Introduction to Distributed Morphology

Externalization of Thought in an Ideal World: the Humboldt's ideal

Assuming that there is a universal conceptual inventory including predicative contents and functional features or categories, in an ideal world the externalization of thought (Berwick and Chomsky 2016) could be achieved in the following way:

Each unit of this inventory is associated with an index—an externalization medium. Let's call it an exponent.

The unit composed of a conceptual unit and its matching exponent is a **Saussurean sign**. Let's call it a morpheme.

Under this definition, morpheme are the smallest meaningful unit of grammar, in other words, the minimal unit of phonology (an exponent): meaning (a conceptual unit) correspondence.

Externalization of Thought in an Ideal World: the Humboldt’s ideal

The narrow lexicon of grammar contains a list of morphemes.

Combinatorial rules could combine them transparently into sentences.

This would be a maximally “simple” system, where there would one-to-one form: meaning correspondence.

Such a language would satisfy what is traditionally referred to as the Humboldt’s ideal. In such a language, sentences are directly constructed by composition of morphemes through syntactic merge.

This would be an ideal fully transparent, analytic language.

In this ideal language, generation of complex conceptual structures is achieved by direct composition of morphemes by syntax.

THE REAL WORLD

Unfortunately, real languages deviate from Humboldt's ideal in many ways.

I. Words

A first deviation from this ideal analytic language is the existence of intermediate units between the morphemes and the sentence: the words.

One could hypothesize that the faculty of language requires the grouping of morphemes into sensori-motor units that can be pronounced and stored in the auditory system.

Once we have words, in addition to morphemes and syntax, we need an operation that generates words; let's call it word formation (m-merger, head movement).

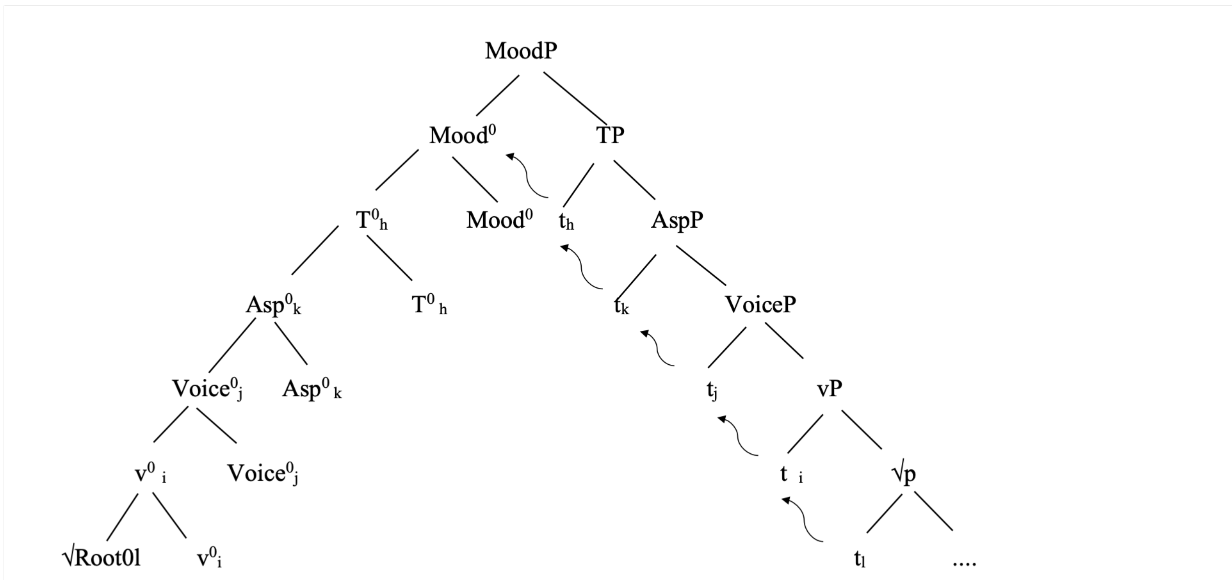
Ideal" word formation generates an "ideal" fully transparent agglutinative language.

Both the "ideal" analytic and agglutinative language satisfy what can be called the Humboldt's ideal of one form –one meaning.

Word-Forming Head Movement

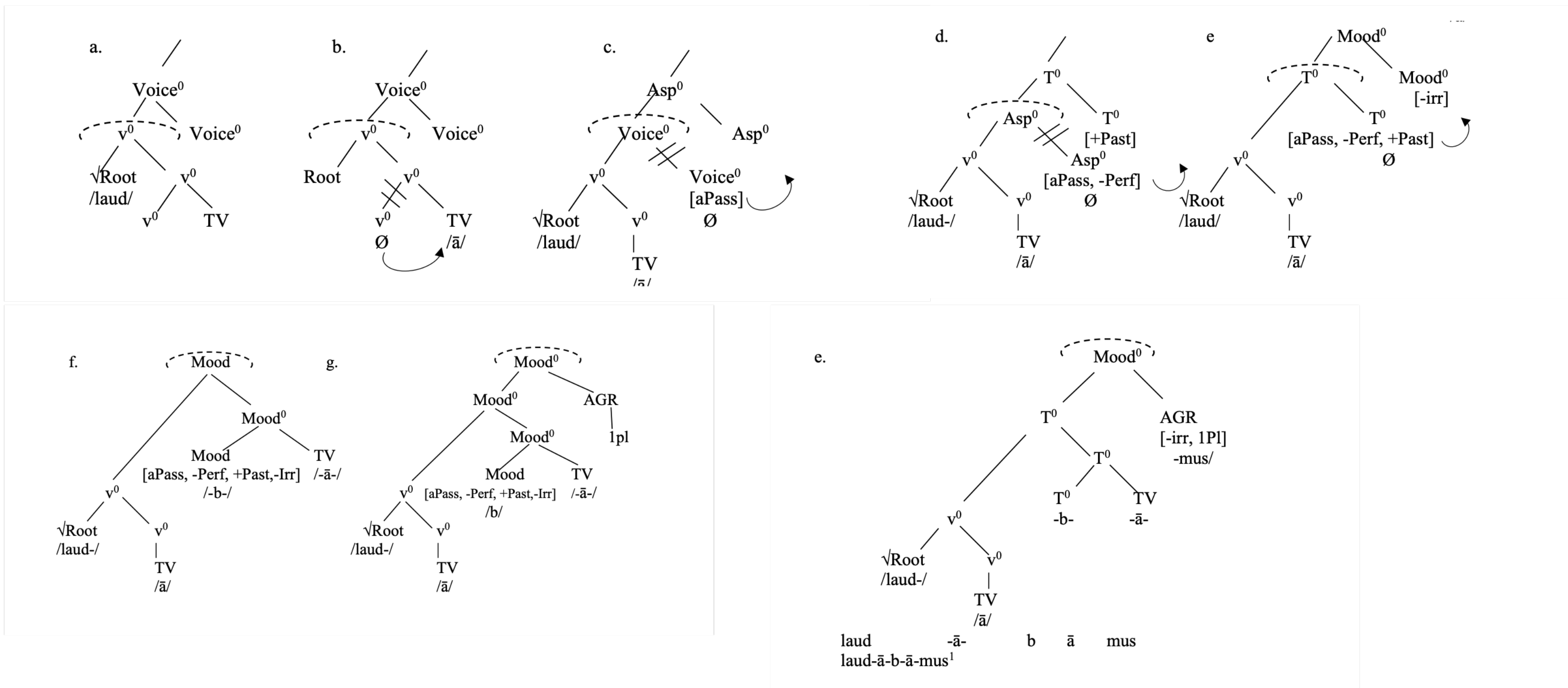
The universal hierarchical structure formed by the combination of (10) and (11) is then mapped onto surface morphological units (i.e., X^0-complexes; Embick and Noyer 2001) via cyclic, iterated head-rollup movement of Root through *v*, Asp, T, M. Such a procedure will create the structure in (12) (For graphic simplicity, I will omit the CausP until it is relevant for the analysis further below):

(1)



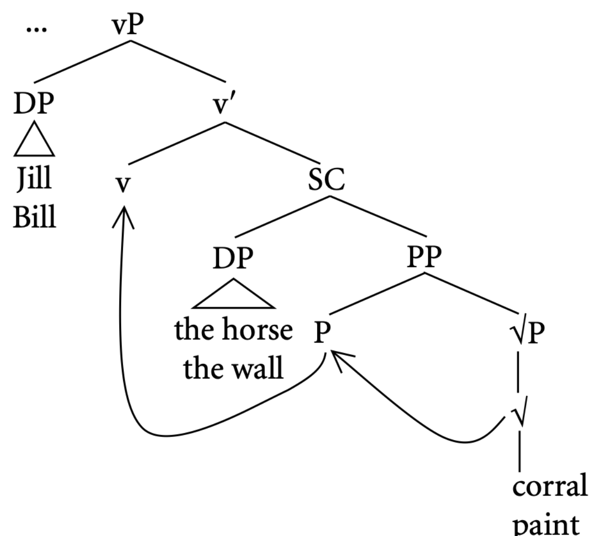
Latin Imperfect indicative *Laudābāmus*

(2)



Lexical syntax

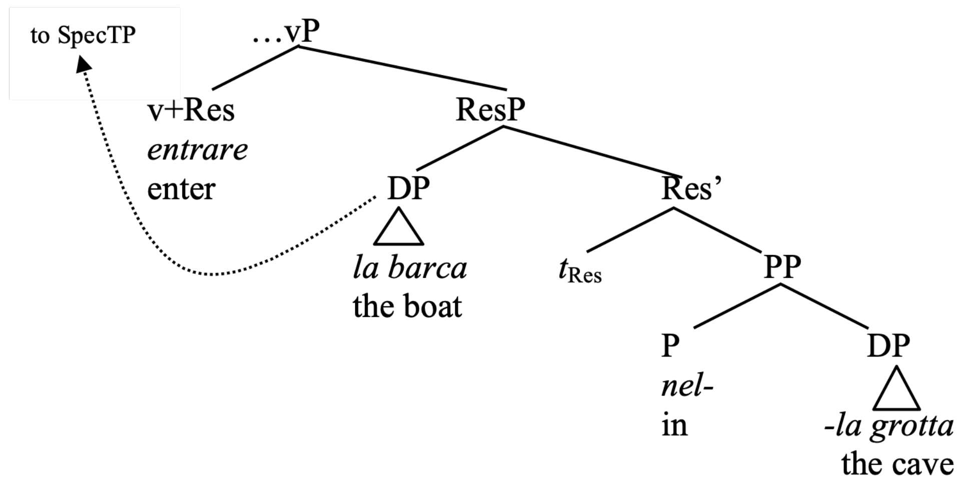
(3)



‘Jill corralled the horse.’
‘Bill painted the wall.’

Lexical syntax

(4) b. Italian



La barca entró nella grotta
The boat entered in.the cave

THE REAL WORLD

In the real world, unfortunately, and perhaps more interestingly, many more deviations from this ideal state of language can be found, especially in word morphology.

II. Portmanteau/fusional morphology

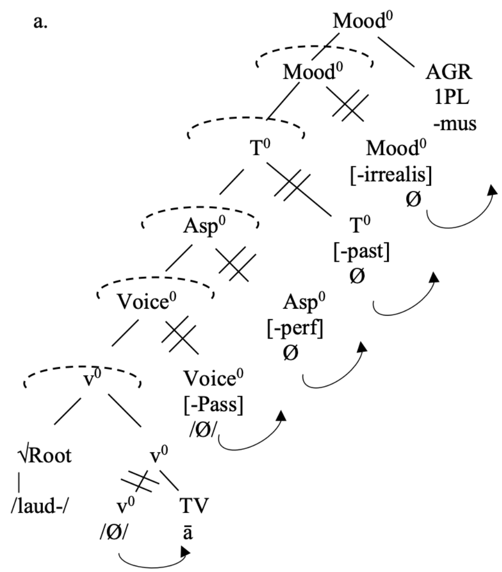
If in a pure "ideal" agglutinative language, there is a biunivocal relation between morphological categories and exponents (see case and number in Turkish: *adam-lar-i*n "man-PL-GEN"), there are many languages in which a single exponent is associated with more morphological categories (Portmanteau morphology in Fusional languages, see case and number in Latin: *homin-um* "man-PL.GEN").

The derivation of Latin verbal form *laudāmus*

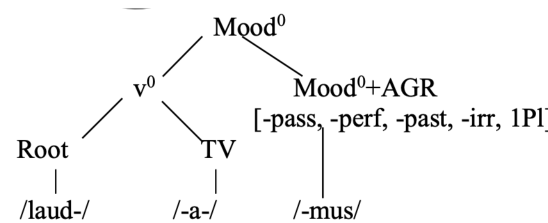
The Present Indicative form *laudāmus* is derived in (58a); the output is in (58b).

(5)

a.



b.



THE REAL WORLD

III. Multiple exponence

There are also cases in which what is a single cumulative exponent in one situation becomes multiple in another situation. Furthermore, there are also cases in which a single meaning has multiple exponents in the same word (e.g. circumfixation).

IV**. Process morphology**

There are also cases of process morphology (ablaut, reduplication, truncation). Process morphology violates Humboldt's ideal insofar as in this case meaning is not expressed through transparent addition of an exponent but through phonological modification of that exponent. This also occurs in the case of templatic morphology where meaning is expressed by changes in the syllabic structure of the stem. One result of this type of morphology is another instance of extended exponence in which a single meaning is associated with multiple elements: the process and a concomitant independent exponent, a worse violation of Humboldt's ideal.

V. Ornamental morphology

There is also Ornamental Morphology. In the case of ornamental morphology, we have exponents, which are meaningless, i.e., they are not associated with predicative contents or functional features. Typical instances of this type of morphology are the thematic vowels of Latin and the Romance languages.

THE REAL WORLD

V. Infixation

A different type of violation of Humboldt's ideal is infixation where the discreteness of morpheme boundaries is violated by inserting a morpheme inside another morpheme, like a root, for example.

VI. Syncretism and inflectional synonimy

Another phenomenon that is not expected according to this ideal is syncretism: homophony in inflection, in which the same exponent is associated with more meanings. At the same time there is inflectional synonymy, i.e., situations in which a single meaning is associated with many exponents, cf. the English plurals, where in addition to regular *dog/dogs, cat/cats, kiss/kisses*, there are also *ox/oxen, fish/fish, foot/feet, alumnus/alumni, datum/data, cherub/cherubim, octopus/octopi* or *octopodes*.

VII. Contextual allomorphy

Another violation of the Humboldtian ideal is the existence of contextual allomorphy. In an ideal Humboldtian world, the exponent of a morpheme should always have the same shape regardless of the context in which the morpheme occurs: exponents may have alternant or varying pronunciations in different morpho-phonological contexts. In most languages the exponent of the same morpheme can appear in various phonetic shapes in different words, i.e. morphemes may have more than one phonetic realization.

Realizational models of PF

All of these violations of the Humboldt's ideal have recently led to the development of realizational models of PF.

Given the many-to-many relations that are possible between exponents and meaning, it has been proposed to dissociate the one-to-one relation between exponents and meaning units.

Under this view, the structuring of exponence is independent of the structuring of meaning: syntactico-semantic structures are generated independently of the exponents.

The insertion of exponents and changes in exponence are governed by principles proper to the PF component and independent of those of the syntactico-semantic component.

There is therefore an immediate account of the mismatches between the syntactico-component of language and PF characterizing the violations of Humboldt's ideal.

In a realization model, the generation of syntactico-semantic structures involves only the composition of units including abstract categories or features without any phonological content. Phonological content is inserted only later through vocabulary insertion, and subsequent morphophonological and phonological processes.

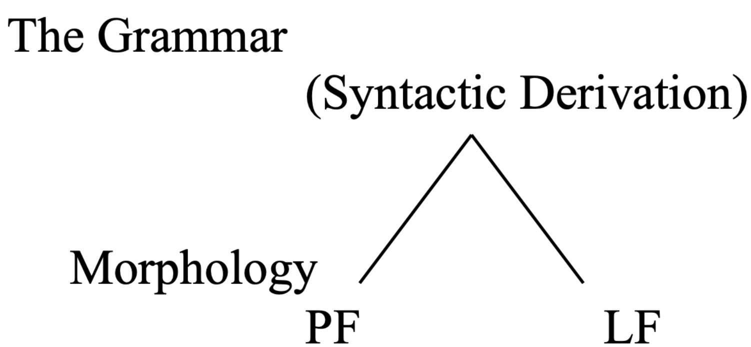
The theory of Distributed Morphology (Halle & Marantz 1993)

The theory of Distributed Morphology (Halle & Marantz 1993)

The input to morphology are syntactic structures where morphosyntactic and semantic features (or feature bundles) are distributed over nodes forming morphemes.

Morphology manipulates these syntactic structures and eventually converts them into linear sequences of phonological representations.

(6)



Distributed Morphology

The derivation of all morphological forms then takes place in accordance with the architecture given above.

2. A language chooses a subset of these features to use in grammar.

3. The language further chooses a subset of this subset for the computational system of syntax and `decides how to package them into the terminal nodes of the syntax (for Chomsky, "lexical items", for Distributed Morphology 1999, "morphemes").

4. These morphemes contain only the features relevant to the computational system of syntax and so lack phonological and purely morphological features, which are inserted in the (morpho)phonology

5. These morphemes are combined into larger syntactic objects, which are moved when necessary (Merge, Move).

6. Words, i.e., X^0-complexes, are generated by head movement operations. These X^0-complexes are the (abstract) morphosyntactic representations which are the input to phonological spell out.

7. During phonological spell out, phonological realizations are assigned to the terminal nodes via the cyclic application from the inside out of rule of exponence.

BASIC ASSUMPTIONS IN ALLOMORPHY

(7) Rules of exponence:

i) Vocabulary items: each vocabulary item includes a phonological exponent and an associated set of features that governs its insertion in the terminal nodes of the morpho-syntax.

ii) Morpho-phonological (MP) rule: morpho-syntactically conditioned phonological rules (also called Readjustment Rules in Distributed Morphology).

iii) Plain phonological (P) rules and repairs triggered by phonological filters.

These rules eventually determine the surface allomorphy of words.

MORPHEMES & TERMINAL NODES

The terminal nodes provided by the syntax consist exclusively of morphosyntactic/semantic features and lack phonological features.

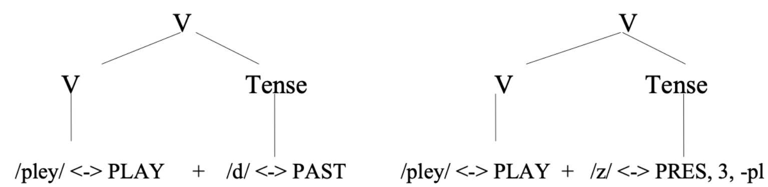
The morphosyntactic features at this level are drawn from a set made available by UG. As mentioned before, he semantic features and properties of terminal nodes created in the syntax will also be drawn from UG and perhaps from language-particular semantic categories and concepts.

BIPARTITE COMPOSITION OF MORPHEMES

Thus, morphemes have two parts: a set of grammatical features and a set of phonological features. We know that the suffixes /d/ and /z/ represent the past and present tense, respectively and we know from syntax that Tense is a terminal element

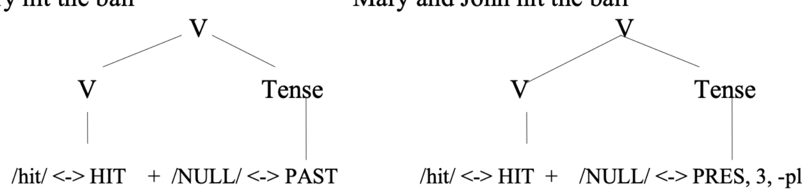
(8) Mary play-ed in the yard Mary play-s in the yard

(9)



(10) Mary hit the ball Mary and John hit the ball

(11)



MORPHEMES & TERMINAL NODES

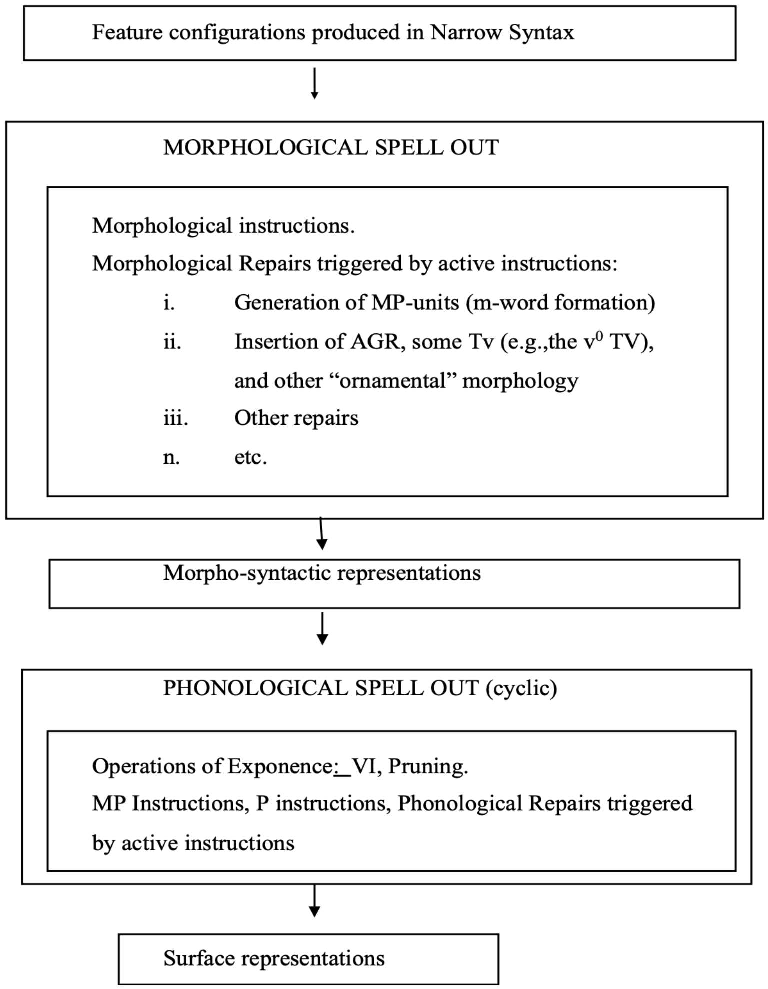
The morphemes in (9) and (11)--i.e., the terminal nodes of the trees-- are composed of two parts, a sequence of phonemes and a set of grammatical features.

These two parts of the morphemes play a radically different role in the syntax than they do in the phonology. As far as the syntax is concerned, it makes no difference whether the phonetic marker of PAST is /d/, /t/ or NULL: the syntax is only interested in the grammatical properties of the morphemes and it has no interest in the phonetic markers.

The phonetic markers of morphemes are of course crucial for the phonology, and it is the morphology that is the bridge between the syntax, which has no interest in the phonetic properties of morphemes and the phonology, where the phonetic properties of morphemes occupy center stage.

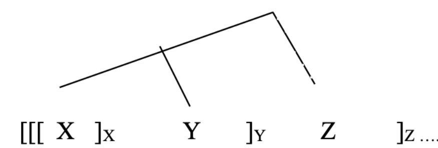
DM Architecture

(12)

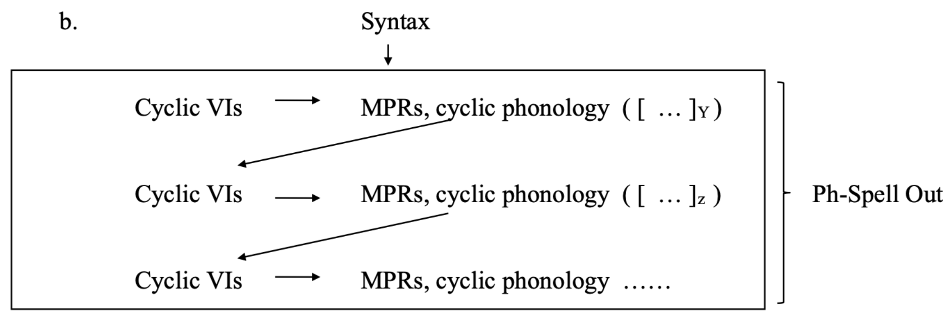


Vocabulary Insertion and all other rules of morpho-phonological exponency: morpho-phonological rules and phonological rules are sensitive to morphological boundaries apply cyclically:

(43) a.



b.



Verbal Functional Structure

The verbal functional structure below expresses the basic core temporal, aspectual, and modal structure of eventualities :

(13) $[_{MoodP}$ $Mood^0$ $[_{TenseP}$ T^0 $[_{AspP}$ Asp^0 $[_{VoiceP}$ $Voice^0$ $[_{vP}$ v^0 $[_{\surd p}$ $\surd Root^0$]]]]]]

[+/-irrealis] [+/-past, +/-future] [+/-perfect., +/-result.] [+/-passive]

The Vedic verbal system

Vedic Sanskrit verbs are traditionally described as being organized into ‘systems’ based on the different “stem” forms. The basic stem systems are the so-called Present, the Perfect and the Aorist:

(14)

	Present	Aorist	Perfect
perfective	–	+	+
stative	–	–	+

Other traditional stem systems are the future, the passive, the causative, the desiderative, the denominative, and the intensive (not discussed here)

Basic analytical steps

After collecting and examining the verbal forms of the language, the semantic features that determine their uses must be established.

Assuming a universal set of semantic features, the features that are contrastive in the morpho-syntax of the language are established, where a feature X is contrastive in a language L if there is at least a form in L whose distribution is determined by the features [aX].

We can say that this contrastive feature is morphologically realized in the language.

In the case of verbs, the set of the morphological realizations of the verbal semantic contrastive features of L define a paradigm for a given verb.

By the standard strategy of form comparison and subsequent extraction of the recurring parts with similar featural distribution (Nida 1970), a basic segmentation of forms can be achieved.

The Vedic verbal system

Each system includes mood distinctions: indicative, subjunctive, optative, and imperative, summarized in (7) The indicative mood moreover has Tense distinctions between present and past ([Past]). There are also Voice [NonAct] and Subject Agreement Person/Number distinctions.

(15)

Aspect	Present				Aorist				Perfect			
perfective	–				+				+			
stative	–				–				+			
Mood	Ind	Subj	Opt	Imp	Ind	Subj	Opt	Imp	Ind	Subj	Opt	Imp
irrealis	–	+	+	–	–	+	+	–	–	+	+	–
desiderative	–	–	+	+	–	–	+	+	–	–	+	+

The Vedic verbal system

The paradigms of the root *kar* ‘make’ in (8) illustrate the basic verbal forms for the three verb stems.

(16) The Vedic Sanskrit verbal system, √*kar* ‘do’, 3sg.act. forms

		Pres. stem		Perf. stem		Aor. stem	
Asp	[pfv]	–		+		+	
	[stat]	–		+		–	
T	[pst]	–	+	–	+	+	
		Pres.	Ipf.	Perf.	Pluperf.	Aor.	Mood
	Ind.	***kr̥-ṇó-ti***	***á-kr̥-ṇo-t***	***ca-kā́r-a***	***á-ca-kar-/t/*** [***ácakar***]	***á-kar-/t/*** [***ákar***]	[−irr, −des]
	Subj.	***kr̥-náv-a-t***		*****cā-kár-a-t***		***kár-a-t***	[+irr, −des]
	Opt.	***kr̥-nu-yā́-t***		*****ca-kr-[i]-yā́-t***		***kr-[i]-yā́-t***	[+irr, +des]

Alternations due to accentuation and zero grade

The forms of this table display a complex surface allomorphy: the root appears as [kar/kr̥/kr] and the internal suffix as [no/nu/nau].

In order to understand this pattern, one needs to look into Vedic Sanskrit morphophonology, and in particular the phenomena of accentuation and [a]-syncope.

A typical example is the alternation between *kr̥-nó-t-i*/ *kr̥-nu-yā́-t* which also displays the mobile accentual pattern that is typical of this language. Here, the vowel -*a*- of the suffix -*nau*- (which surfaces as [no] by monophthongization and [nau] by resyllabification) is syncopated when an accented syllable such as –*yā́*- follows.

The alternant that displays the vowel (-*náu*- > -*nó*-) is traditionally called the "full grade", and the alternant where the vowel is syncopated (-*nu*-) is called the "zero grade"

The relationship between accentual shifts and full/zero grade will not be dealt here (See Calabrese & Grestenberger (Forthcoming).

The basic morphosyntactic structure of Vedic Sanskrit verbs

The basic morphosyntactic structure of Vedic Sanskrit verbs can be observed in the present subjunctive and optative forms in (7). The endings display not only person and number contrasts but also voice and tense distinctions..

(17) a. Surface structure of Vedic verbs

kr̥-ṇu-yā́-ta	b. *kr̥-ṇáv-a-dhve*
[[[[kar]$_{root}$ -nau] –yā] -ta]	[[[[kar]$_{root}$ -nau] -a] -dhve]
make-IPFV-OPT-ACT.2PL.PAST	make-IPFV-SUBJ-MID.2PL.PRS

In addition, T[+past] is marked by a prefixal vowel. We assume that this is a case of an antitropal realization of the T exponent as a prefix — when lexically marked as antitropal, the exponent of an affix is phonologically inserted in the side opposite to that expected by the morphosyntax, e.g., the affix is prefixal instead of being suffixal as regularly expected

(17) b. á-kr̥-ṇo-t

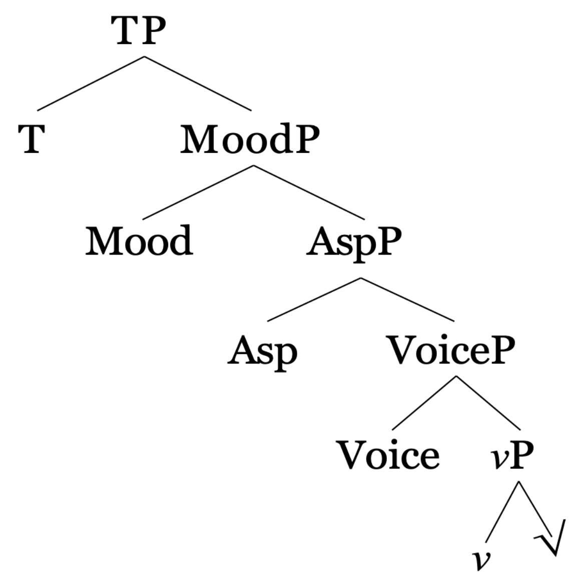
[[a [[kar]$_{root}$ -nau]] -ta]

PAST-make-IPFV-ACT.2PL.PAST

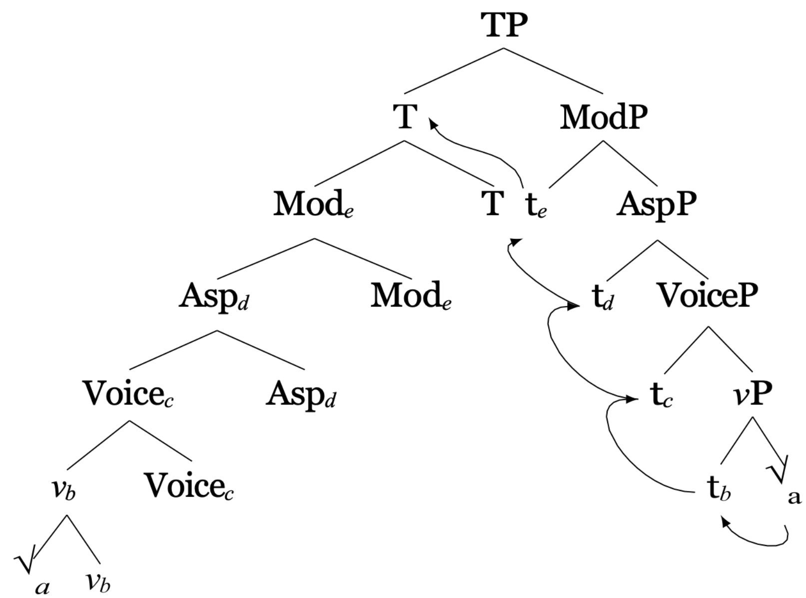
Complex heads via head movement

The universal hierarchical structure in (20) is then mapped onto surface morphological units (i.e., X0-complexes; Embick and Noyer 2001) via cyclic, iterated head-rollup movement of the root through *v*, Voice, Asp, Mood, and T, although we will not linger on the technical aspects of head-movement and their justification here. Such a procedure will create the structure in (21).

(17)



(18) **Head movement and synthetic verb formation**



The Synthetic morphology principle

Here I will implement the operation of word-formation in terms of the theory using filter and repairs.

In synthetic languages, functional heads tend to be realized as bound morphemes, as affixes. Calabrese (2019) assumed that this followed from the principle in (3) , which is parametrically active in synthetic languages.

(19) Synthetic morphology principle: A functional head Y^0 must be bound to a root or to X^0 complex including a root.

MORPHOLOGICALLY BOUND = PART OF A COMPLEX HEAD FORMED BY ADJUNCTION

I hypothesize that the affixal properties of functional heads follow from the interaction between this morphological requirement and the fact that they express a closed set of functional meanings (i.e., they are the exponents of a small set of functional features). If we also assume that category-defining X^0 nodes are also functional elements, we can also account for why they must be adjoined to roots.

Nota bene: An obvious question is how we know that we are dealing with a synthetic language at early stage sof the derivation before vocabulary insertion. A possible idea: roots are complex entities including VIs and diacritics. They are inserted early in the derivation probably already in narrow syntax. It is the roots that carry the language-identifying indices that determine the setting of parameters such as that in (3)

M-word formation as a repair

In Calabrese (2019) syntactic representations in violation of are repaired by m-word formation through the two operations in (from Harizanov and Gribanova (2018)

(20) A syntactic complementation relation $[X^0\ [_{YP} \ldots Y^0\ \ [_{ZP} \ldots]\ ]$ may be realized in the morphology as a complex head by:

a. **Head Raising**:

$[_{XP} \ldots X^0 \ldots [_{YP} \ldots Y^0\ \ [_{ZP} \ldots]\ ]\ ] \rightarrow [_{XP} \ldots [_{X^0}\ Y^0\ X^0]\ [_{YP} \ldots [_{ZP} \ldots]\ ]\ ]$

(where Y^0 and X^0 are heads, X^0 c-commands Y^0, and there is no head Z^0 that c-commands Y^0 and is c-commanded by X^0)

b. **Head Lowering**

$[_{XP} \ldots X^0 \ldots [_{YP} \ldots Y^0\ \ \ \ [_{ZP} \ldots]\ ]\ ] \rightarrow [_{XP} \ldots [_{YP} \ldots [_{Y^0}\ Y^0\ X^0]\ [_{ZP} \ldots]\ ]\ ]$

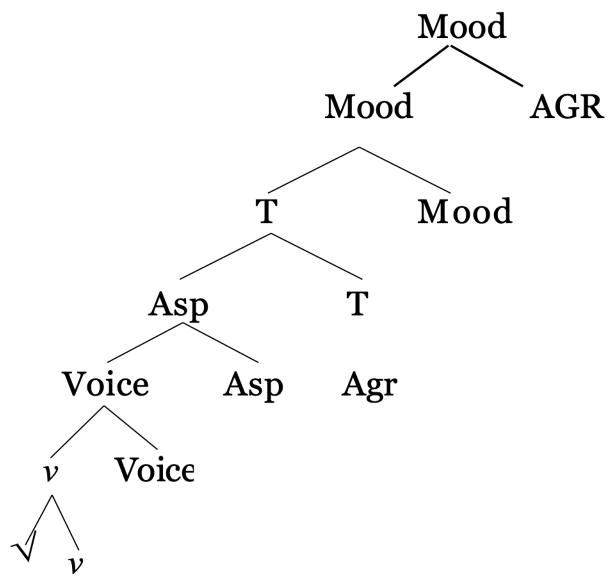
(where Y^0 and X^0 are heads, X^0 c-commands Y^0, and there is no head Z^0 that c-commands Y^0 and is c-commanded by X^0)

As all repairs, head raising and head lowering are ranked in order of application. Head raising is the preferred repair and applies first. If it fails to converge, and the derivation it starts crashes, head lowering applies.

AGR-Insertion

Following Halle and Marantz (1993) and Bobaljik (2008), the node AGR is inserted as an ornamental morpheme in the morphology by a rule that adjoins AGR to the highest X^0 in the complex verbal X^0 head. Application of this rule in the case of a complex head structure that is the output of (21) generates (22).

(21)



The structure in (22) then undergoes Vocabulary Insertion. The resulting structure undergoes VI, i.e. the procedure of association between (bundles of) morphosyntactic features (=functional nodes) and phonological content (exponence), which is assumed to occur cyclically from inside out (Bobaljik 2000).

A problem

Now there is a problem. The X^{0}-complex generated in (22) should include at least six affixes in addition to the root; only three pieces maximally appear in the actual forms in Table (5), as also shown in (23).

(22) a. *kr̥-nu-yā́-ta*
make-pres-opt-3sg.act ‘s/he would make’
b. *kr̥-ṇáv-a-t*
make-pres-subj-3sg.act
‘(so that) s/he make; s/he should make’

Against node syntactic inactivity

A first possibility is that the lack of overt realization of certain nodes is simply due to syntactic inactivity, i.e., the nodes that not overtly marked are simply absent in the morphosyntactic representation.

Embick and Halle (2005) and Halle (2019) proposed that unmarked features specifications are just syntactically absent in morphosyntactic representations.

A striking case discussed by Calabrese (2023) is the exponent of the Latin 1sg.act AGR node, which, putting aside the perfect /-ī/ (cf. laudāvī), is -ō in present and future indicative forms (laud-ō 'I praise', laudāb-ō 'I will praise', laudāver-ō 'I will have praised'), otherwise -m (laudāba-m 'I used to praise', laudāvera-m 'I had praised', laude-m '(that) I praise (subj.)', laudāre-m '(that) I praise (subj.)', laudāvisse-m '(that) I used to praised (subj.)', laudāveri-m '(that) I praised (subj.)').

The only way to characterize the distribution of -ō is to refer to the feature specifications [-past,-irrealis]. These are clearly unmarked, yet they must be present when Agreement exponence is inserted.

This could not be accounted for if the relevant nodes were not syntactically projected.

Zero-exponence (conventionally signaled as -Ø-) seems therefore unavoidable in sit- uations where distributional properties of exponents can only be accounted for by referring to phonologically silent (i.e., null) feature specifications (see Calabrese 2023 for further arguments).

Problems with surface zero exponence

On the other hand, adoption of zero-exponence leads to theoretical underdetermination, which allows for a proliferation of equally possible theories (among others, Anderson 1992, Marantz 1997, Trommer 2012, Dahl and Fábregas 2021).

To exemplify, consider the Sanskrit present suffix -nau-. It appears in root-adjacent position and competes with other pieces occurring in the same position; thus not only with the other present system markers in Table (5) and the "VP-shell" derivative suffixes such as the passive and the causative suffix (see...) , but also the aorist markers -Ø- and -s-, as we will see below.

Once the featural distribution of an item such as -nau- is established, one must determine its structural distribution: the insertion site of the exponent.

If the specification(s) belong(s) to a single terminal node, there is no issue: the exponent is inserted into this terminal node. The situation is more complicated if the Vocabulary Item includes specifications belonging to multiple terminal nodes. This is the case of -nau-.

In this form, at least three nodes are realized by the single piece -nau-; so, potentially only one node is overtly realized; the other two are silent. At its current version, DM in principle allows the practitioner to analyze -nau- as realizing any of the existing nodes above the root, (23), with no independent way to prefer one solution to the others (Pullum and Zwicky 1992):

(23) a. ... [Asp [Voice [v [kar] -nau]v -Ø]Voice] -Ø]Asp ...
b. ... [Asp [Voice [v [kar] -Ø]v -nau]Voice] -Ø]Asp ...
c. ... [Asp [Voice [v [kar] -Ø]v -Ø]Voice] -nau]Asp ...

Node Fusion

To account for situations like this, an operation of node fusion, which bundles nodes together, was proposed in early models of DM (e.g., Halle and Marantz 1993).

(24) ... $[_{\text{Asp+Voice+v}}$ [*kar*] -*nau* $]_{\text{v+Voice+Asp}}$ ...

Fusion, however, suffers from a look-ahead problem insofar as it must occur before VI so to allow insertion of the relevant fused exponent (Chung 2007, 2009).

Null-node pruning

An alternative operation that can directly connect fusion and zero exponents is null-node pruning.

This procedure consists of delinking nodes where Ø's are inserted when independently-motivated terminal nodes lack a phonological realization.

As Calabrese (2019) shows, null-node pruning is independently required to (a) simplify the phonological realization of morphosyntactic structures; and (b) account for the convergence of complex morphosyntactic structures and their simplex phonological realization.

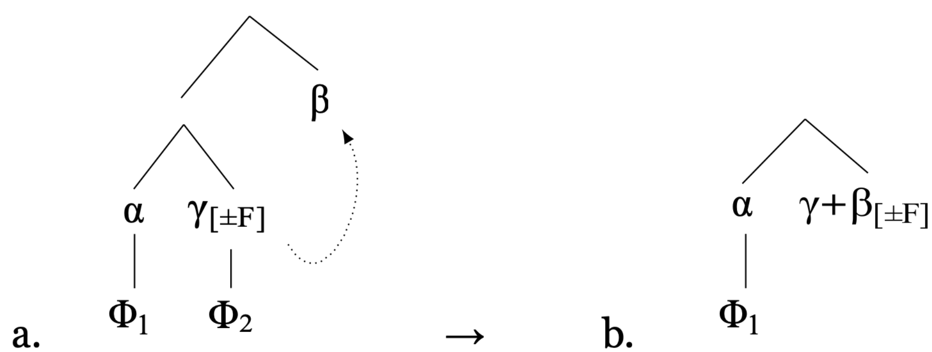
Additionally, (c) it explains why phonologically null exponents — regardless of their marked/unmarked status — appear not to act as interveners for morphological locality.

Null-node pruning

Pruning applies as in (25). First, the terminal node Y^0 is assigned a Ø during VI. As such, Y^0 is detached from its immediately dominating node, but is not deleted.

Its features become floating and are merged with an adjacent higher terminal node if there is one. The lower adjacent node X^0 has already undergone VI — in (25); Φ is an overt exponent that is already inserted; the higher adjacent node Z^0 is instead yet to be realized, thus being the available docking site. As a result, the nodes Y^0 and Z0 fuse into $Y^0 + Z^0$, which undergoes VI again and get assigned the exponent ω.

(25) Zero node pruning & fusion



Φ_1 and Φ_2 are exponents, Φ_2 is phonologically empty.

Vedic Sanskrit Verbal form VIs

Considering the Vedic Sanskrit forms in (17) (repeated here as (26), one observes that only Asp and Mood have overt exponence as in (27); the other nodes are then assigned zero exponence.

(26) a. *kr̥-ṇu-yā́-ta*
[[[kar]$_{root}$ -nau] –yā] -ta]
make-IPFV-OPT-ACT.2PL.PAST

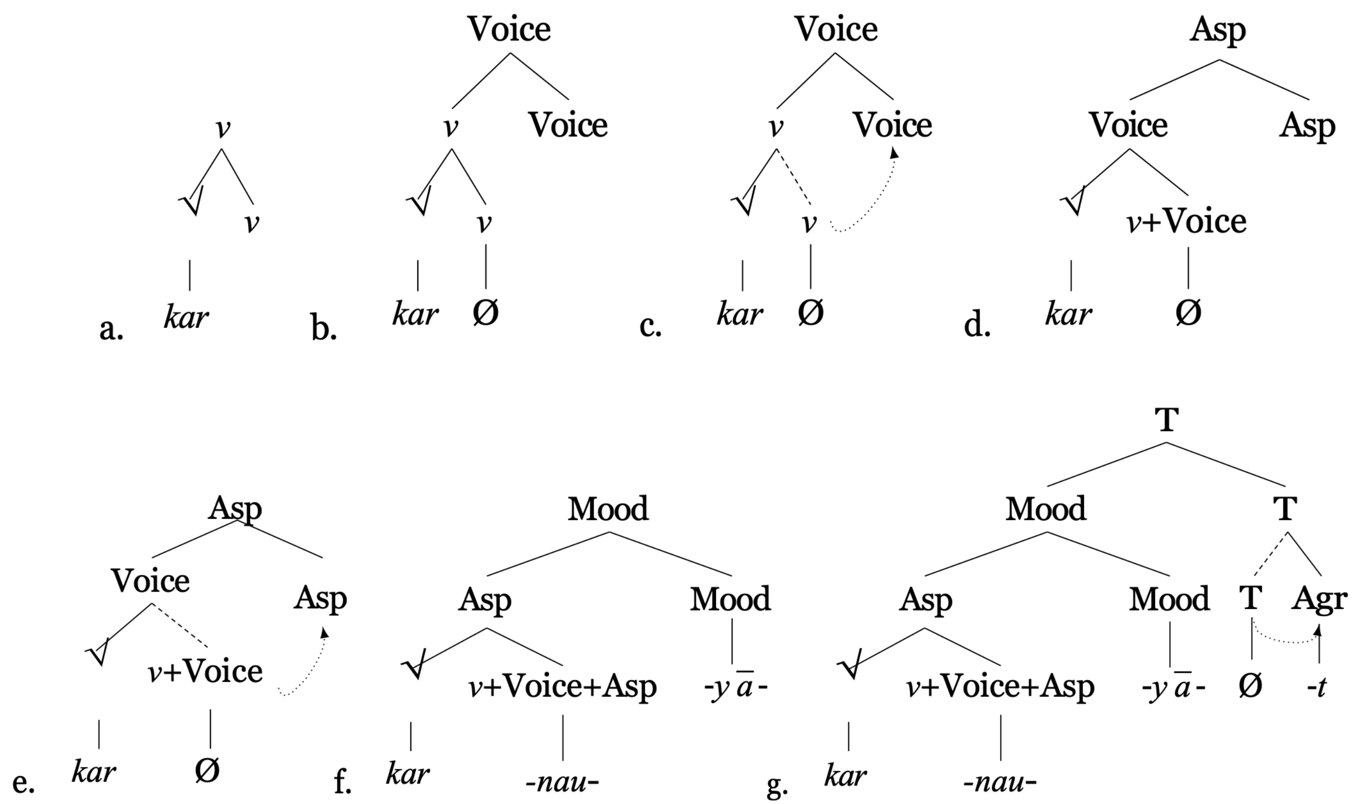
b. *kr̥-ṇáv-a-dhve*
[[[[kar]$_{root}$ -nau] -a] -dhve]
make-IPFV-SUBJ-MID.2PL.PRS

(27)
a. Asp[–pfv] → -nau- / RootV, VIII ___
b. Mood[+irr, –desid] → -a-
c. Mood[+irr, +desid] → -yā-
d. Ø elsewhere

By cyclic pruning, v and Voice are delinked and fused with the higher local Asp node. Further, T is pruned and fused with Agr.

Cyclic Derivation of kr̥-ṇu-yāt

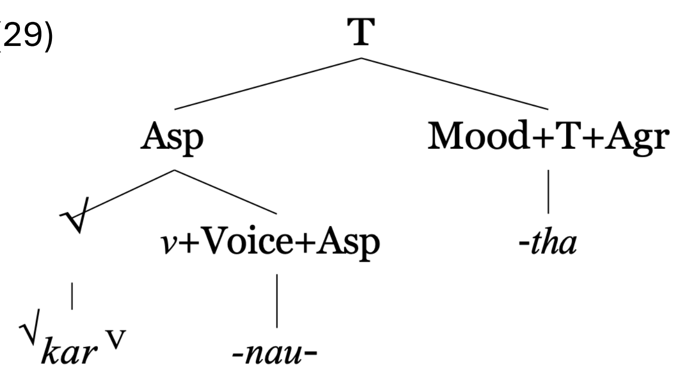
(28) Cyclic derivation of *kr̥-ṇu-yā-t* /kar-nau-yā-t/ make-ipfv-opt-3sg.act.pst 'may he/she make'



Cyclic Derivation of kr̥-ṇu-thá

If Mood also has zero exponence, for example, in indicative forms where only Asp has overt realization, the resulting structure is derived as in (29).

(29)



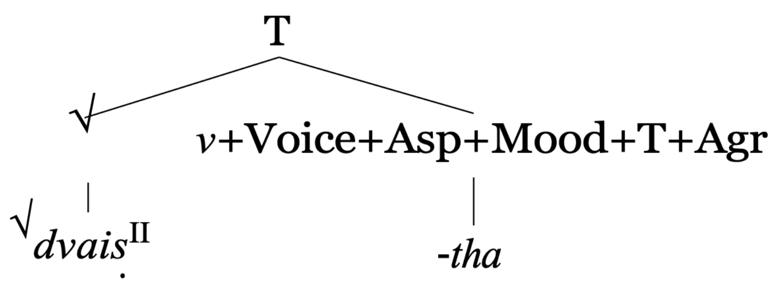
Again, this is the underlying form before the computation of processes such as accent and zero grade. The surface form after application of all phonological rules is given in (30).

(30) [kr̥-ṇu-thá] do-ipfv-act.2pl.prs (After application of accent and Ø-grade)

Cyclic Derivation of dviṣ-thá

There are also cases in which the *v*/Voice/Asp complex is non-overt. Therefore, all functional nodes are fused as in the sample form in (31), a class II (root) present.

(31) *dviṣ-thá* hate-prs.act.2pl 'you (pl.) hate'



Pruning criterion

Thus, what matters is the pattern of overt exponence.

Given an independently motivated morphosyntactic structure such as that in (20), once the distribution of overt exponents has been established, zero exponent insertion and bottom-up cyclic pruning will generate the relevant simplified structures.

In the Vedic verbal system, (v/Voice)Aspect, [past] Tense, and [+irrealis] Mood have overt exponents, as well as Agr (the inflectional endings).

Putting the exponence of [+past] Tense aside for now, this readily accounts for the structures of Vedic Sanskrit verbal forms in a simple and syntactically motivated manner.

the exponence of the root-adjacent nodes

In PIE, as in ancient IE languages such as Vedic Sanskrit and Greek, the present (imperfective) system displayed a wide variety of affixes (cf. LIV2; Fortson 2010; Weiss 2020), which are traditionally divided into different present classes and involve root-dependent realization of Asp[−PFV], see below table which lists a selection of Vedic verb classes and their PIE and Ancient Greek (AG) correspondences.

Imperfective stem-forming morphology in Vedic and Indo-European

Vedic class	Root	Meaning	Ved.	PIE	AG
I	*as*	‘be’	-Ø-	-Ø-	-Ø-
II	*bhav*	‘become’	*-a-*	*-*e/o*-	*-e/o-*
IV	*(s)paś*	‘see’	*-ya-*	*-*i̯e/o* -	*-ye/o-*
V	*kar*	‘make’	*-nav* -	*-*neu̯*-	*-nŭ̆f-*
VII	*yauj*	‘yoke’	*-n(a)-*	*-*n(e)*-	(-*n*-)

The standard analysis is that these stem-forming elements were originally Aktionsart/lexical aspect markers that were reanalyzed as aspectual markers at some late stage of PIE (e.g., Hoffmann 1970; Rix 1986; Strunk 1994; LIV2; Fortson 2010; Meier-Brügger 2010). Diachronically, these elements would then have been reanalyzed as exponents of the formerly Ø-exponed head Asp (cf. Ringe and Eska 2013:167–177 for further examples of terminal node reanalysis).

The morphosyntactic position of aspectual elements

These suffixes cannot occur together: They seem to compete for the same structural position. This also holds for the perfect and aorist system: exponents of Asp[+pfv,-stat] (aorist stem) and Asp[+pfv,+stat] (perfect stem) compete for the same position as (and never occur together with) exponents of Asp[_pfv] (present stem). This is illustrated in the table below again for Vedic.

Vedic present, aorist, and perfect stem-forming morphology (stem formant = underlined)

Present	Aorist	Perfect	Meaning
ján-a-	*ján-iṣ-*	*ja-jan/ja-jñ-*	‘beget, create’
bháv-a-	*bhú-/bhú-Ø-*	*ba-bhū̆-*	‘become’
dá-dhā/dá-dh-	*dhā-/dh-Ø-*	*da-dhǎ-/da-dh(i)-*	‘put’
vrh-á-	*vrk-śa-*	*va-várh-*	‘tear’
bhi-ná-d-/bhi-n-d-	*bhéd- / bhid-Ø-*	*bi-bhéd-/bi-bhid-*	‘split’
kr-náv-/kr-nu-	*kár-/kr-Ø*	*ca-kǎ́r- / ca-kr-*	‘make’

Exponence complexities

Note that there are fewer stem classes in the aorist than in the present stem: a root aorist with Ø-exponence, the s -aorist (with three–four subclasses), the thematic aorist with the theme vowel -a-, and the reduplicated aorist.

The perfect is even more constrained, as there is only one type, the reduplicated perfect (with one exceptional unreduplicated perfect, véda 'knows').

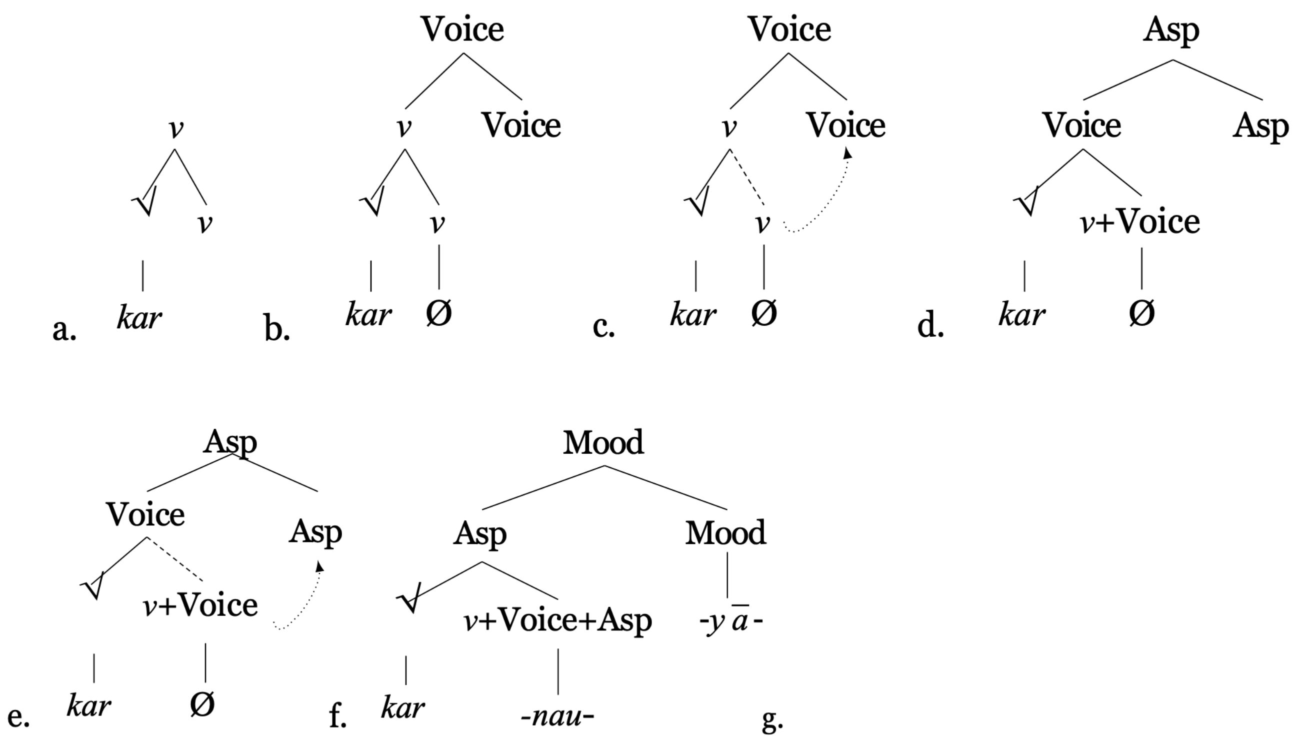
One and the same root is often associated with different types of present and aorist stems (Joachim 1978; LIV2), and different combinations of present + aorist classes are attested.

Moreover, the selection of stem formants is root-dependent, and specific modifications of the root (such as ablaut grade) in turn are triggered by adjacent stem-forming suffixes.

Some are moreover associated with meaning or argument structure properties that are closer to Aktionsart than to aspect, such as transitivization and causative, inchoative, or iterative meaning (see Grestenberger 2022c on AG).

V+ voice+Asp node

Calabrese and Petrosino (2023) propose that these Aktionsart features associated with v are inherited by Asp after zero node pruning and fusion (see (8)).



VIs for PIE aspect

The respective stem-forming affixes are then inserted in a node that results from the bundling of v and Voice with Asp. One can therefore propose that Asp is realized through different root-dependent VIs, as in Table below (we use root diacritics to refer to a group of particular roots.

VIs for PIE aspect

a.	$\text{Asp}_{[+\text{PFV},-\text{STAT}]}$	→	**-s -*	/ $\sqrt{\ }^{\textit{-s-}}$ _	(sigmatic aorist)
b.	$\text{Asp}_{[+\text{PFV},+\text{STAT}]}$	→	$*\text{Ø}^{\text{RED}}$		(perfect)
c.	$\text{Asp}_{[-\text{PFV}}$	→	**-i̯e/o –*	/ $\sqrt{\ }^{\textit{i̯e/o-}}$ _	*(*-i̯e/o*-present)
d.	$\text{Asp}_{[-\text{PFV}]}$	→	**-ne -*	/ $\sqrt{\ }^{\textit{-ne-}}$ _	(nasal infix present)
e.	$\text{Asp}_{[-\text{PFV}]}$	→	**-neu̯-*	/ $\sqrt{\ }^{\textit{neu̯-}}$ _	(*-*nu* -present)
f.	$\text{Asp}_{[-\text{PFV}]}$	→	**-sk̂e/o-*	/ $\sqrt{\ }^{\textit{sk̂e/o}}$ _	*(*-sk̂e/o*-present)
g.	Asp	→	*-Ø-	/ $\sqrt{\ }^{\text{Ø}}$ _	(root present/aorist)
h.	Asp	→	**-e/o-*	/ $\sqrt{\ }^{\textit{-e/o-}}$ _	(thematic present/aorist)

Note that we treat the theme vowel of the simple thematic conjugation (*-e/o-) and the zero exponent of root aorists/presents as underspecified with respect to [±pfv] because they appear in both the perfective and the imperfective stem, albeit with aspectually conditioned ablaut differences. They still need to be specified with respect to root context, however.

Periphrastic Constructions

Periphrastic constructions. Why?

(32) Compound tenses in Italian:

	Synthetic			Periphrastic	
	Present	Imperfect	Simple Perfect	Present Perfec	Pluperfect
	ama	amava	amo	ha amato	aveva amato
Past	-	+	+	-	+
Perfective	-	-	+	+	+
Resultative	-	-	-	+	+

On Italian periphrastic constructions and past participles

If synthetic forms are due to operations leading to complex head formation, one can plausibly assume that, in contrast, periphrastic forms involve failure of application of these operations.

One can therefore assume that the structure for the periphrastic Present Perfect and pluperfect form is due to a morphosyntactic inability to create a complex head including all of the relevant features (see a.o., Arregi (2000), Bjorkman (2011), Embick (2000), Fenger (2018), Kiparsky (2004.) and in particular Pietraszko (2016) for similar proposals).

For example, in Pietraszko (2016), periphrasis is due to the impossibility of word formation (c-selection in her model) to combine marked privative features in terminal morphological nodes in synthetic structures. (see Pietraszko (2016:chapt.IV) for details).

On Italian periphrastic constructions and past participles

In the model adopted here, following what proposed earlier with regard to head raising/head lowering, I formalize this impossibility in terms of filters disallowing combinations of functional head binary features (see also Fenger (2018)).

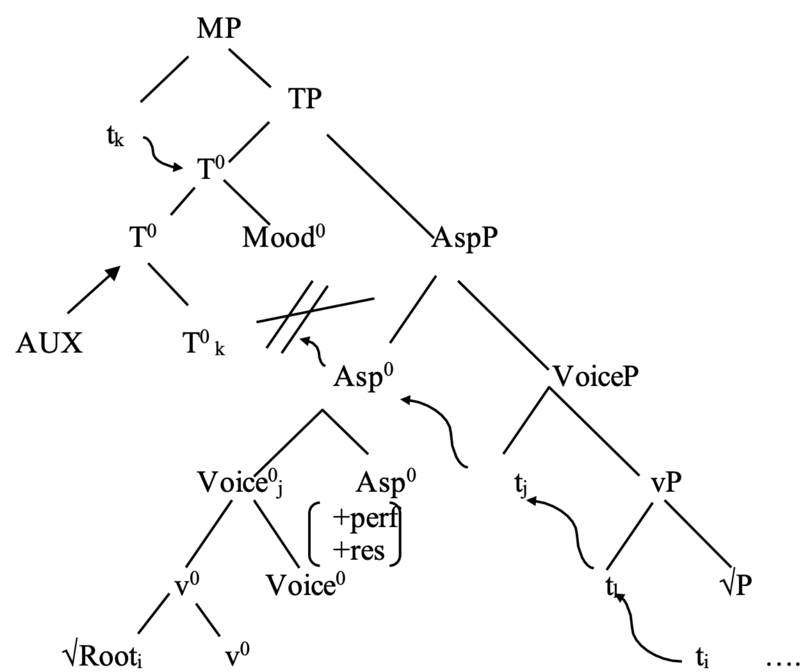
I thus assume that the periphrastic verbal forms are the result of filters that block m-word merger of heads to other heads. In the case of the compound perfect in Italian, the relevant filter is that in (51) that blocks head raising of the complex root +Asp to the higher T.

(33) *[T^0, +resultative]$_{\text{m-word}}$

Blocking of m-word formation

Crucially, differently than the filters in (14), (51) also blocks head lowering of the higher T^0 node to the lower complex with Asp^0. It follows that the higher T^0 is in violation of (3). A dummy root—the AUX root—is therefore inserted as a "holder" for T^0 (Bjorkman (2011). Assuming that here one is dealing with Italian, M^0 is lowered to T^0 and (52) is generated.

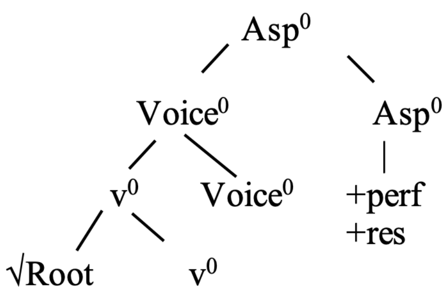
(34)



The participle

Given the analysis just proposed, the participle has the basic structure in (35). It is essentially a tenseless, moodless verbal form.

(35)

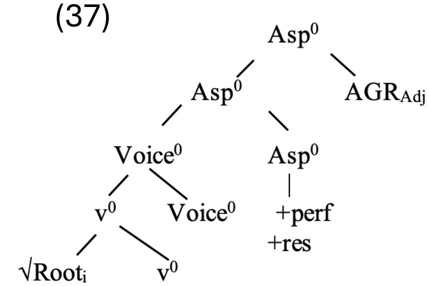


The participle

Characteristically, verbal forms of these type, especially in the Indo-European languages, have adjectival properties, specifically the agreement morphology typical of adjectives. One must account for the appearance of these adjectival morphology. For now, I will assume that this is due to the type of AGR that is inserted in the verbal m-word. AGR is inserted in absence of inherent Phi features, which are found only in nouns: AGR_V probes for person and number features, AGR_{Adj} probes for gender and number features (and case features in languages with overt morphological case). One can then hypothesize that AGR_V is inserted only when there is T in the same m-word, otherwise AGR_{Adj} is inserted, i.e., AGR_{Adj} is the default AGR (I assume that only nouns can carry inherent Phi-features):

(36) Given a MP unit U not including inherent Phi-features,
a. Adjoin AGR_V to its highest X^0 if U contains T^0
Otherwise:
b. Adjoin AGR_{Adj} to its highest X^0.

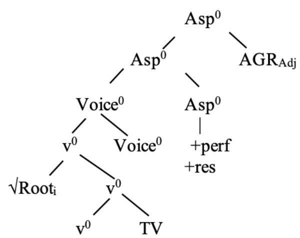
(37)



The participle

Insertion of the v^0 TV at this point generate (38) :

(38)

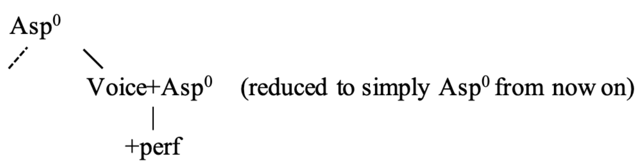


Pruning of null v^0 will apply to this structure. Furthermore, if we assume that in general the voice head, is always assigned a zero exponent in Italian (39) , null exponent pruning will always fuse the sequence $Voice^0$ and Asp^0 as in (40).

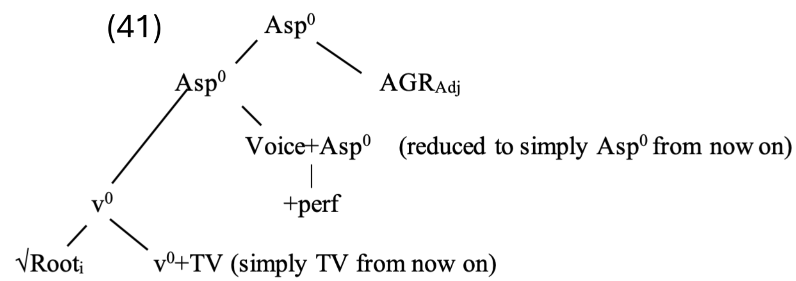
(39) $[\ \ ]_{Voice} \rightarrow \emptyset$

This will generate :

(40)



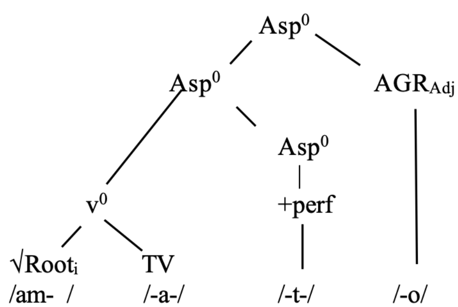
(41)



The participle

This structure accounts for the surface shape of past participle (see (42) where full exponency is assigned to all pieces:

(42)



The filters blocking word-formation

An obvious central issue of the model outlined here is that of the filters blocking word-formation, and subsequently also of the features that are assumed for the functional heads. In not relying on privative features, the theory outlined here allows a simpler account of parametric variation in verbal periphrastic constructions in so far as, although universal, the filters may be deactivated on a language specific basis.

So, a verbal form that is synthetic in a language, e.g. the present perfect in Latin (amavi) can be periphrastic in another language (Italian: ho amato).

A tentative list of the relevant filters for Romance varieties is given in (62) given the feature assignment for the functional categories in (42):

(42)	*[Mood^0, T^0]$_{\text{m-word}}$	Periphrastic modality
	*[T^0, +resultative]$_{\text{m-word}}$	Periphrastic perfect forms (present perf., pluperf., etc.)
	*[AspPerf^0, +progressive]$_{\text{m-word}}$	Periphrastic progressive
	*[ASP(prog^0), +passive]$_{\text{m-word}}$	Periphrastic passive
	(*[T^0, FUT^0]$_{\text{m-word}}$	Periphrastic future

The filters and the features proposed in (50) are obviously tentative. A detailed investigation of this point is not really important in the theoretical architecture of this book which focus on the entire derivation of surface verbal forms and will be left to future research.

Italian Passive participles

The analysis given above accounts for the past participle that appears in the verbal paradigms of Italian present perfect and pluperfect forms of transitive and intransitive (inaccusative and agentive) verbs as in (9):

(43) Mario ha mangiato una mela ‘Mario has eaten an apple’
Mario ha lavorato ‘Mario has worked’
Mario e’ partito ‘Mario has left’

A past participle form, however, also appears also in the case of passive constructions:

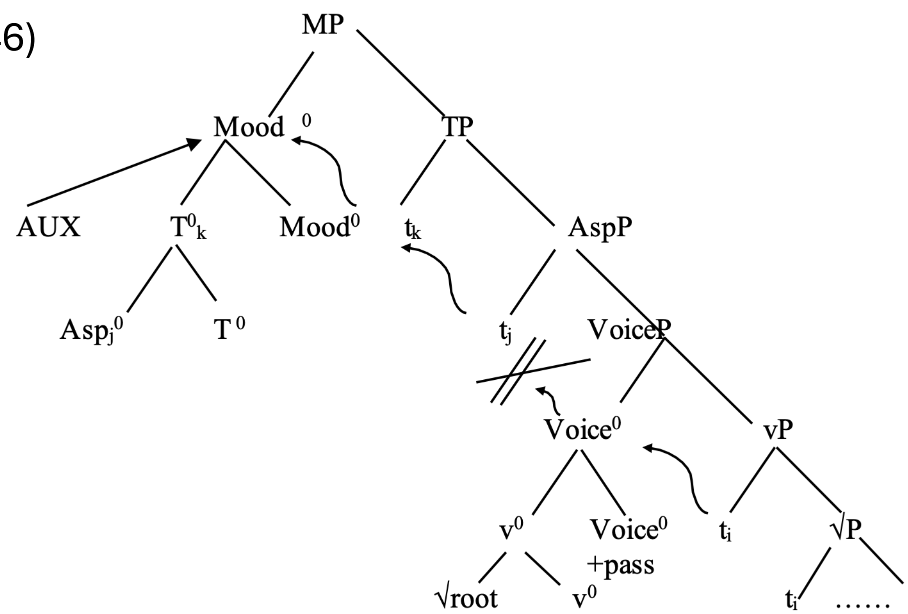
(44) La mela fu/venne mangiata da Mario ‘The apple was eaten by Mario’

Italian Passive participles

We can now assume that passive participles are also generated in the same way due to the filter in (91) that blocks head raising of the root merged with V(oice) to ASP:

(45) *[Asp, +passive]$_{m\text{-}word}$

(46)



Italian Passive participles

This immediately accounts for the fact that the auxiliary found in passive constructions is characterized by more inflectional distinctions and specifically by aspectual distinctions with respect to the auxiliary found in active perfect participle constructions (putting aside future, subjunctive and conditional distinctions, which are morphologically realized in the auxiliaries of both constructions):

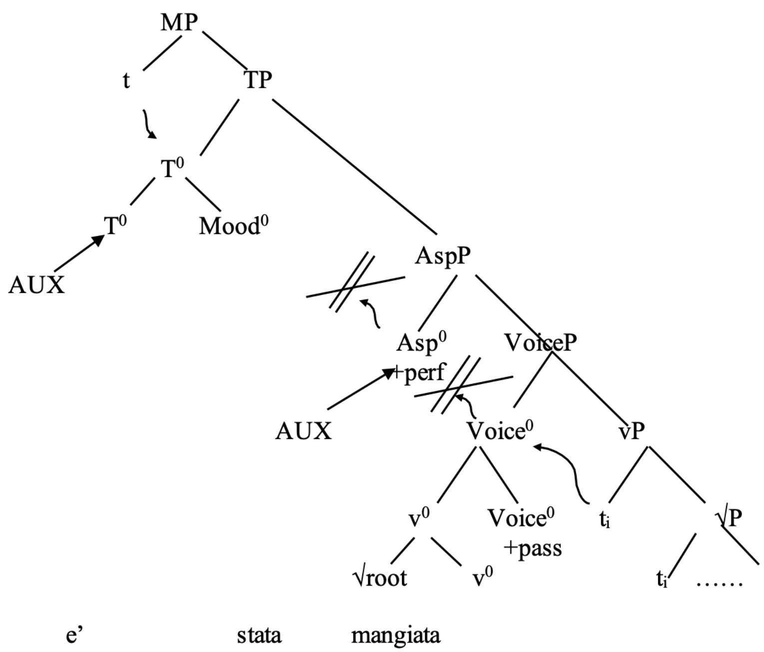
(47) a.	Mario	ha aveva	mangiato	la mela	'Mario	has had	eaten the apple
b.	La mela	e' era fu e' stata era stata	mangiata		'The apple	is was being was has been had been	eaten'

The auxiliary in active perfect participle constructions displays only tense distinctions (see (93a) as expected given the structure in . The auxiliary in the passive construction in contrast displays the full range of the distinction generated by the combination of tense and aspect features (see (93b). This expected given (92) .

Italian Passive participles

The analysis of the perfect and passive participle forms just proposed also predicts a perfect passive structure with two participles: one for the ASP that does not merge with T and the other for the low voice. This prediction is correct (cf. *la torta e' stata mangiata* 'the cake was eaten') (see also Ippolito (1999), as illustrated by the tree in (94) (before the insertion of TV and AGR elements and pruning operations):

(48)



Perfect and Passive participle forms

Given the structures in (34) and (46), there should be two different morphological types of participle: the passive one and the perfect one. As a matter of facts, however, these two participles are always morphologically realized in the same way in Italian.

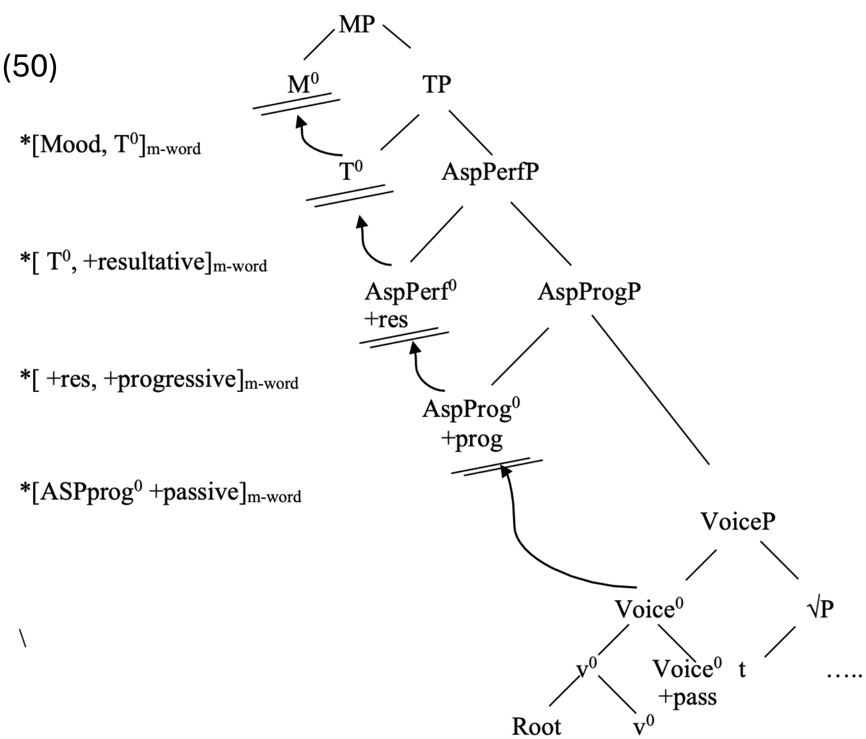
(49)	a.	Carlo ha <u>mangiato</u> il gelato	'Carlo ate the ice cream'
	b.	Il gelato e' mangiato da Carlo	'The ice cream is eaten by Carlo'

This convergence will be discussed in another class

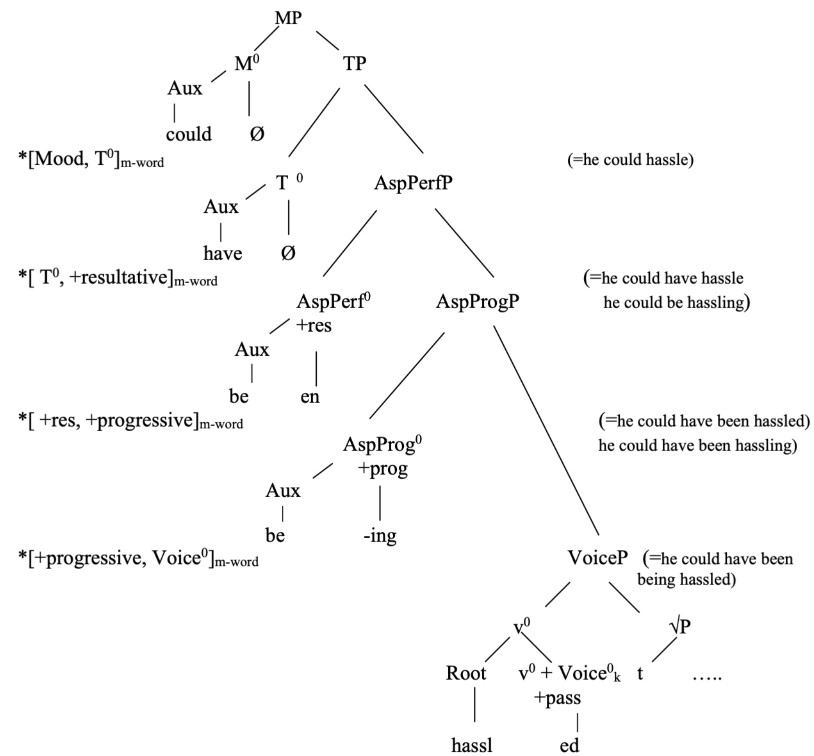
English periphrastic verbal morphology

The relevant head lowering operations are blocked as in (114). VIs and auxiliary insertion will then generate (115) and the rigid ordering of English auxiliaries is readily accounted for.

(50)



(51)



IRREGULAR MORPHOLOGY

Irregular morphology=morpheme-specific morphology

Morphological operations dependent on morpheme specific information.

Regular morphology

Italian Imperfect marker
always the same across verbs

amavo/battevo/partivo

[+imperfect] → /-v-/

Irregular morphology
Italian perfect marker /s/ or /X/ appearing only with certain verbal roots

persi , venni

[+perfect]→ /-s-/ / $root^s$ ___ ($root^s$ =perd, etc.)
[+perfect] → /-X-/ / $root^s$ ___ ($root^s$ =ven, etc.)

Most typical case: morphological operations dependent on root specific information

Morpheme Exponence dependent on root specific information=root based contextual allomorphy

Root-based contextual allomorphy accounted for by: Vocabulary Items and MP Rules (morpho-phonological rules) including root-information in their structural description.

Root based contextual allomorphy in Italian perfect and past participle forms.

Basic observation :
A correlation between presence vs. absence of regular morphology and presence vs. absence of thematic vowels.

Root based contextual allomorphy occurs only when the thematic vowel is absent.

(52)

Irregular:	*vs.*	**Regular:**
$[[[[\textit{pérd}]_{\text{root}}\ \_\_\ ]_{\text{V}}\ \text{-}\textit{s}\text{-}]_{\text{T}}]\ \textit{i}]_{\text{AGR}}$		$[[[\textit{part}]_{\text{root}}\ \textbf{-i}_{\textbf{TV}}\text{-}\ ]_{\text{V}}\ \text{-Ø-}]_{\text{T}}\text{-}\textit{sti}]_{\text{AGR}}$
pérsi ‘lose-Perf-1sg’		partisti ‘leave-Perf-1sg’
$[[[[\textit{pérd}]_{\text{root}}\ \_\_\ ]_{\text{V}}\ \text{-}\textit{s}\text{-}]_{\text{T}}]\ \textit{o}]_{\text{AGR}}$		$[[[\textit{part}]_{\text{root}}\ \textbf{-i}_{\textbf{TV}}\text{-}\ ]_{\text{V}}\ \text{-t-}]_{\text{T}}\text{-}\textit{o}]_{\text{AGR}}$
pérso ‘lose-PerfPart-MscSg’		partito ‘leave-PerfPart-MscSg’
Athematic		**Thematic**

Root-based contextual allomorphy and athematic morphology

Root-based contextual allomorphy is observed only in athematic morphology.

Why:
Morpheme-to-morpheme interactions are local (cf. Embick (2010) reformulated by Calabrese (2015) as in (53)

(53) a. An overt node a morphologically interacts with an overt node b iff a, b are local.
b. a, b are local if no node intervenes.

Locality in Kiowa distributives (Adger et al. 2003)

(54) a. á-kˈúú
3.PL.AN -sit
'They sit.'
b. hɔ́n á- kóp –gɔ̂ɔ
NEG 3.PL.AN -sit -NEG
'They don't sit.'

In (3) we can see that the lexical verb root kˈúú 'sit' undergoes suppletion conditioned by NEG: kˈúú is realised as kóp when NEG is present in the structure. However, when the distributive marker yɔ́ is added to the structure, which, as shown in (4) is located between the verb and NEG, root-suppletion is blocked:

(55)

		[$_v$ ROOT]	DISTR	NEG	
a.	á-	kˈúú			'They sit.'
b.	á-	kóp		gɔ̂ɔ	'They don't sit.'
c.	á-	kˈúú	yɔ́		'They sit about.'
d.	á-	kˈúú	yɔ́	gɔ̂ɔ	'They don't sit about.'
e.	*á-	kóp	yɔ́	gɔ̂ɔ	'They don't sit about.'

Specifically, in (4d) NEG-driven root-suppletion is blocked and instead the elsewhere form of the root surfaces, ku´u´. Adger et al. (2003) argue that this is because the distributive morpheme intervenes between the verbal root and NEG.

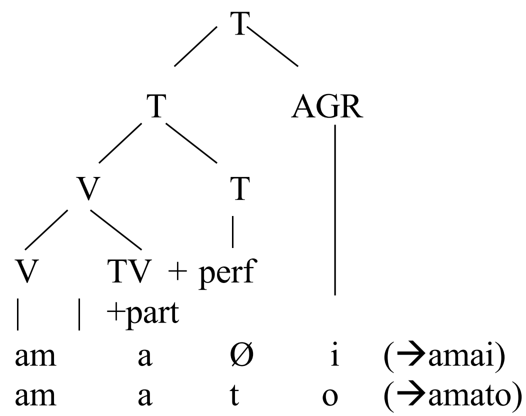
(56) [[[v √SIT] DISTR] NEG]

They argue that root allomorphy is necessarily conditioned by adjacency, and thus the root cannot supplete for NEG due to the intervention of distributive y´O.

Analysis of Perfect/past participle forms

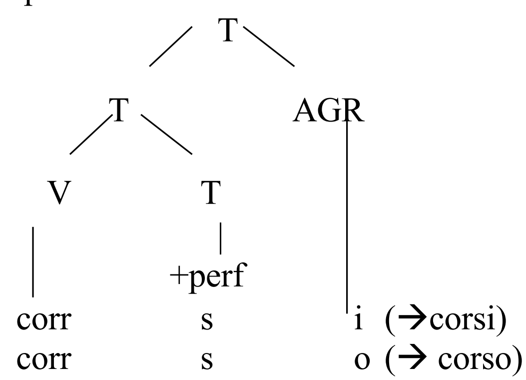
Thematic perfect/ past participle

(57)

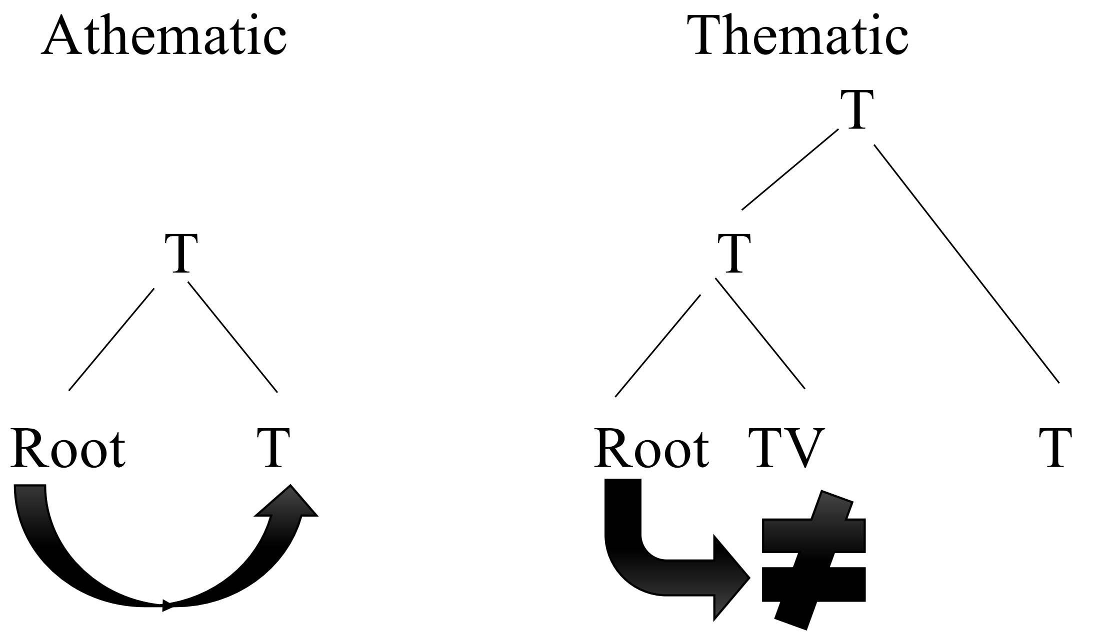


Athematic perfect/past participle:

(58)



Locality



LOCALITY

The rules accounting for the Passato Remoto allomorphy appear to be governed by linear adjacency. Thus, they apply only if there is this type of adjacency between the location of their application and the irregular root. If this adjacency is disrupted by the presence of a Thematic Vowel, only regular unmarked morphology will occur.

For example, the special exponent /-s-/ for the [+past] Tense of some of the irregular Passato Remoto forms is inserted by the Vocabulary Item in ( ) where the symbol ^ indicates linear adjacency):

(59) Vocabulary Item for the Passato Remoto Tense morpheme (provisional)
[+Past]T $\rightarrow$ /s/ / Roots ^____ (where Roots = corr, val, perd, met, etc.)

Given the adjacency requirements, (7) can apply only when the Thematic Vowels are missing. So it can apply in the structure in (8) but not in (9) where the TV is present:

(60) a. [[[corr^{S}]$_{root}$] +PAST]$_{T}$ +part,-+auth, -pl] $_{AGR}$
b. [[[corr^{S}]$_{root}$] s]$_{T}$ i]$_{AGR}$
cor-s- i

(61) a. [[[corr^{S}]$_{root}$ TV] +PAST]$_{T}$ +part,-+auth, -pl] $_{AGR}$
b. [[[corr^{S}]$_{root}$ e] s]$_{T}$ i]$_{AGR}$
*corr-e-s-i

Against affixless theories

Affixless theories (Anderson 1992, Aronoff 1976, 1994, Maiden 2004, Matthews 1972, Stump 2001 a.o)

(62) Affix–based theories
[[[root ± W] ±X] ±Y]

(63) Affixless theories

$$\text{Root} \begin{pmatrix} \pm W \\ \pm X \\ \pm Y \end{pmatrix}$$

Embick (2013): Affixless theories make no predictions about the locality of morpho-phonological interactions since all of the features are equally local in representations like (11)