence to decide among various imaginable implementations of this, so we will leave it here as an open question, but we will discuss some evidence in Section 7 which suggests that the agent is associated with SpecTransP even in the passive.

5 Introducing the *ko⌊* morpheme

There is one additional morpheme we will be concerned with here which can occur in the *v* domain of the Tamil verbal sequence. Common practice in the literature is to refer to it simply as *ko⌊*, based on its form, because its syntax and semantics are rather difficult to characterize and thus have not suggested a more contentful category to assign it to.[7] We will follow this practice initially for convenience, though we will ultimately suggest a better label for it. Specifically, based on the detailed discussion in Sundaresan (2012), we will argue that *ko⌊* has a middle-like semantics. Its distribution, syntax and semantics and, in particular, how it interacts with the other markers we have seen so far, will be extremely instructive to

[7] As will be apparent in the examples below, its form is also subject to extensive variation based on the morphophonological context, but *ko⌊* is a good approximation of the underlying form that feeds into phonological processes.

our understanding of the shape of the v domain in Tamil. The conditions of where it appears are somewhat more complicated than with Trans and Pass, however — simply saying "middle-like" isn't particularly illuminating, given that the middle has been associated with many functions and meanings — so we will be a bit more detailed in describing it here.

One typical environment for *koɭ* is with reflexives. Indeed, clauses involving co-argument reflexivity usually require *koɭ* to be suffixed to the verb, as shown in (9):

(9) Sri$_i$ tann-æ$_{\{i,*j\}}$ aɖi-čču-kko-ɳɖ-aan/*aɖi-čč-aan.
Sri ANAPH-ACC hit-ASP.TR-*koɭ*-PST-3MSG/*hit-ASP.TR-3MSG
'Sri$_i$ hit himself$_{\{i,*j\}}$.'

koɭ is also frequently found suffixed to unaccusatives, as in (10):

(10) Paanæ oɖæ-nʤu-kko-ɳɖ-adu.
Pot break-ASP.INTR-*koɭ*-PST-3NSG
'The pot got broken.' (rough translation)

So the distribution of *koɭ* seems, at first blush, reminiscent of the (partial) syncretism between reflexive and unaccusative structures observed in a number of languages, e.g. with 'non-active' verb morphology in Greek (Embick, 2004) or with 'reflexive' clitics and weak pronouns in Romance languages, Slavic and German (Sportiche, 1998; Schäfer, 2008; Medová, 2009, and many others). Tamil would thus lend additional support to popular analyses (in the works just cited and elsewhere) according to which reflexives and unaccusatives share a structural subcomponent, and

the *koḷ* data would bolster the idea that reflexivity is a species of voice phenomenon. In fact, Lidz (2001, and subsequent) explicitly proposes as much based on similar data from the closely related Dravidian language Kannada. Lidz argues, in particular, that *koḷ* spells out a specifierless Voice head in both unaccusative and reflexive structures, very much along the lines of Embick (2004)'s analysis of non-active morphology in Greek.

However, closer inspection reveals that *koḷ* (at least in Tamil) realizes something distinct from the head responsible for the causative/inchoative alternation which we have been calling Trans, and corresponding to Lidz' notion of Voice. First, *koḷ* is actually optional on unaccusatives, as illustrated in the minimal pair below.

(11) a. Paanæ oḍæ-nʤ-adŭ.
Pot break-PST.INTR-3NSG
'The pot broke.'

b. Paanæ oḍæ-{nʤŭ/*ččŭ}-kko-ŋḍ-adŭ.
Pot break-{ASP.INTR/*ASP.TR}-*koḷ*-PST-3NSG
'The pot got broken.' (rough translation)

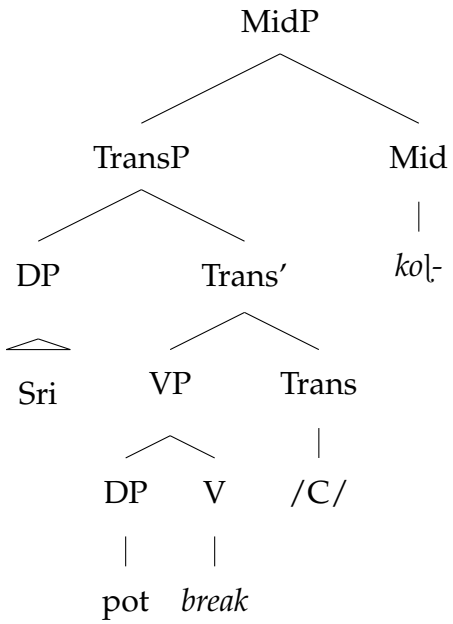
In other words, the presence of *koḷ* is not a necessary condition for the unaccusativity or intransitivity of (11b), since (11a), where there is no *koḷ*, is also unaccusative/intransitive. Second, *koḷ* not only appears on reflexives and unaccusatives but is also readily available on (non-reflexive) transitives, as in (12), showing that *koḷ* is also not a sufficient condition for unaccusativity (or reflexivity for that matter) :

(12) Sri paanæ-jæ oɖæ-{čču/*nʤu}-kko-ŋɖ-aan.
Sri pot-ACC break-{ASP.TR/*ASP.INTR}-*ko*ɭ-PST-3MSG
'Sri got the pot broken.' (rough translation)

Third, the verb form is already marked as either unaccusative or transitive independent of — and linearly prior to — *ko*ɭ-suffixation. This is illustrated by the obligatorily non-geminated form of the aspect morpheme in (11b) and its obligatorily geminated form in (12).

We'll consider the syntactico-semantic contribution of *ko*ɭ momentarily. The important message to take away from the data just presented is that, regardless of how it is ultimately analyzed, its distribution has nothing to do with the valency of the predicate it attaches to. I.e. *ko*ɭ must represent a syntactic head distinct from Kratzerian Voice/Trans, i.e. the head responsible for introducing the external argument. We'll call this head Mid (for 'middle'), a label we'll motivate in Section 6 below. Syntactically speaking, Mid must be higher than Trans (as per the Mirror Principle) since it linearly succeeds the locus of the gemination that indicates the position of Trans in the series of verbal suffixes. Thus, we can posit the preliminary structure (to be revised slightly in the next section) in (13) for a transitive clause with *ko*ɭ:

(13)



6 Motivating the 'middle-like' nature of *ko̲*

Turning now to the syntactic and especially the semantic contribution of *ko̲*, the data presented so far do not paint a very clear picture, especially given the apparent optionality of the morpheme in unaccusatives. Indeed, this is not just an artifact of our presentation. In the Dravidianist literature, *ko̲* is regarded as a notoriously tricky morpheme. A range of disparate and apparently contradictory meanings are identified, including self-benefaction, self-affectedness, volitionality, accident, inchoation from a state, and the simultaneity or completion of an action (Schiffman, 1999; Annamalai, 1999; Steever, 2005). It is perhaps most often described as a reflexive marker, given the fact that its presence is typically oblig-

atory to obtain co-argument binding, but the data above have already revealed this characterization to be too simple.

In order to shed new light on this confusing situation, Sundaresan (2012) investigates the effect of adding *ko*⌊ to a range of predicates culled from the verb classes identified by Levin (1993). She conducts a survey among native Tamil speakers of the acceptability of the sentences thus created against different discourse scenarios, thus honing in on its meaning contribution. Here, we briefly summarize the results of that study.

ko⌊ is subject to both thematic and aspectual restrictions. The former may be described as follows. Verbs that aren't compatible with a result state, either because they lack an eventive component (e.g. inherent statives), or because they actively resist the addition of one (e.g. involuntary emissives), are incompatible with *ko*⌊. However, telic change-of-state/location predicates (inchoatives) frequently co-occur with *ko*⌊. All other verbs which lack a result state but are compatible with one (e.g. most activity verbs) can optionally take *ko*⌊ when such a result state is added. The following thematic restrictions act in tandem with the aspectual ones described above. *ko*⌊ is incompatible with predicates that already have an intrinsic mental/spatial viewpoint (psych- and path-verbs), and with verbs that are incompatible with the holding of a mental/spatial viewpoint toward their result state by one of their arguments (e.g. involuntary directed motion verbs). Conversely, verbs that are especially conducive to such viewpoint-holding (postural and grooming verbs) fre-

quently co-occur with *ko*|.

This leads Sundaresan (2012) to the following proposal. *Ko*| attaches to the (derived) result state of a main event predication, with the interpretation that the highest argument of that predication (the Agent in transitives/unergatives and the Theme in unaccusatives) comes to hold/gets this result state in its mental/spatial locus, such that this result state comes to be evaluated from the mental/physical space of this argument. Thus, *ko*| introduces a semantics much like those of Sells (1987)'s SELF (described as "one whose mind is being reported") and PIVOT ("if someone makes a report with Mary as the PIVOT, that person is understood as literally standing in Mary's shoes") roles. The precise lexical entry for *ko*| is given below:

(14) $[\![ko|]\!] = \lambda Q_{<s_s,t>} \lambda x \lambda e' \exists s.Q(s) \wedge Get(e') \wedge Locus(e', x) \wedge Theme(e', s)$

(14) states that *ko*| takes a stative proposition as its argument and binds off the state. It further takes an individual and a(n) (sub-)event and relates them to the result state. Specifically, x-*ko*|-[QP] means "x comes to hold the derived result state denoted by QP".[8] Note, finally, that the Locus predicate is underspecified with respect to whether it is interpreted as a mental or spatial locus, with the information contextually supplied. This is modelled as an intrinsic vagueness in the lexical entry of *ko*|, yielding meanings that seem mutually incompatible. An interpretation of Locus

[8]This is incidentally a nice result because it is very close to the meaning of *ko*| when it served as an independent lexical verb in older stages of the language, and meant *hold*.

along the mental dimension yields an agentive effect – compatible with a self-benefactive reading. If the Locus reading is interpreted along the spatial dimension, however, we get a more Patient-like reading, compatible with one of physical self-affectedness.

We can illustrate this contribution of *koɭ*, including the relevance of the vagueness between a mental and a spatial Locus, by means of the pair of examples below:

(15) a. Mansi paal-æ uutt-in-aaɭ.
Mansi milk-ACC pour.TR-PST-3FSG
'Mansi poured the milk.'

b. Mansi paal-æ uutt-i-kko-ŋḍ-aaɭ.
Mansi milk-ACC pour.TR-ASP-*koɭ*-PST-3FSG
'Mansi poured the milk for herself.' —or—
'Mansi poured the milk on herself.'

The simple transitive sentence in 15a has the straightforward meaning that Mansi poured milk. The addition of *koɭ* to the verb in 15b adds the information that the result state of the milk-pouring comes to be evaluated from Mansi's mental or physical Locus space. Informally, we get the reading that Mansi either poured the milk *for* herself (mental Locus: self-benefactive reading), or that she poured it *on* herself (physical Locus: physically self-affected reading).

In a sense, therefore, *koɭ* associates the highest argument with an extra semantics, much like an additional θ-role. We adopt Sundaresan (2012)'s proposal that *koɭ* is thus a (thematic-)raising predicate (along the lines of

Ramchand, 2008), which raises the highest argument in the event predication in its scope to its Spec.[9] As Sundaresan shows, this raising property is crucial in explaining the obligatory presence of *ko*ɭ in typical cases of co-argument binding in Tamil. Briefly, the addition of *ko*ɭ to a sentence like (9) allows the agentive DP to raise to a position where it can scope over the entire event predication, giving it the right perspectival semantics required to antecede the anaphor. Anaphoric antecedence in Tamil is perspective-driven as it is (at least in part) in many other languages, including Japanese, Norwegian, Icelandic and Italian (see Kuno, 1987; Sells, 1987; Hellan, 1988; Sigurðsson, 1990; Bianchi, 2003). I.e. the antecedent must denote a sentient individual in the evaluation context that holds a particular mental/spatio-temporal perspective with respect to the minimal predication containing the anaphor. This explains why, in structures where such a perspectival semantics is already encoded as part of the verb meaning, as with psych-predicates, the use of *ko*ɭ is prohibited and co-argument binding can obtain in its absence:

(16) Raman$_i$ tann-æ$_{\{i,*j\}}$ virumb-in-aan/*virumbi-kko-ɳɖ-aan.
Raman ANAPH-ACC love-PST-3MSG/*love-*ko*ɭ-PST-3MSG
'Raman$_i$ loved himself$_{\{i,*j\}}$.'

[9] Alternatively, we could analyze it as a kind of control predicate, as has been proposed for adjectival passives by Bruening (2014).

6.1 Crosslinguistic concerns for *ko⌊* = Mid

As discussed in detail by Sundaresan, the distribution and meaning-contribution of *ko⌊* are, in fact, very similar to certain uses of *get* in English. We would add here that there are also clear parallels to middles in languages like Ancient Greek.[10] Anagnostopoulou and Sevdali (to appear) report the following uses of the AG middle: direct (17a) and indirect reflexivity (17b), reciprocals and examples like (18b), where, in contrast to (18a) "[t]he presence of middle morphology... does not signify a change in the argument structure of the verb, but rather the fact that the subject is affected to a greater extent by the action denoted by the predicate" [p. 13].

(17) a. Ho stratio:te:-s lou-etai.
the soldier-NOM wash-MID.PRES.3SG
'The soldier washes himself.'

b. Ho stratio:te:-s lou-etai ton hippo-n.
the soldier-NOM wash-MID.PRES.3SG the horse-ACC
'The soldier washes the horse (for his own interest).'

(18) a. Ho Achilleu-s pher-ei to depa-s.
the Achilles-NOM carry-ACT.PRES.3SG the goblet-ACC
'Achilles is carrying the goblet.'

b. Ho Achilleu-s pher-etai to depa-s.
The Achilles-NOM carry-MID.PRES.3SG the goblet-ACC
'Achilles is carrying the goblet (as a winner)/Achilles wins the goblet.'

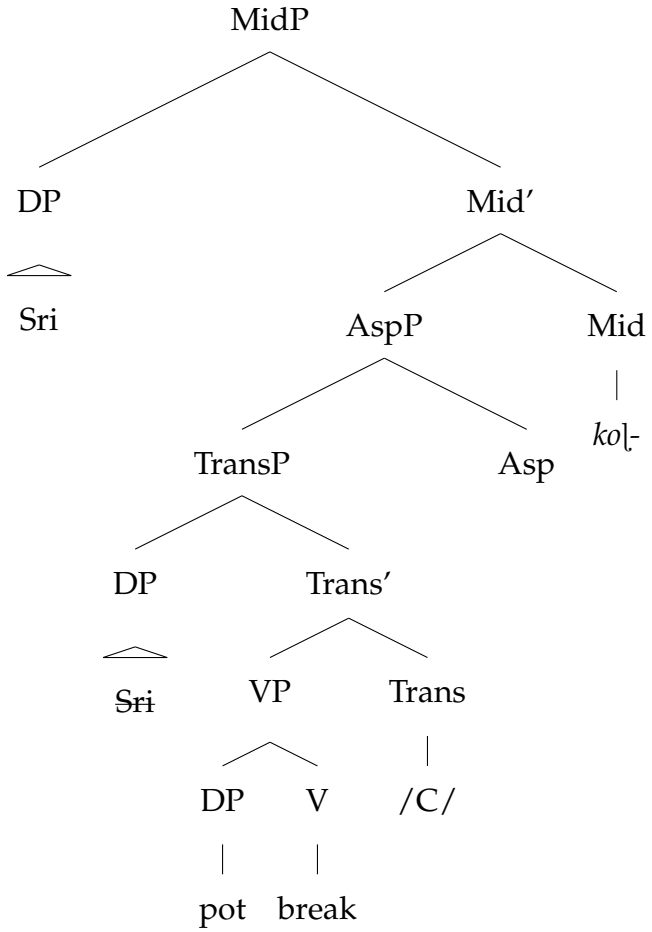
[10] We thank Elena Anagnostopoulou (p.c.) for drawing our attention to the extent of the parallels with Ancient Greek.

All of these are contexts where *ko*⌊ would be either required or felicitous in Tamil.

Thus assuming a real parallel here and regarding *ko*⌊ as a species of middle marker is justified, and it is for these reasons that we refer to the head that *ko*⌊ spells out as Mid(dle). Given this semantic contribution of *ko*⌊ and the tree in (13), we propose that a transitive *ko*⌊-sentence actually has the structure given in (19). Note that we posit a head Asp below Mid which derives the result state from the main predication below it and which shows up as the morphological marker labeled as ASP in the examples above between the verb root and *ko*⌊ (e.g. the *-čču-* in 12 and the *-i-* in 15b).[11]

[11]We are not saying any more about this marker here (and in particular are not including it in our main discussions of the hierarchy of projections needed in the *v* domain, e.g. in Section 8 below) because we are still working on its proper analysis, e.g. whether it is always present or only appears in combination with *ko*⌊ and how it relates to the homophonous 'past' tense marker. See Amritavalli and Jayaseelan (2005) for some relevant discussion. Furthermore, Sundaresan (2012) argues for the presence of an additional Persp(ective) head between Asp and Trans, which plays a crucial role in anaphora and other perspective-related phenomena. We are leaving it out here simply to keep the trees from getting too complicated.

(19)



7 Putting it all together

We have now seen two markers that appear above Trans — the passive morpheme *paɖ-* and the middle morpheme *koɭ*. The question now is how these two interact with each other. One possibility — indeed a reasonable one, given the idea that passive and middle are alternative values of a broader category of 'voice' in the traditional sense — is that they are two competing instantiations of the same structural position, which would be

something like a single Voice head above our Trans. If this is the case, then they should not be able to co-occur. The other logical option, of course, is that they realize distinct structural positions. If this is correct, then it should be possible to have both *ko*̩ and the passive marker co-occur in a single verb form.

As it turns out, they *can* in fact co-occur, as demonstrated in (20a):

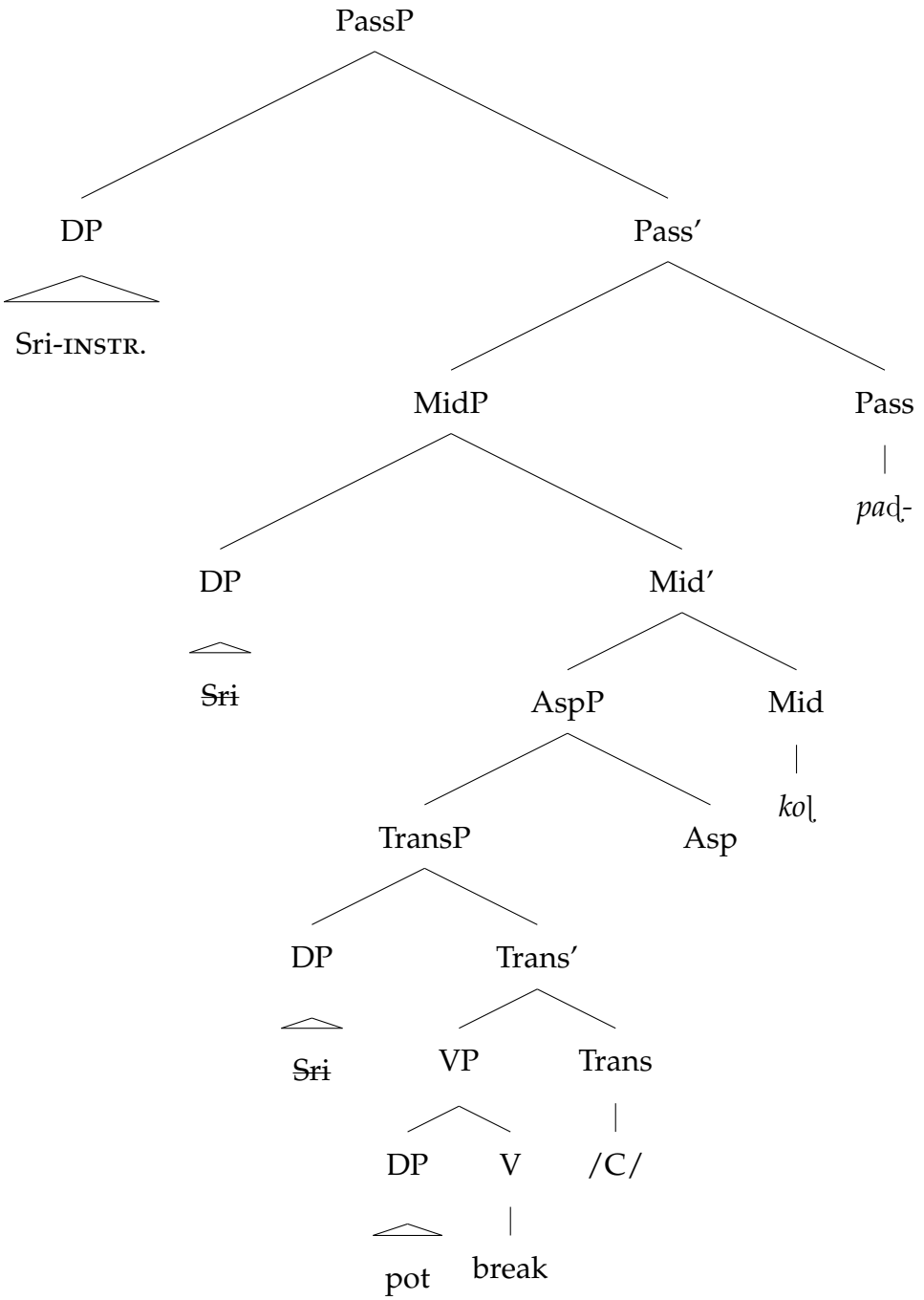
(20) a. Paanæ Sri-aal oḍæ-čč̆ŭ-**kko**̩̩**æ-pa**ṭ-ṭ-adŭ̆.
Pot Sri-INST break-TR-MID-PASS-PST-3NS
'The pot got broken by Sri.'

b. * Paanæ Sri-aal oḍæ-čč̆ŭ-**pa**ḍ**æ-kko**-ṇḍ-adŭ̆.
Pot Sri-INST break-TR-PASS-MID-PST-3NS
'The pot got broken by Sri.' (Intended)

The possibility of a sentence like (20a) shows clearly that *ko*̩ and *pa*ḍ must realize distinct heads in the *v* domain. Furthermore, we see that their relative ordering is fixed: *ko*̩ must linearly precede the passive morpheme as in (20a), never succeed it as in (20b). Again, according to the Mirror Principle this indicates that Pass is higher in the syntactic structure than Mid, which in turn is higher than Trans.

We thus have a (nearly) final picture involving an even higher degree of articulation in the *v* domain, as illustrated by the structure for (20a) in (21) below:

(21)



Note that the DP that gets associated with the *koɭ* reading — i.e. the one

that is thematically raised into its specifier — is the AGENT 'Sri', not the THEME 'the pot'. This is what we would expect based on minimality if the external argument is still introduced in SpecTransP even in the passive, and moves from there to SpecMidP. Deriving this fact would be quite a bit more difficult, at the least requiring some additional assumptions, under a theory where the normal external argument position is empty and 'by'-phrases are projected elsewhere, e.g. in SpecPassP.

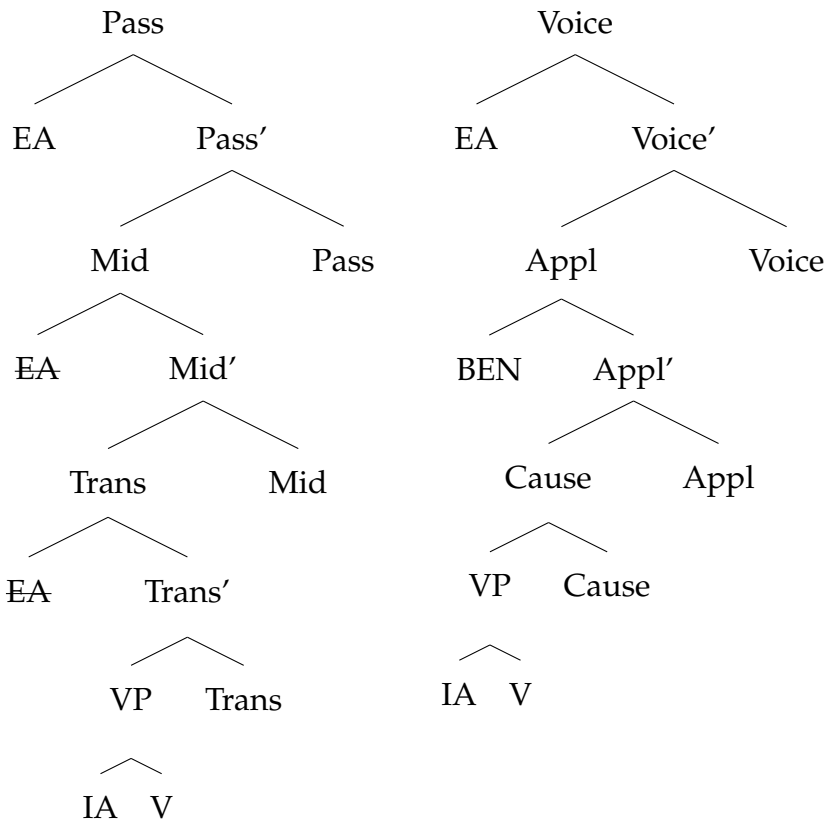
A puzzle for what we have posited in 21 is how the determination of instrumental case-marking works. Since the agent is only marked instrumental in the passive, the case presumably **is** associated with Pass and not with Trans, yet it has the hallmarks of an inherent case rather than a structural one, so we would not expect it to be assigned after movement from SpecTransP (potentially via SpecMidP) to SpecPassP. An alternative would be to assume that instrumental-marked agents **are** introduced in SpecPassP, being assigned the inherent instrumental case in much the way that inherent dative case is assigned in many languages to arguments introduced the SpecApplP. This would then control or bind an obligatorily null DP which is first-merged in SpecTransP and then raises to SpecMidP when this is present. This solves the case puzzle, but raises the question of what sort of null DP it is that is appearing here (pro, PRO, something else?) and what enforces the requirement — which must come from PassP — that it obligatorily be this special null element rather than an arbitary DP.

8 Comparison: motivating a Cause head

We thus have relatively straightforward morphosyntactic evidence for three distinct heads in a fixed hierarchy in the v-layer in Tamil: Pass > Mid > Trans. The interesting question to ask then is how this fits in with the Voice > Cause heads identified by Pylkkänen (2002), Alexiadou et al. (2006) and others. On the left, in (22a), we have our proposed structure for Tamil, and on the right, in (22b) Harley (2013)'s structure for Hiaki, which augments the basic Voice > Cause structure with the addition of an Applicative head.

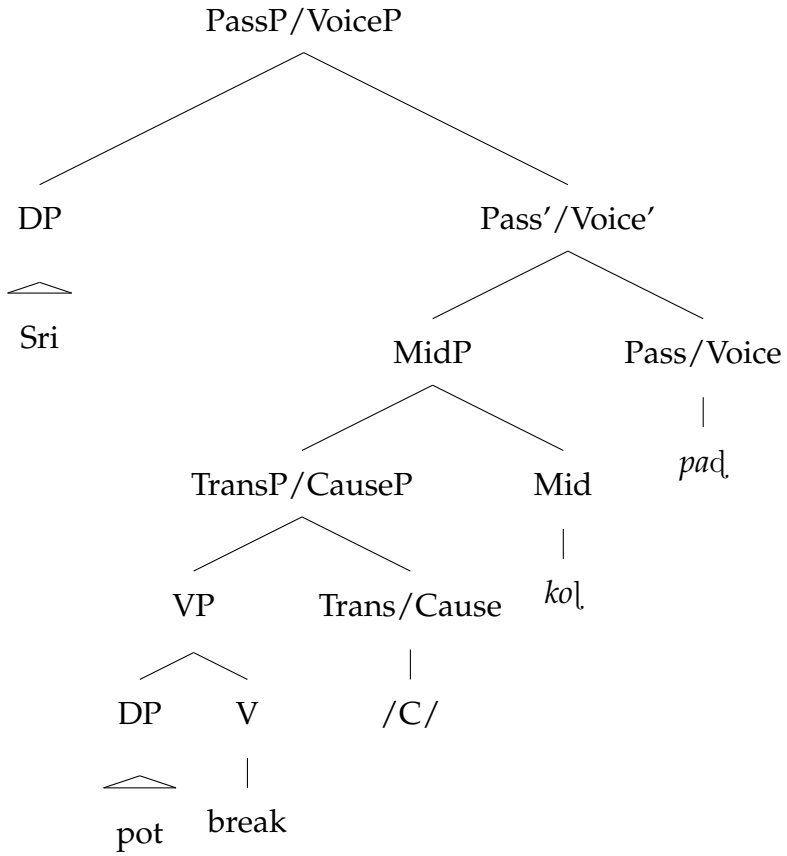
We won't be too concerned about Mid and Appl in the subsequent discussion, since we don't have clear evidence at the moment about where to place them relative to each other. Instead, we'll focus on our Pass and Trans, and their Voice and Cause.

(22) a. Current Proposal for Tamil b. Harley (2013) for Hiaki



Now, an obvious possibility would be that our Pass is really just their Voice, and our Trans is really just their Cause. External Arguments would then actually introduced be by Pass/Voice (as proposed by Pylkkänen etc.) rather than by Trans/Cause, i.e. like this:

(23) Possible structure (to be rejected):



However, we think there are good reasons to reject this idea, illustrated by the tree in 23 above. First, as discussed above, *ko* has a thematic relationship with the sole argument of unaccusatives, but with the EA of transitives. This can be straightforwardly explained in our account by having the minimally closest DP raise to the specifier of MidP, as we've proposed, but for this to work, the EA must be introduced lower than Mid, hence not in Pass/Voice. Second, the alternation instantiated in

our Trans head really tracks transitivity, not the causative/eventive semantics attributed by Pylkkänen et al. to their Cause head. So *oḍæ-nʧ-* 'break (intr.)' and *oḍæ-čč-* 'break (trans.)', for example, are alike in the fact that they are verbal, eventive and involve a change-of-state, i.e. both are 'causative' in the relevant sense. The difference is in whether there is an agent responsible for the causative event, and this is a matter for the higher Voice head of Pylkkänen and the others, not for their Cause head.

These considerations lead us to propose instead that our three heads essentially expand what previous researchers had identified as a single Voice head, and thus that all three of them come on top of their Cause:

(24) Pass > Mid > Trans > Cause > $\sqrt{\ }$

Indeed, there is solid evidence for an additional Cause head in Tamil below the three heads that have been the focus of the earlier sections of this paper, which comes from constructions built out of a light verb above a nominal or adjectival root. Consider the following examples:

(25) a. Saṭṭæ segappŭ
shirt red
'The shirt is red.'

b. Saṭṭæ segapp-aag-i-ččŭ
shirt red-LV.INTR-PST-3NSG
'The shirt became red'

c. Naan saṭṭæ-jæ segapp-aakk-in-een
I shirt-ACC red-LV.TR-PST-1SG
'I made the shirt red.'

(25a) shows that a basic adjective like *segaappŭ* 'red' can be used uninflected and without a copula as the main predicate of a clause, with a purely stative meaning. In (25b), we see an inchoative built on this by the addition of *aagŭ* (which we are glossing for the moment simply as LV for light verb), translatable roughly as 'become'. Being verbal, *aagŭ* is inflected for tense and subject agreement. We can then make a transitive version of this as in (25c), showing the regular transitivity alternation characterized by gemination, in this case applying to the light verb, which again inflects for tense and agreement. Now, if the gemination is the realization of Trans, and *segappŭ* is the root, then it is reasonable to take the light verb *aagŭ* to be the realization of a Cause head with much the same properties as posited by Pylkkänen (2002); Alexiadou et al. (2006, etc.).

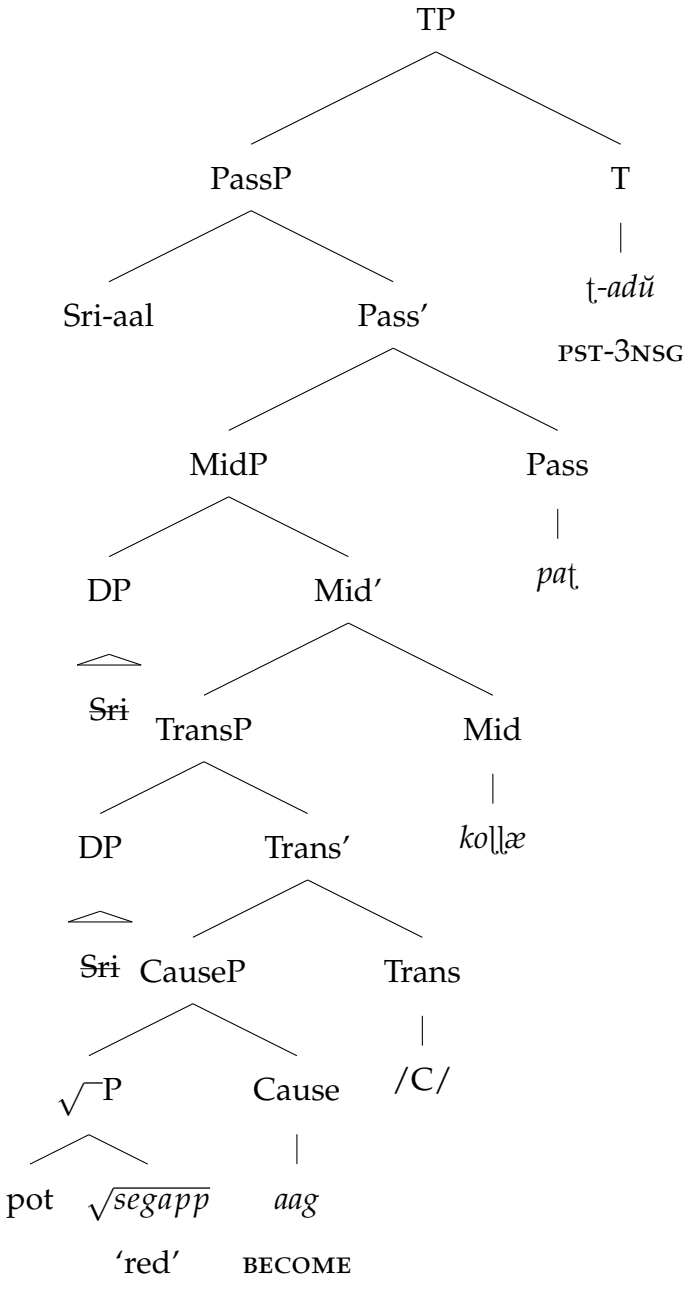
That is, it verbalizes the structure and introduces the eventive semantics, building a change-of-state out of a pure state, with the specification that the change-of-state is caused, but crucially not indicating that this cause necessarily involves an Agent. The addition of an Agent comes only with the addition of the Trans head. More accurate glosses for the verb forms in (25b) and (25c) above, would thus have Cause in place of LV. Once again, all of the morphemes of interest here can be strung together in a single verb form — in exactly the order predicted under the analysis above:

(26) Sri-aal paanæ segapp-aakk-i-ko|̣|æ-pa|̣-|̣-adŭ.
Sri-INSTR pot.NOM red-CAUSE.TR-ASP-*ko*|̣-PASS-PST-3NSG

'The pot was gotten reddened by Sri.' (rough translation)

The sentence in (26) thus shows the full articulation of the functional spine above the root — $\sqrt{}$-Cause-Trans-Mid-Pass(-T-Agr) — as shown in tree form below:

(27)


TP
PassP
T
ʈ-adŭ
PST-3NSG
Sri-aal
Pass′
MidP
Pass
paʈ
DP
Mid′
Sri
TransP
Mid
koɭɭæ
DP
Trans′
Sri
CauseP
Trans
/C/
√P
Cause
pot
√segapp
‘red’
aag
BECOME

9 Summary and outlook

We have presented what we take to be clear evidence, morphological, syntactic and semantic, that there are at least four distinct syntactic functional heads in the region between the (verbal) root and the T head in Tamil. Starting just above the root and moving upwards, these are a Cause head, responsible for eventive, change-of-state semantics; a Trans head, which introduces agentive semantics and the external argument; a Mid head, which indicates that the result state of the predicate is evaluated from the mental or physical space of the highest argument; and finally a Pass head, which suppresses or demotes the agentive argument. This builds on prior work that first split off an agent-introducing functional head from the main verb and then subsequently split this further, to distinguish a Cause head responsible for causative-eventive semantics and a Voice head responsible for introducing the external argument as well as traditional 'voice' phenomena. We have argued that it is this latter Voice head which must be exploded so that the role of agent-introduction is separated from those of passivization and middle formation. The straightforward take-home message, much in line with recent developments, is that we need to recognize what might be called a v layer or domain, much like the C layer and the Infl layer that have been the subject of much work in the cartographic tradition.

The findings we have reported here raise a number of interesting ques-

tions which remain open to be pursued in future work. Perhaps the most important ones are of a comparative nature. At the most fundamental level, while we think the evidence for the full articulation of the *v* layer in Tamil is quite secure, this does not necessarily imply that all languages will make full use of it. One could imagine e.g. that some languages 'bundle' two or more of the heads together as Pylkkänen (2002) suggested for Voice and Cause, or even that some languages simply lack a given projection entirely. Either of these might be a reasonable analytic strategy to pursue e.g. for languages that do not have a morphosyntactically distinct middle formation.

A related question is how seriously we should take the specific labels we've proposed for the functional heads here, in particular in the application of the hierarchy to the analysis of other languages. It is well known e.g. that what are called 'passives' do not have uniform properties cross-linguistically, and thus it may well be that the correct analysis of some 'passives' should not involve the Pass head at all. One specific constellation that we are interested in from this perspective are the various constructions built with auxiliary *get* in English, some of which behave similarly to passives, while others look quite similar to the examples with Tamil *ko*ļ, as noted above. It is to be hoped that the additional articulation of the *v* domain we have proposed here will be able to shed light on some such puzzles.

Yet another question, which commonly arises in work that posits ad-

ditional functional heads, is the extent to which the full range of heads in the v layer are present in all verbal structures within a given language, or whether at least some of them can be left off. More concretely, is the Pass head simply lacking in an active clause, or is it present, but with some specification making it essentially inactive for the purposes of morphology, syntax and semantics? It is actually quite difficult to find clear ways to test these alternatives, but the evidence presented here is of some relevance. It would have been in some sense easier to assume the view that all heads are always present if there were a single Voice head responsible for all 'voice' interactions, with different feature specifications deriving transitive actives, unaccusatives, passives and middles. In other words, there would be no need to assume a deactivated passive head in middles, but rather a middle specification of the Voice head instead of a passive one.

However, since it is clear that we must split up these various functions over distinct heads, in particular given the fact that they are not in complementary distribution but can indeed co-occur, such a simple assumption is no longer possible. We are forced, if we insist on the complete immutability of the functional sequence, to accept that certain functional heads can be present in the structure while apparently having no discernible effect at any grammatical level. Again, this does not necessarily mean that this is the wrong way to go, given the difficulty of finding clear empirical tests to distinguish between the two options, but it gives

us some additional clarity about what it means to say that a head, once posited, is always there.

There are also a number of important questions about the specific analysis of Tamil that we have not yet been able to settle. As we noted at a couple of points in the discussion above, there are a number of additional morphological markers that can appear in the string of suffixes on the verb which we have not analyzed in full, including the 'infinitive' suffix that comes inside of the passive and the marker that comes inside of *koḷ* which we have labelled Asp. It remains to be seen to what extent these bits of morphology can be correlated with substantive bits of syntax and semantics, in particular in the case of the former marker. In addition to these, there are several other aspectual forms and at least one apparent applicative suffix, *-kuɖu* (Sundaresan, 2006), all of which might be expected to be somewhere in what we are calling the *v* layer. It will be instructive to see how these interact with the other heads identified here, also for the purposes of working out how to fully integrate our results for Tamil with those of Harley (2013) for Hiaki, among others.

References

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer. 2006. The properties of anticausatives crosslinguistically. In *Phases of interpretation*, ed. Mara Frascarelli, 187–211. Berlin: Mouton.

Amritavalli, R., and K.A. Jayaseelan. 2005. Finiteness and negation in Dravidian. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 178–220. Oxford: OUP.

Anagnostopoulou, Elena, and Christina Sevdali. to appear. Case alternations in ancient greek passives and the typology of case. *Language* .

Annamalai, E. 1999. Lexical anaphors and pronouns in Tamil. In *Lexical and anaphors and pronouns in selected South Asian languages: a principled typology*, ed. Barbara C. Lust, Kashi Wali, James W. Gair, and K.V. Subbarao, 169–216. Mouton de Gruyter.

Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–415.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Temps et point de vue/Tense and Point of View*, ed. Jacqueline Guéron and L. Tasmovski, 213–246. Nanterre: Université Paris X.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.

Bruening, Benjamin. 2014. Word formation is syntactic: adjectival passives in English. *Natural Language and Linguistic Theory* 32:363–422.

Bye, Patrick, and Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In *The morphology and phonology of exponence*, ed. Jochen Trommer, 427–495. Oxford: OUP.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT

Press.

Christdas, Prathima. 1988. The phonology and morphology of Tamil. Doctoral Dissertation, Cornell.

Embick, David. 2004. Unaccusative syntax and verbal alternations. In *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, ed. Alexiadou et al, 137–158. Oxford: OUP.

Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125:34–57.

Hellan, Lars. 1988. *Anaphora in Norwegian and the theory of grammar*, volume 32 of *Studies in Generative Grammar*. Dordrecht: Foris Publications.

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*, ed. Johan Rooryck and Laurie Zaring, 109–137. Dordrecht: Kluwer.

Kuno, Susumo. 1987. *Functional syntax – anaphora, discourse and empathy*. Chicago: Chicago University Press.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.

Levin, Beth. 1993. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press.

Lidz, Jeffrey. 2001. The argument structure of verbal reflexives. *Natural Language and Linguistic Theory* 19.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Penn Linguistics Colloquium*.

Medová, Lucie. 2009. Reflexive clitics in the Slavic and Romance languages. Doctoral Dissertation, Princeton University.

Parsons, Terence. 1990. *Events in the semantics of English: a study in subatomic semantics*. Cambridge, MA: MIT Press.

Pesetsky, David. 1989. Language-particular processes and the earliness principle. Ms., MIT.

Pylkkänen, Liina. 2002. Introducing arguments. Doctoral Dissertation, MIT, Cambridge, Mass.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: a first phase syntax*. Cambridge: CUP.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Schäfer, Florian. 2008. *The syntax of (anti-)causatives. external arguments in change-of-state contexts*, volume 126 of *Linguistik Aktuell*. Philadelphia: John Benjamins.

Schiffman, Harold. 1999. *A reference grammar of spoken Tamil*. Cambridge: Cambridge University Press.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18:445–479.

Sigurðsson, Halldór Ármann. 1990. Long distance reflexives and moods in Icelandic. In *Modern Icelandic syntax*, ed. Joan Maling and Annie Zaenen, 309–346. New York: Academic Press.

Sportiche, Dominique. 1998. *Partitions and atoms of clause structure: subjects, agreement, case, and clitics*. Psychology Press.

Steever, Sanford B. 2005. *The Tamil auxiliary verb system*. London: Routledge.

Sundaresan, Sandhya. 2006. The argument structure of verbal alternations in Tamil. In *Proceedings of WCCFL 25*, ed. Donald Baumer, David Montero, and Michael Scanlon, 390–398.

Sundaresan, Sandhya. 2012. Context and (co)reference in the syntax and its interfaces. Doctoral Dissertation, University of Tromsø and University of Stuttgart, Tromsø.

Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 1:81–114.