How Outsiders discovered CTLP‘s Point of View

(CT = Classical Tamil)
(LP = Language Professionals)

Jean-Luc Chevillard
Tamilex (Hamburg) and HTL (Paris)
Budapest (21 and 28 september 2026)

FIRST PART

(21st september 2026)

„History of Linguistics“, as a 50 year old field

This presentation will be partly based on my 30 years of experience as a member of a French CNRS team called HTL, i.e. Laboratoire d‘Histoire des Idées Linguistiques. During that period, I was for 12 years as the editor of a scientific journal called HEL i.e. „Histoire Épistémologie Langage“, published by the SHESL , i.e. „Société d‘Histoire et d‘Épistémologie des Sciences du Language“. The founder for both HTL and HEL is Sylvain Auroux. These three entities (SHESL, HEL and HTL) can also be seen as part of an international network, which has existed for almost 50 years, which periodically organizes the ICHoLS conferenc, i.e. „International Conference for the History of the Language Sciences“ (https://ichols.org)

Describing the languages of the world, after describing those of Europe:

„Extended Grammars“

(workshop at ICHoLS 2017)

https://gokulam.net/my_articles/2020/Beiträge_30-1_Foreword.pdf

Cristina Muru
Grammaire Latine Étendue. Two Portuguese missionary *Tamil Arte* (17th cent.) .. 59

Gonçalo Fernandes
Latin sources of the first printed Konkani and Marathi grammars (1640–1859) ... 74

Thi Kieu Ly Pham
Les deux premières grammaires vietnamiennes écrites en latin (1651 et 1838) 88

Claudia Schweitzer
Les traces des grammaires latine et grecque dans le traitement de l'accent typographique chez les grammairiens et musiciens français vers 1700 103

Karla Janiré Avilés González
De la grammaire castillane étendue aux langues nahuas du Mexique 119

José Edicarlos de Aquino
Un cas de « grammaire latine étendue ».
Holmes Brazileiro ou Grammatica da Puericia, de Júlio Ribeiro (1886) 132

Specimen

Beiträge zur Geschichte der Sprachwissenschaft

Extended Grammars

Themenheft



30.1 (2020)

Describing the languages of the world, after describing those of Europe: „Extended Grammars“ (workshop at ICHoLS 2017)

tions illustrating what Sylvain Auroux has discussed inside the section « La grammaire latine étendue » on pp. 82–85 of his 1994 book *La Révolution Technologique de la Grammatisation*. Briefly explained, “Extended Grammars” are grammatical descriptions of various languages which make use of tools initially developed in and/or for another language.

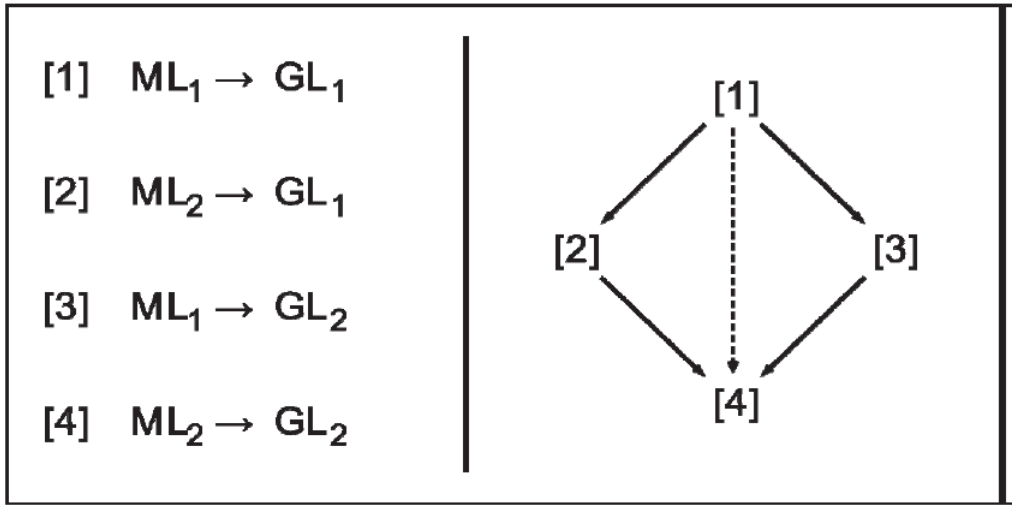


Figure 1:
Going from one grammar to another one, step by step (after Auroux 1994: 82–85)

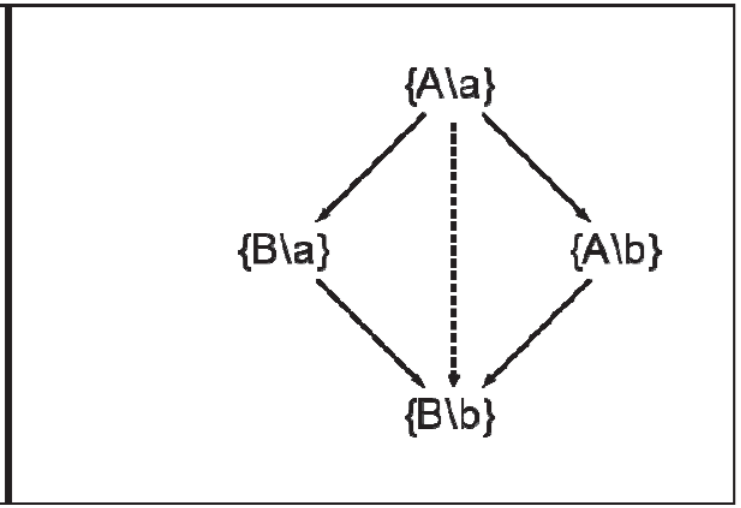
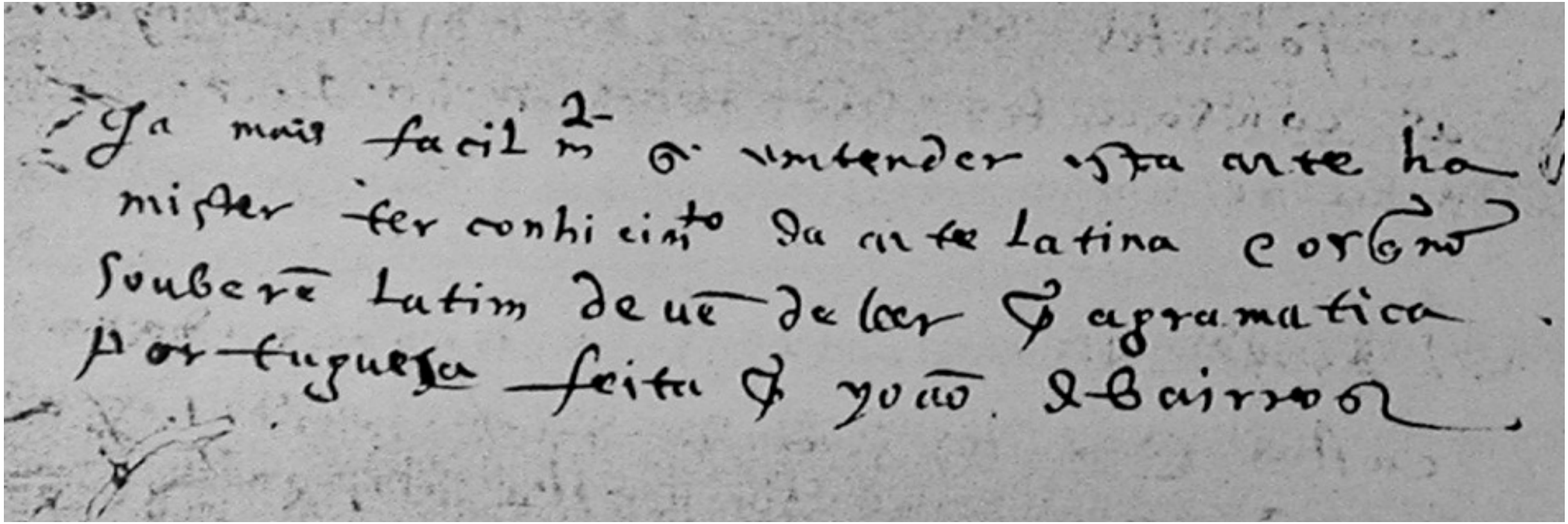


Figure 2:
compact notation for the Auroux diagram (lowercase letters stand for object languages and capital letters stand for the corresponding metalanguages)

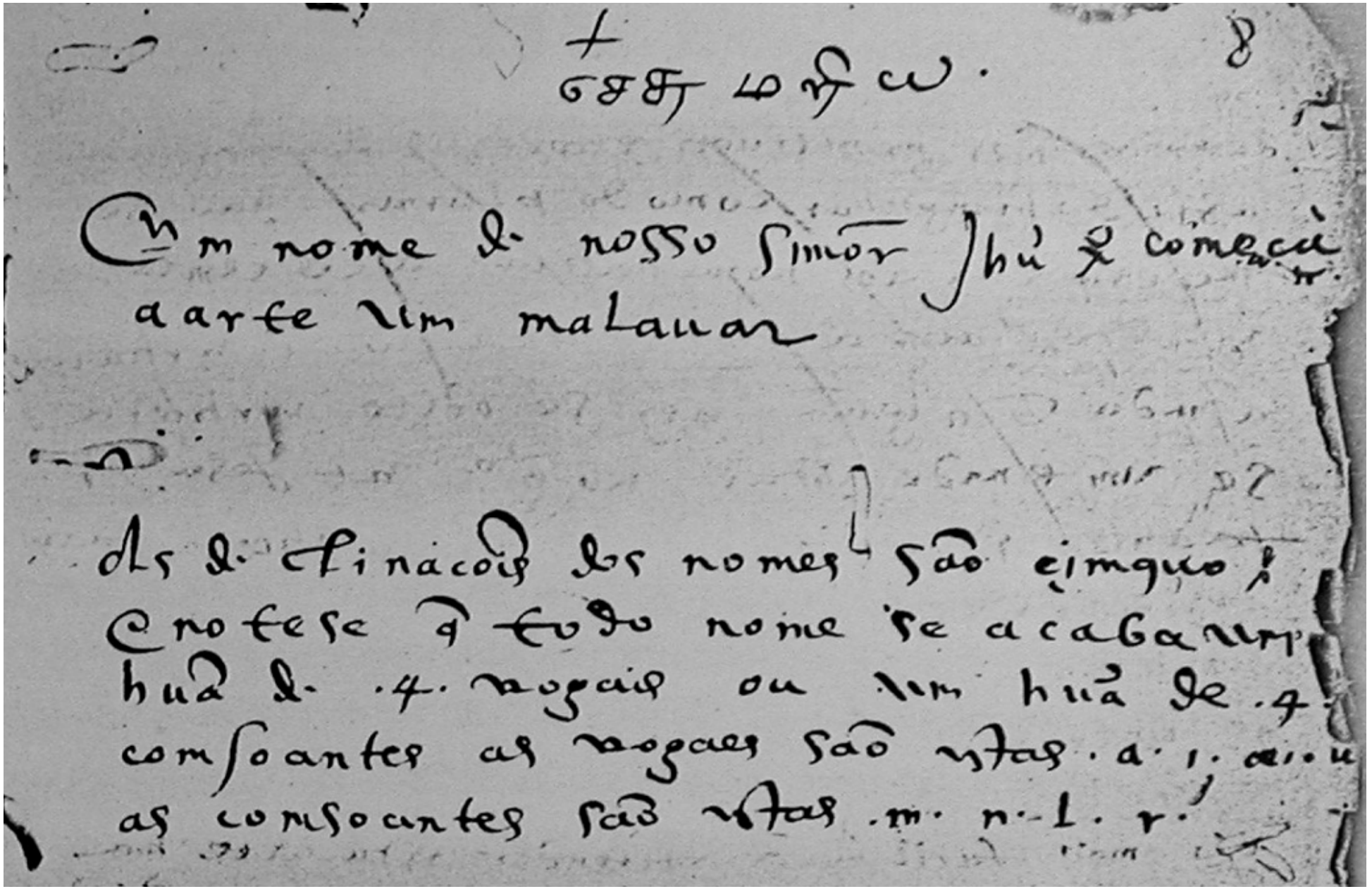
Describing the languages of the world, after describing those of Europe:



Describing the languages of the world, after describing those of Europe:

Para mais facilmente se emtender esta arte ha mister ter conhicimento da arte latina, e os que nõ souberẽ latim deuẽ de leer por a gramatica portuguesa feita por Yoaõ dBairros.

Describing the languages of the world, after describing those of Europe:



Describing the languages of the world, after describing those of Europe:

[fol. 8 r]

Em nome de nosso sinior Jhesu Christo começa a arte em malauar.

As declinaçoīs dos nomes saõ cimquo e notese que todo nome se acaba em huã de 4 vogais ou em huã de 4 comsoantes. As vogaes saõ estas: a, i, æi; as comsoantes saõ estas: m, n, l, r.

Describing the languages of the world, after describing those of Europe:

A primeira declinaçã he dos nomes das castas e estes saõ de duas maneiras, scilicet para homẽs e outra para molheres. Os que seruẽ para os homẽs se acabaõ em æn, e no plurar ẽ ær, aos quais se reduzẽ algũs outros nomes, scilicet nomes de oficios e algũs outros semelhantes.

Os que seruẽ para as molheres se acabaõ em i, e no plurar ẽ iæl ou ẽ gal. Os da 2ª declinaçaõ sam de duas maneiras, scilicet acabados em vogal e outros acabados em cõsoamte. As vogaes

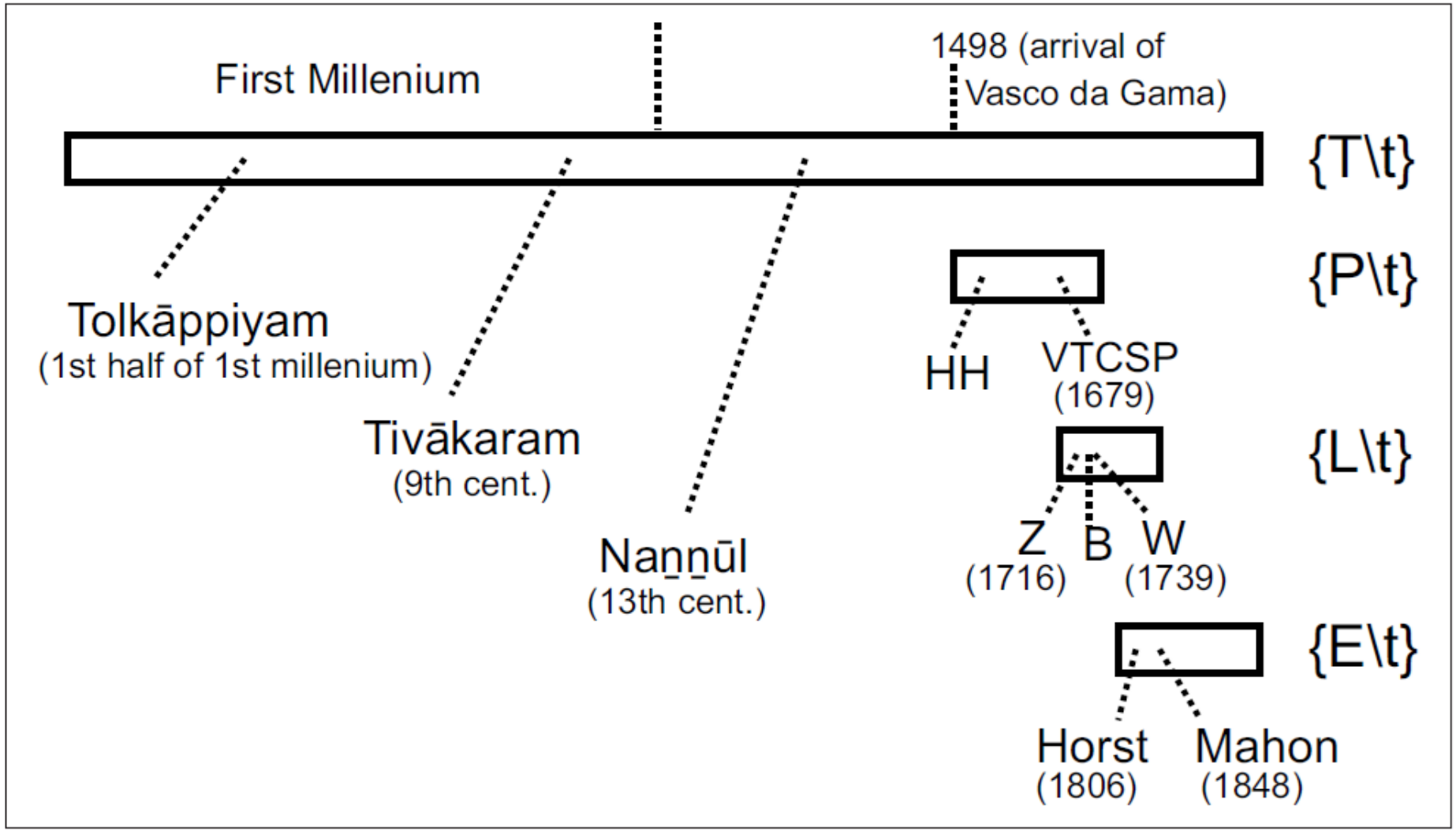


Figure 5: Brief chronology of the early stages in the description of Tamil HH = Henrique Henriques; Z = Ziegenbalg; B = Beschi; W = Walther; P = Portuguese; L = Latin; E = English; t = Tamil as object-language; {T\t} stands for a description of Tamil (=t) written in Tamil (=T).

Beschi as a discoverer of Poetical Tamil (Trento-Chevillard 2025)

Jesu-Christi missionariis // offero; quâ ipsis utilitate futurum sit, facile noverint // omnes, si paulisper animadverterint in his regionibus, // monumenta Deorum, Fabularum figmenta, scientiarũ // præcepta, Poëtarum carmina, astronomiæ calculos // medicinæ leges, musices, choreæq.e regulas, omnia deniq. // vel ipsa prima Grammaticæ rudimenta, elegantiori // hoc idiomate ab antiquis scripta fuisse. Quare nil // omnino de eorum Diis ac fabulis, de eorum artibus // ac scientiis radicitùs, ac sine animi hæsitatione quis // scire possit, si hoc idioma penitùs ignoret. Tunc enim // fidendum erit Indigenis, qui, et Bragmanes prae-// =cipuè, ne quid ignorare se fateantur, nil hæsitando // turgentia verba trutinantes, quæ primùm menti oc- // =currunt, figmenta ex cathedrâ proferre non dubitant; // quæ passim eorum auctores legenti, quàm falsa sint, // manifestè patet. Præterea

Beschi as a discoverer of Poetical Tamil (Trento-Chevillard 2025)

can easily know, if they have observed for a moment how in these regions the chronicles of the Gods, the fictions of their fables, the precepts of science, the poems of the poets, the calculations of astronomy, the rules of medicine, the rules of music and dance, in short everything, and even the first rudiments of grammar, were written by the ancients in this more elegant register. Wherefore no one may know fully and without hesitation of mind anything at all of their Gods and fables, of their arts and sciences, if they are deeply ignorant of this idiom. Indeed, in that case one is forced to trust the natives, who, especially the Brahmins, would not admit to ignore something, and without hesitation balancing bombastic words, which first come to their mind, they have no scruples to utter such imaginations *ex*

Beschi as a discoverer of Poetical Tamil (Trento-Chevillard 2025)

their mind, they have no scruples to utter such imaginations *ex cathedra*; but for whoever reads passages of their authors far and wide, it is plainly evident how false these [invented answers] are. Besides, if I were to show how much the things they say about the gods and their fables are in contrast to reason; they would not hesitate to say that all those things, perceived through the power of reason, are all false and imaginations of the common people; but indeed, if I were to object to those same things by citing a passage in the books of their ancient authors, not for a moment would they dare deny [my argument], and silently and with shame they would be compelled to condemn all those things together with me. And indeed, since all Indians agree more through authority, than by reason; a procedure

Beschi as a discoverer of Poetical Tamil (Trento-Chevillard 2025)

all those things together with me. And indeed, since all Indians agree more through authority, than by reason; a procedure beyond doubt, through which anyone would be convinced, will be rather to cite the scriptures of the ancients. And indeed, the ancient Tamils have talked about God and virtue in a sensible way. But how can you [convince anyone], if you are completely ignorant of the elegant register of the Tamil language, in which truly all the things are written?

This will be enough to realize how much this work will be useful to the missionaries of Jesus Christ, and to all the others who either want to call back the Indians to the faith of Christ away from their ancient error, or at least wish their ancient inventions to be investigated in depth. God, most good and most great, who

The Tamil triglossia

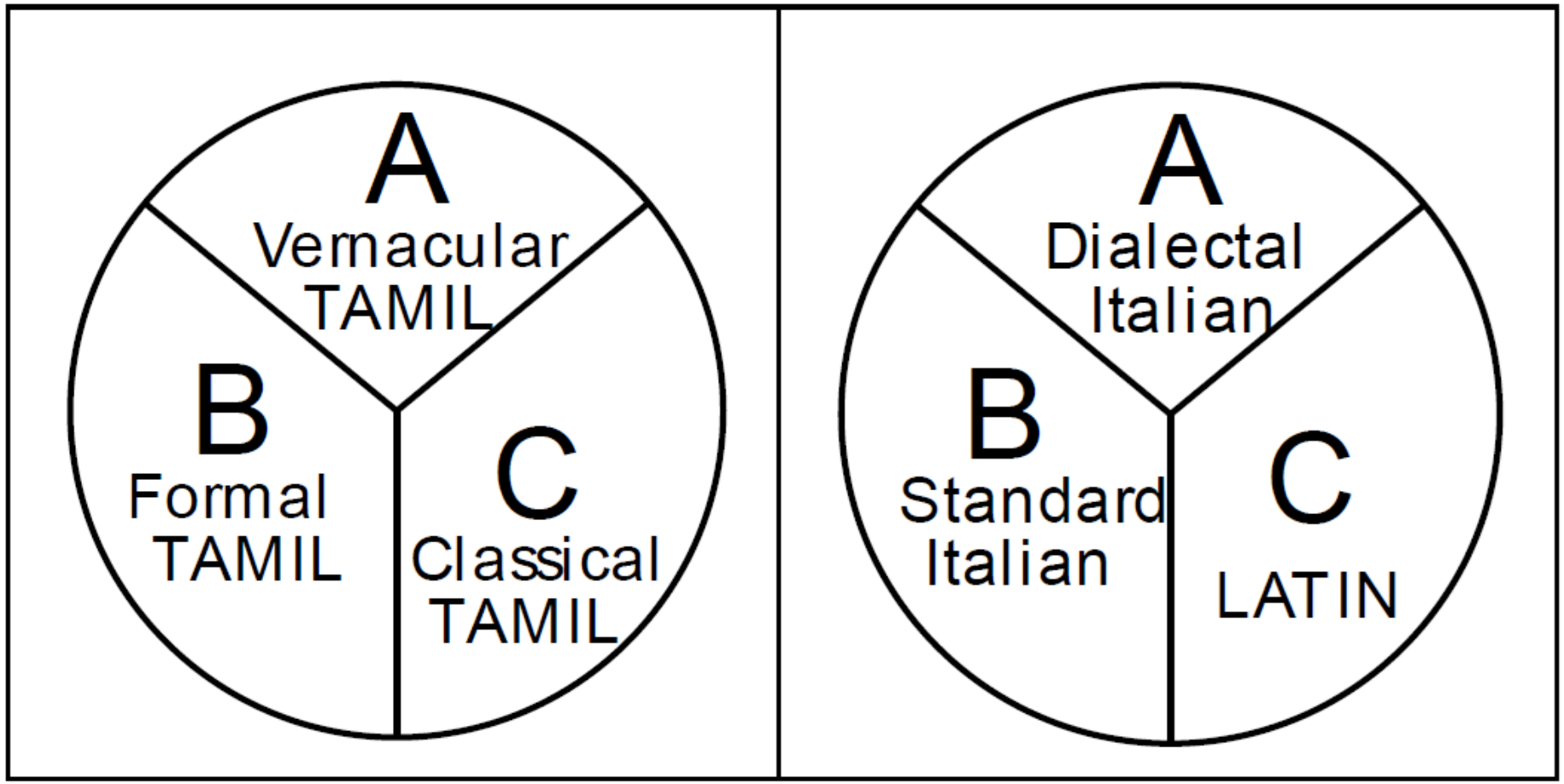
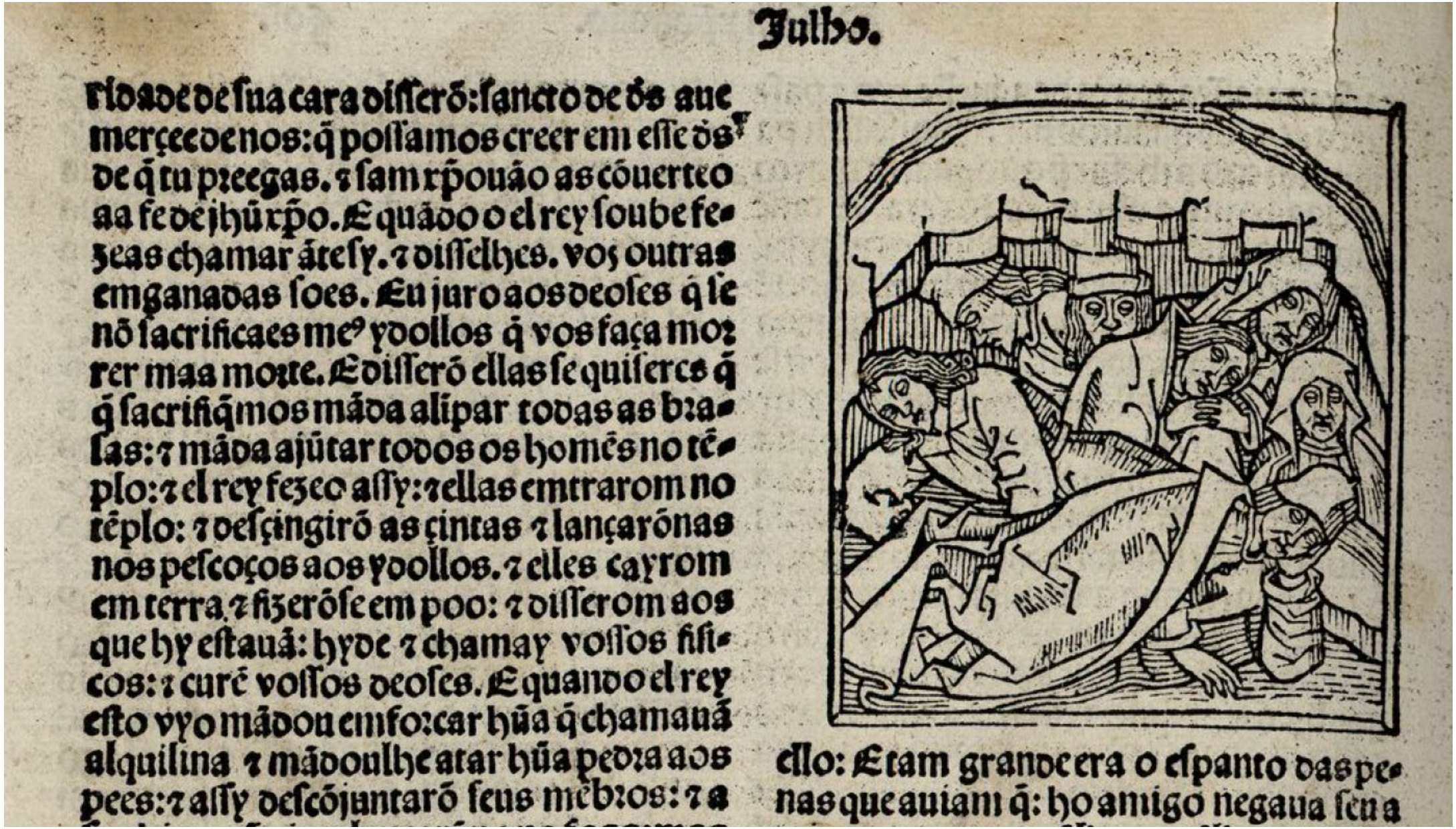


Table 2: Comparing the linguistic situation of Tamil Nadu and the linguistic situation of Italy as two examples of "TRIGLOSSIA"

Flos Sanctorum 1503, folio cviii-v: A vida dos fete dormientes

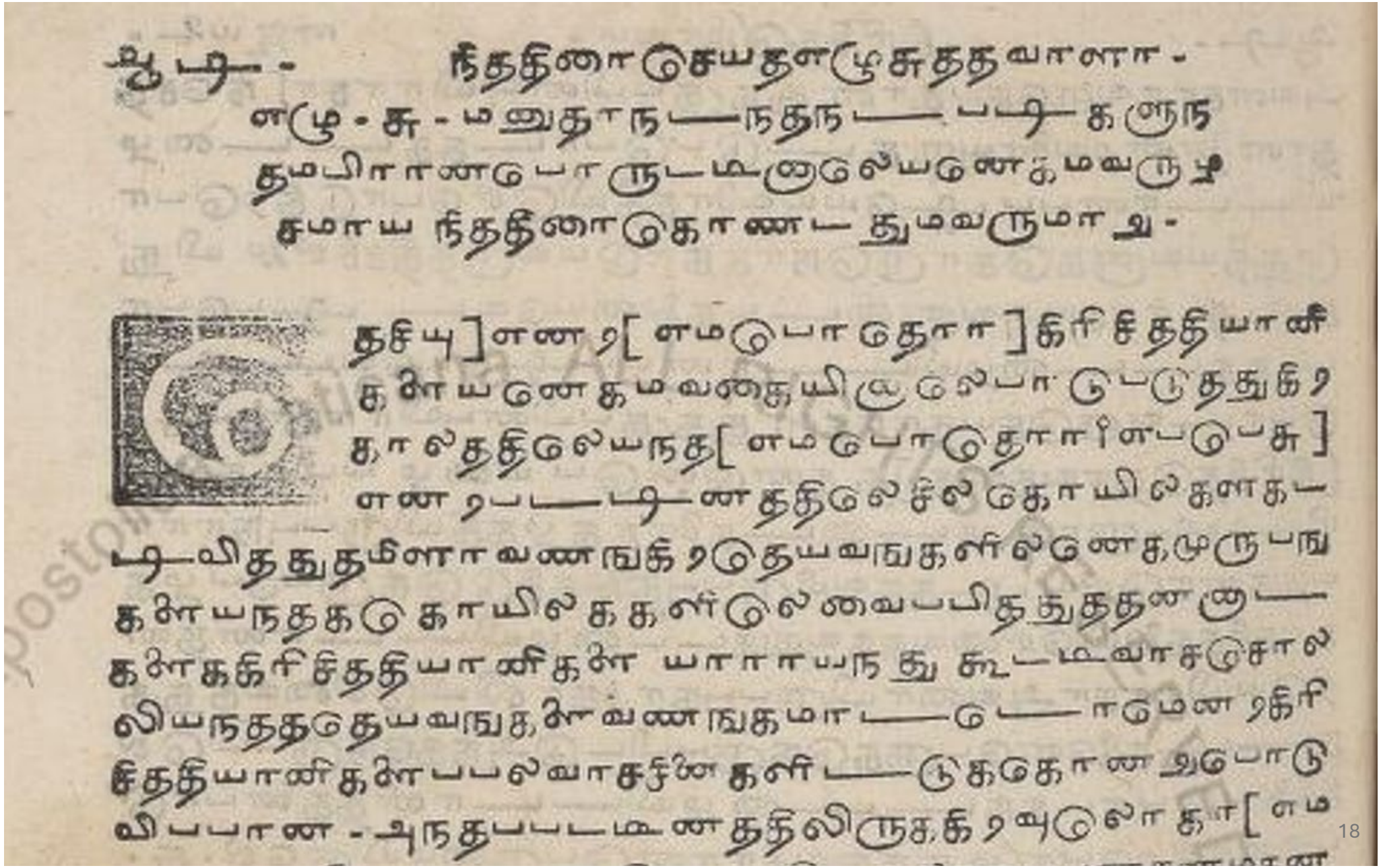


Julho.

rioade de ſua cara diſſerõ: ſancto de ds auemerçee de nos: q̃ poſſamos creer em eſſe ds de q̃ tu preegas. ⁊ ſam rpouão as cõuerteo aa fe de jhũ xpo. E quãdo o el rey ſoube fezeas chamar áteſy. ⁊ diſſelhes. vos outras emganadas ſoes. Eu juro aos deoſes q̃ ſe nõ ſacrificaes me⁹ ydollos q̃ vos faça morrer maa morte. E diſſerõ ellas ſe quiſercs q̃ q̃ ſacrifiq̃mos mãda alipar todas as braſas: ⁊ mãda ajũtar todos os homẽs no tẽplo: ⁊ el rey fezeo aſſy: ⁊ellas emtrarom no tẽplo: ⁊ deſçingirõ as çintas ⁊ lançarõnas nos peſcoços aos ydollos. ⁊ elles cayrom em terra ⁊ fizerõſe em poo: ⁊ diſſerom aos que hy eſtauã: hyde ⁊ chamay voſſos fiſicos: ⁊ curẽ voſſos deoſes. E quando el rey eſto vyo mãdou emfo:car hũa q̃ chamauã alquilina ⁊ mãdoulhe atar hũa pedra aos pees: ⁊ aſſy deſcõjuntarõ ſeus mẽbros: ⁊ a

ello: E tam grande era o eſpanto das penas que auiam q̃: ho amigo negaua ſeu a

Flos Sanctorum 1586 (Vat.ind.24_0282_fa_0139v_m): நித்திரை செய்த ஏழு சுத்தவாளர்



SECOND PART:
discovering the technical literature

(28 th september 2026)

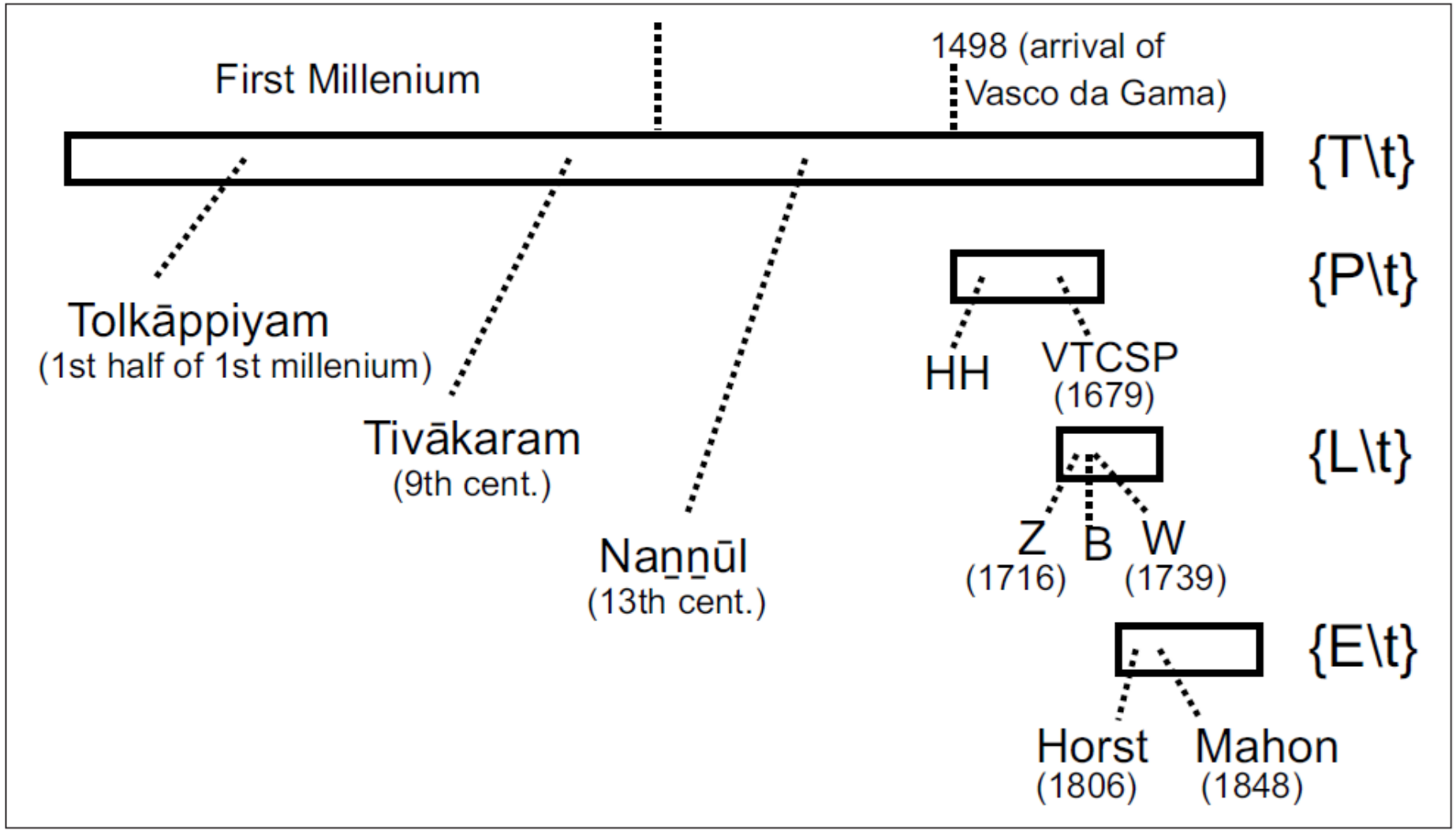


Figure 5: Brief chronology of the early stages in the description of Tamil HH = Henrique Henriques; Z = Ziegenbalg; B = Beschi; W = Walther; P = Portuguese; L = Latin; E = English; t = Tamil as object-language; {T\t} stands for a description of Tamil (=t) written in Tamil (=T).

How the 1679 VTCSP mostly ignores Tamil grammatical terminology and displays sanskrit terms (after Chevillard [2015: 122])

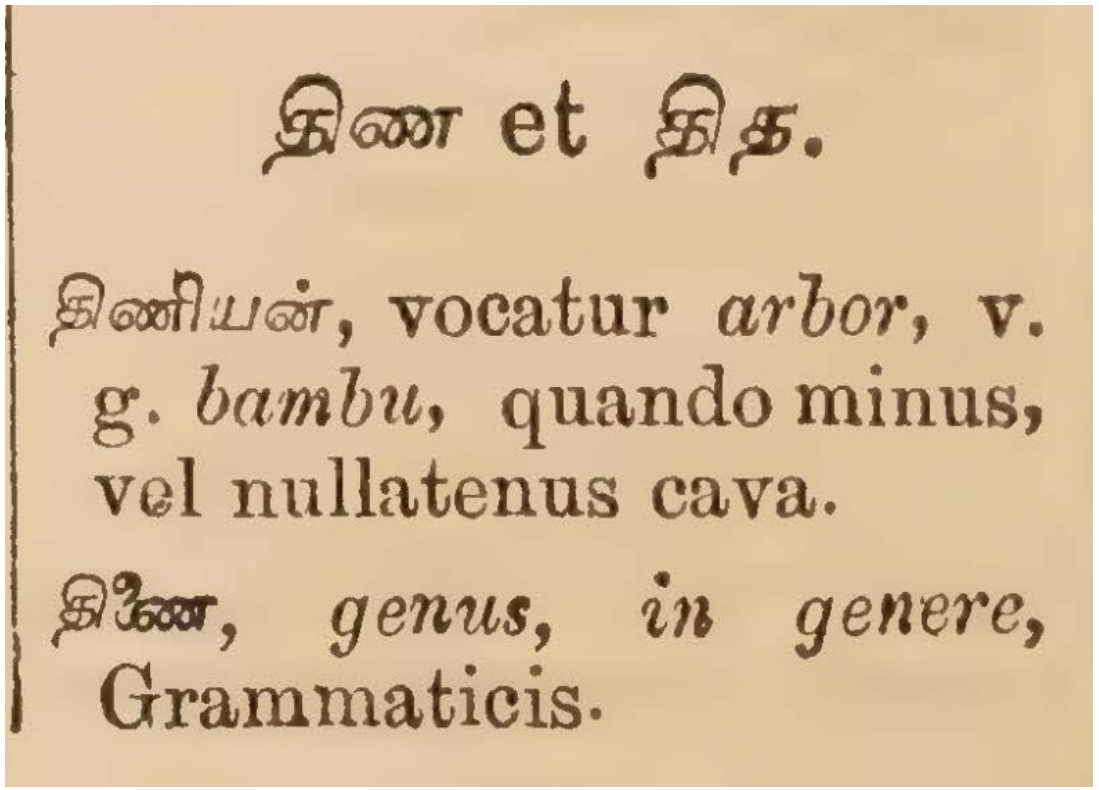
16a. **Sanskrit terms:** (44Rc) அக்ஷி#[1]ம [akṣaram ← akṣara] "Letra, character"; (384Rk) உசசரிததல [uccarittal; from skt. uc-car] "Pronunciaçaõ"; (276Rb) கவி [kavi] "verso"; (79Ra) கி#நதகறன [kirantakaṟaṉ ← granthakara?] "Compositor do liuro."; (79Rj) கீறவ#ணம [kīṟvāṇam] "A lingoa grandonica" (cf. MTL946 : kīrvāṇam [gīr-vāṇa]); (450Re) சபதச#{ஸதர}[2]ம [captacāstram ← śabdaśāstra] "Gramatica"; (474Rj) ச#ததி#ம. Quod ச#{ஸதர}ம; (243Ri) பெ#{ஸத}கம [postaka ← pustaka] "Liuro"; (474Re) ச#{ஸதர}ம [cāstram ← śāstra] "Sciencia, aliquando agouro"; (65Rf) துவிபப#ஷி [tuvippāṣi ← dvibhāṣi] "Topàs, interprete. Homẽ que ſabe falar duas linguas"; (158Lm) ந#மம [nāmam ←

How the 1679 VTCSP mostly ignores Tamil grammatical terminology and displays sanskrit terms (after Chevillard [2015: 122])

aliquando agouro"; (65Rf) துவிபப#ஷி [tuvippāṣi ← dvibhāṣi] "Topàs, interprete. Homẽ que ſabe falar duas linguas"; (158Lm) நா#மம [nāmam ← nāma] "Nome"; (215Rj) ப#ஷை. [pāṣai ← bhāṣā] "Lingoa, que se fale"; (246Ls) புலிங்கம [puliṅkam ← puṃliṅga] "Genero Masculino"; (250Rb) பூதக#ல [pūtakālam ← bhūta kāla] "Tempo preterito"; (107Lb) லிங்கம [liṅkam ← liṅga] "Sexo, ou genero"; (106Ro) லிபி [lipi] "letra, carecter"; (107La) லிபிககிறது [lipikkiṟatu] "Escreuer"; (415Ra) ... வறதம#ன க#லம [vaṟtamāṉa kālam ← vartamāna kāla] "Tempo presente"; (447Lq) விசெஷியம [viceṣyam ← viśeṣya] "Sogeito do predicado"; (436Lq) விய#ககிய#னம [viyākkiyāṉam ← vyākhyāna] "Comento";

திணை [tiṇai]

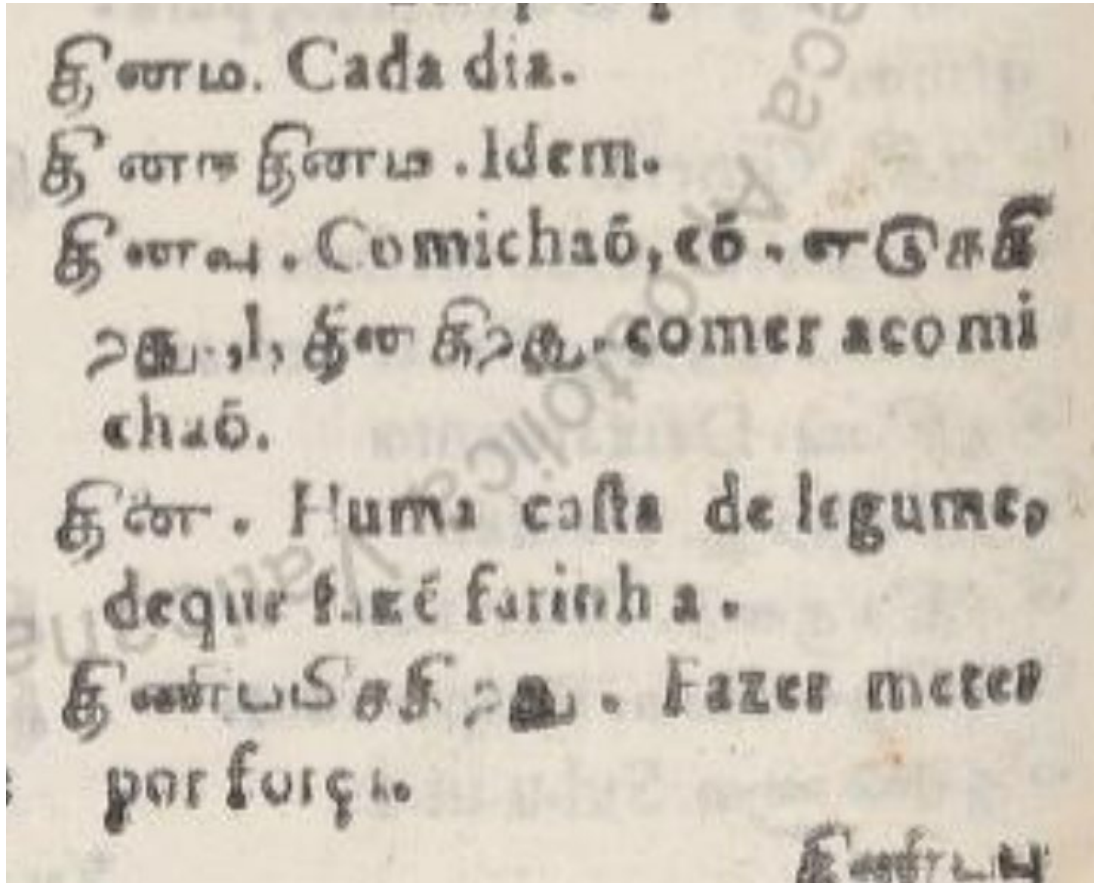
present in Beschi [1742/43] vs. absent in VTCSP (1679)



திணை et திக.

திணியன், vocatur *arbor*, v. g. *bambu*, quando minus, vel nullatenus cava.

திணை, *genus*, *in genere*, Grammaticis.



தினமு. Cada dia.

தினதினமு. Idem.

தினவு. Comichaõ, எ. ன. ஒருகி ஆல், தினகிறது. comer a comichaõ.

தின. Huma casta de legume, deque faz farinha.

தினப்பகிறது. Fazer mete por força.

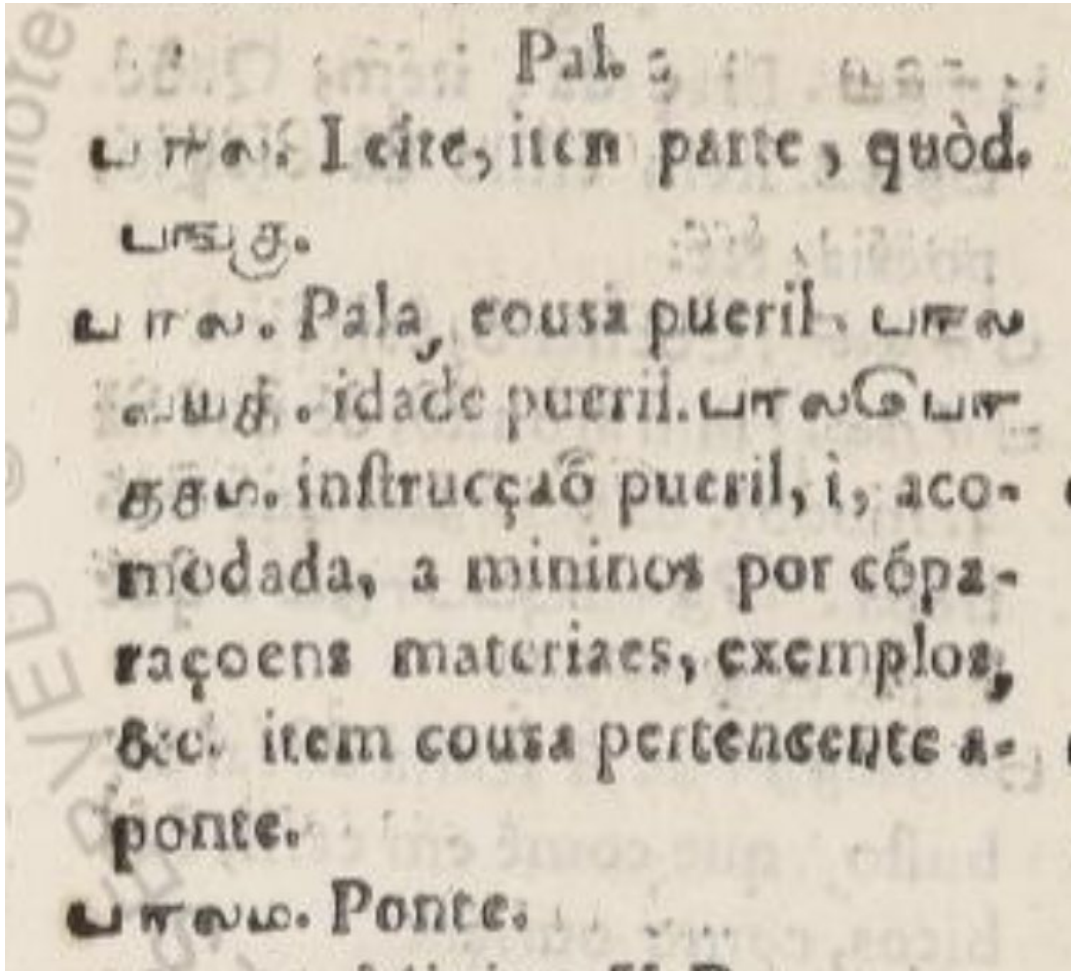
பால் [pāl] (as a TECHNICAL term)

present in Beschi [1742/43] vs. absent in VTCSP (1679)

பால், 1° *lac.* 2° in ling. sublimiori, *genus ac numerus* : quintuplex est.

பால்சோற்றி, vel பாற்சோற்றி, *certa planta.*

பால்மரம், vocatur *quaelibet arbor, quae succum lacteum habet,* ut அத்தி, ஆல், & similia.


Pal.

பால். Lcite, item parte, quòd.

பாகு.

பாலை. Pala, cousa pueril. பாலை யெத்தை. idade pueril. பாலோபதேசம். instrucçaõ pueril, i, acomodada, a minimos por cõparaçoens materiaes, exemplos, &c. item cousa pertencente a ponte.

பாலமை. Ponte.

வினை (viṉai) in Beschi[1742/1743] & VTCSP[1679]

வினை, 1° *actio.* Hinc நல் வினை, தீவினை, *bona, et mala actio.* 2° வினைச்சொல், est *verbum*, ut distinguitur a nomine. 3° absolute *malum.* Sic இப்போ வினைவிளையுங்காலம் : nunc tempus, quo abundant mala. Unde proverb. வினைக்காலம்வந்தால் மனைவழியுந் தெரியாது : instante infortunio, viam, qua domum itur, obliviscimur. 4° aliquando *periculum.*

வினையாகன. (Quasi abq. dño.) he nome de hum idolo, que poé nos caminhos cõ barriga grande, cabeça, e tromba de Elefante.

வினைகாரன. H. tracista. prudente, abil.

வினை. Perigo.

வினையம. Sogeiçaõ, humildade, ou docilidade, item treça de gouerno, ou de viuer, item fagacidade.

வினை (viṉai) in Beschi[1742/1743] & VTCSP[1679]

வினை, 1° *actio.* Hinc நல் வினை, தீவினை, *bona, et mala actio.* 2° வினைச்சொல், est *verbum*, ut distinguitur a nomine. 3° absolute *malum.* Sic இப்போ வினைவிளையுங்காலம் : nunc tempus, quo abundant mala. Unde proverb. வினைக்காலம்வந்தால் மனைவழியுந் தெரியாது : instante infortunio, viam, qua domum itur, obliviscimur. 4° aliquando *periculum.*

தீவெட்டி. Tocha, em que metẽ os panos cõ hum canudo de ferro, e os uaõ seuando cõ arcite:

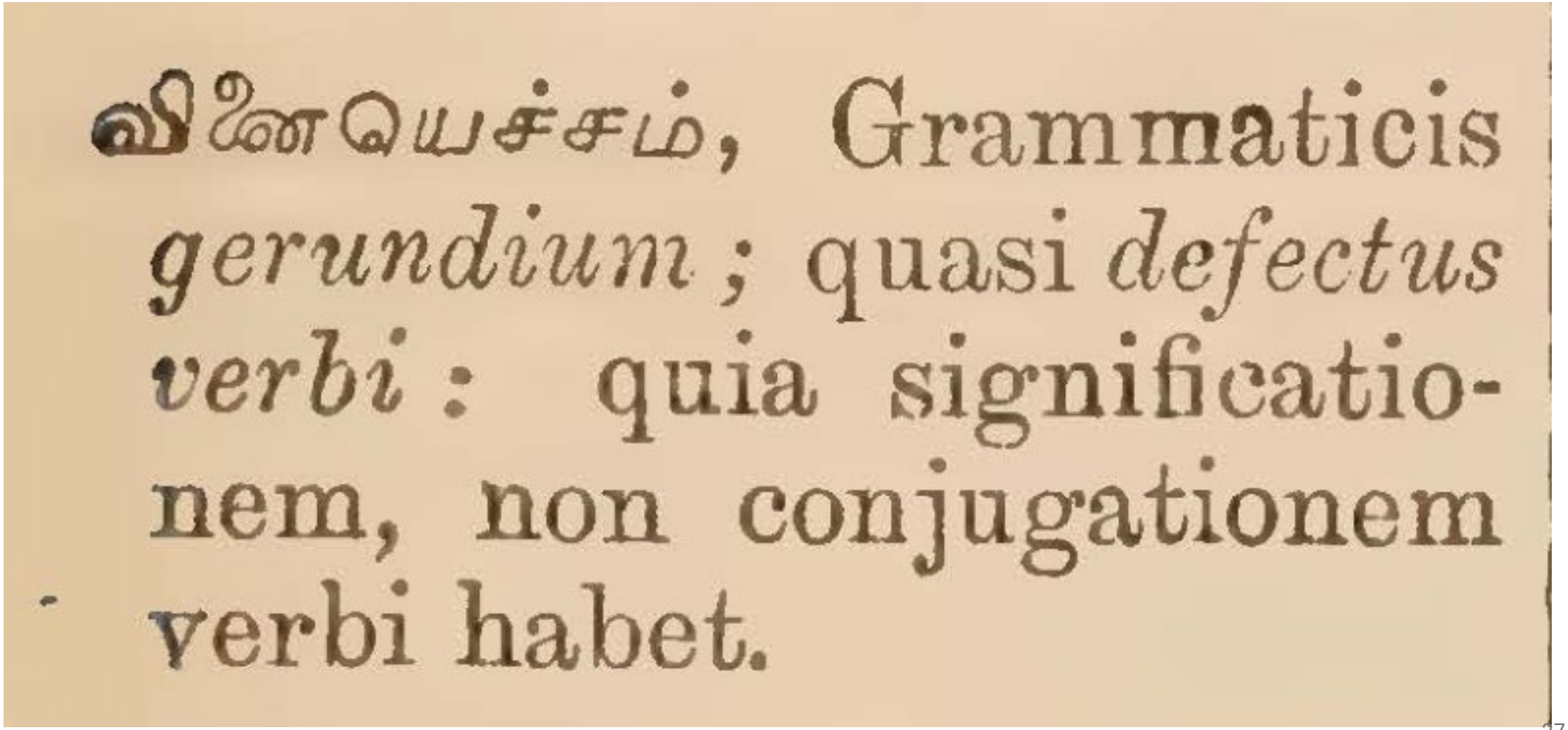
தீவினை. Mal, substantiue.

தீவுதம. Negaça.

தீவுதரூஉ. Idem.

தீனெறி. Doutrina, item profissaõ, que hum faz de ser discipulo.

viṉai-y-eccam (வினையெச்சம்) in Beschi[1742/43] (with non-orthodox explanation)



viṉai-y-eccam (வினையெச்சம்) in Beschi[1738] (with non-orthodox explanation)

§. VI.

De Gerundiis et Participiis.

GErundium vocant Tamulenſes வினையெச்சம் 71
ſive, ac ſi dicerent, *defectivus verbi*: habet enim gerundium vim et ſignificationem verbi; attamen, cum non conjugetur per perſonas et tempora, nonnihil deficere dicitur a ratione verbi. Unicum gerundi-

um

peyar-eccam (பெயரெச்சம்) in Beschi[1738]

cam num. 125, tunc non sequitur hanc regulam.

Participium autem vocant பெயரெச்சம், 71
defectus nominis: habet enim significationem nominis, sed apud Tamulenses non declinatur per casus, genera et numeros, sed eadem omninò vox una participii inservit pro quolibet casu, genere et numero, unde bene deficere dicunt a ratione nominis. Tria sunt participia, præsentis, præteriti et futuri. Participium præsentis fit ablato கு à கிறது, vel கின்றது: sic ex ஆகிறது ஆகின்ற ex செயகிறது, செயகின்ற; ex போகிறது, போக

Beschi becomes a Tamil grammarian:
தொகை in TNV-88/89 (alias தொன்னூல் விளக்கம்)

மூன்றுவது தொகைநிலைத்தொடர்மொழிப்பெயரியல்.

(அஅ-சூ.) தொகைநிலையென்ப தொடரும்பெயரொடு, விளைபெயர்புணர்புளி வேற்றுமைமு தலொழித், தொருமொழிபோற்பல வொன்றியநெறியே.

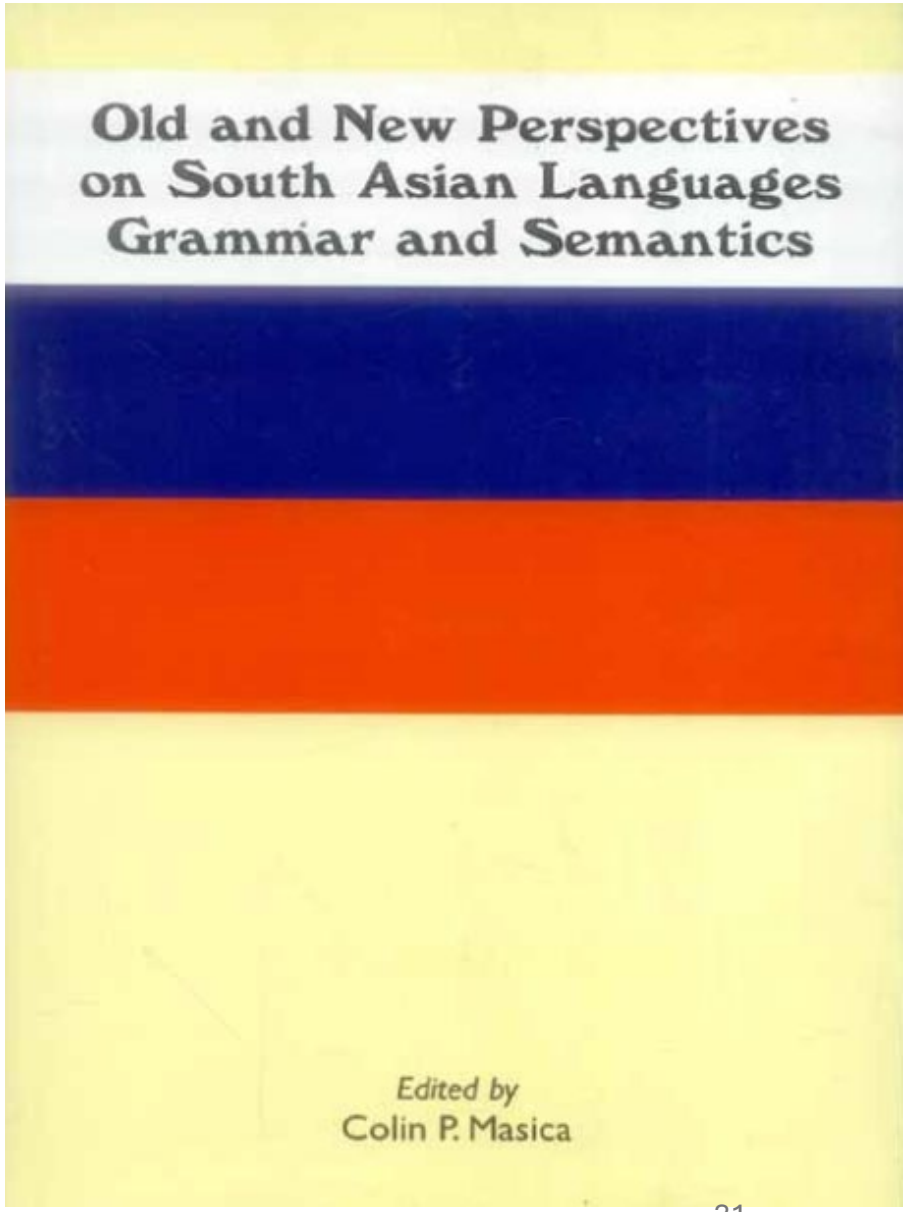
எ-து. மேலேகூறியவிருவகைத்தொடர்மொழிகளினுட் டொகாநிலைவிட்டுத் தொகைநிலையாகத்தொடர்ந்துவருமொழிக்கீண்டுச் சிலவிதியாமாறுணர்த்துதும். பெயருடனேபெயரும், பெயருடனேவினையும்-புணருமிடத்து வேற்றுமைமுதலியவுருபுகள் தோன்றுதொழியநிற்ப இரண்டுசொற்பல சொற்றொடர்ந்ததன்மையால் ஒருபெயர்ச்சொற்றன்மையும் ஒருவினைச்சொற்றன்மையும் வழங்குவன தொகைநிலைத்தொடர்மொழிகளாம். இவற்றை வினைச்சொல்விட்டுத் தொகைநிலையாகத்தொடர்ந்துவ ரும்பெயர்ச்சொற்களாவன. பொருள்-இடம்-காலம்-சினை-குணம்-தொழில்-என்றிவ்வாறுவகைப்பெயர் தொக்குநிற்பனவாம். [உ-ம்] பூண்மார்பன்-மலையருவி-மாரிமாட்பெயிர்-சைம்மா-சருங்குவவா-மெலிடை க்கோமான், எனமுறையே. பொருளாதியாறுவகைப்பெயர்தொக்குநின்று தொகைநிலைத்தொடர்மொ ழிவந்தவாறுகாண்க. இவற்றையிரிக்ச-பூவணையணிந்தமார்பன்-மலையினின்றுவீழ்கின்றவருவி-மாரிகாலி லுள்ளபெயிர்-கையையுடையமா-கருமையையுடையகுவவா-கொடையையியற்றுகின்றகோமான் எ-ம். வரும். எ-று. (க)

(அங-சூ) தொகைநிலைவகைப்படின் றொகும்வேற்றுமைவினை, யுவமைபண்பும்மையோ டன் மொழியாறே.

எ-து. தொகைநிலைத்தொடர்மொழிகளாமாறுணர்த்துதும். வேற்றுமையுருபுதொக்குநிற்பது வேற்றுமைத்தொகையும். காலத்தைக்காட்டாமல்வினையிற்றிரிந்தபெயரொச்சம்தொக்குநிற்பது, வினை த்தொகையும். உவமையுருபுதொக்குநிற்பது, உவமைத்தொகையும் குணப்பெயர்தொக்குநிறபது, பண் புத்தொகையும். உம்மைதொக்குநிற்பது, உம்மைத்தொகையும் சொன்ன ஐவகைத்தொகையிலொன்று வந்ததனிறுதி. மற்றொருபெயரோயாயினும் பதபதவிதியாலுந் தொக்குநிற்பது, அன்மொழித் தொகையும். எனத்தொகைநிலைமொழி- அறுவகைப்படும். [உ-ம்] நிலங்கடந்தான்-பொற்குடம், வேற் றுமைத்தொகை. பொருளம்-வாழ்குடி, வினைத்தொகை. நறுமலர்-ஆனித்திங்கள், குணத்தொகை. பொ

Exploring the topic of தொகை (tokai), [corresponding to *samāsa*] by confronting Tamil grammatical literature (ilakkaṇam) and Tamil belles-lettres (ilakkiyam), while revisiting a 2007 article

- Chevillard, Jean-Luc, 2007, "Syntactic Duality in Classical Tamil Poems", in Colin P. Masica (Ed.), Old and New Perspectives on South Asian Languages: Grammar and Semantics, MLBD Series in Linguistics Vol. 16, Motilal Banarsidass, Delhi, pp. 177-210.



Coulson as an initial inspiration for examining Tamil „compounds“ (tokai)

SYNTACTIC DUALITY IN CLASSICAL TAMIL POEMS

JEAN-LUC CHEVILLARD

"Since the punctuation of compounds is my own responsibility, it is no doubt up to me to try to justify the practice. [...] So far as the grammatical analysis itself is concerned, I am anxious not to give the impression of imagining that experienced Sanskritists stand in the same need of it as an aid to interpretation as I do myself. One might nevertheless urge its merits as a simple convenience to all readers, like word-division, the use of which does not imply that scholars would be unable to separate the words out for themselves. However my case rests on quite different grounds, those supplied by Hertel to justify word-division itself, namely that its use provides the simplest form of running commentary. To justify his choice of readings, an editor must make clear how he construes his own text."

(M. Coulson, *Mālatīmādhava*, introduction, p. liii)

Coulson as an initial inspiration for examining Tamil „compounds“ (tokai):

„To justify his choice of reading, an editor must make clear how he construes his own text.“

"Since the punctuation of compounds is my own responsibility, it is no doubt up to me to try to justify the practice. [...] So far as the grammatical analysis itself is concerned, I am anxious not to give the impression of imagining that experienced Sanskritists stand in the same need of it as an aid to interpretation as I do myself. One might nevertheless urge its merits as a simple convenience to all readers, like word-division, the use of which does not imply that scholars would be unable to separate the words out for themselves. However my case rests on quite different grounds, those supplied by Hertel to justify word-division itself, namely that its use provides the simplest form of running commentary. To justify his choice of readings, an editor must make clear how he construes his own text."

(M. Coulson, *Mālatīmādhava*, introduction, p. liii)

Revisiting a 2007 article: M-syntax vs. U-syntax

conducted[4] and concentrates for the time being on a set of 3 early anthologies of love poetry: the *Kuṟuntokai* (KT, 401 short poems, 4 to 8 lines), *Naṟṟiṇai* (NA, 400 poems, 9 to 12 lines) and *Akanāṉūṟu* (AN, 400 long poems, 13 to 31 lines). The results presented here will mainly be drawn from 4 corpora, listed here by ascending number of lines.

- CK [152 lines] (a subset of KT): the 38 poems having 4-lines inside the KT
- NE [658 lines] (a subset of AN) consisting of the 40 *neytal* poems[5]
- KT [2504 lines] (the complete anthology)
- NA [4180 lines] (the complete anthology)

Revisiting a 2007 article: M-syntax vs. U-syntax

A dual Syntax: M-syntax vs. U-syntax

The main point which I would like to emphasize in this study, is that there is a duality in the syntax of these texts contrasting two poles, which I shall characterize as a **marked syntax** (or M-syntax) opposed to an **unmarked syntax** (or U-syntax), and that most of the poems that belong to the corpus bear traces of these two syntaxes, in ways which I shall try here to quantify. The roots of that phenomenon probably lie in the diglossic situations that have been existing at the various periods when the poems were composed: M-syntax is probably rooted in the

Revisiting a 2007 article: M-syntax vs. U-syntax

the poems were composed: M-syntax is probably rooted in the "vernacular speech" that was in use at that time, although the poetical constraints certainly restricted/filtered the spectrum of what was put inside the mould; on the other hand, U-syntax would draw its strength/impetus/efficiency from the poetical craft of learned men vying with (or emulating) each other in the use (and the enrichment) of a stock of formulae, where the use of what traditional CTamil grammar calls *tokai* (or *tokuti*), usually translated as "compounds", is frequent.

The terminology which I shall use in describing the CTamil texts is ultimately based on the one found in the oldest preserved CTamil grammar, the *Tolkāppiyam* (T), subsequently reworked by later grammars like the *Naṉṉūl* (N), and partly adapted into western languages

Revisiting a 2007 article: M-syntax vs. U-syntax

expression of tense.

My description of M-syntax will be quite incomplete. The reason is that it is much more complex than U-syntax. I shall in fact concentrate on those features in M-syntax which have a dual feature in U-syntax. One such area, in the case of the verbal paradigm, is the use of the two main types of *peyar-eccam* "relative participles", designated here as V^{PE1} and V^{PE2}, which is echoed in U-syntax by constructions using plain verbal roots, designated here as V^0, which could be considered as an unmarked relative participle. Both will be presented, in a preliminary fashion, in the next paragraph.

Revisiting a 2007 article: M-syntax vs. U-syntax

Argument positions around a relative participle

The *peyar-eccam* forms are so called because they are (syntactically) "incomplete" (*eccam*) forms, which have to be completed by a *peyar* "noun", which immediately follows it. We shall refer to this noun, N_1, as the "head noun" and shall consider it as the most prominent syntactic argument for the participle. Many relative participles have other arguments, N_2, N_3, etc., standing on the left, and a typical template for these constructions can be represented as:

$$\{((\ldots N_3) \ldots N_2)V^{PE}N_1\} \text{ (where } V^{PE} \text{ could be } V^{PE1} \text{ or } V^{PE2})$$

The difference between the V^{PE1} and V^{PE2} types is temporal (or aspectual) but syntactically they behave in the same way, although we shall see later that the most frequent semantic roles for their arguments

Revisiting a 2007 article: M-syntax vs. U-syntax

- *utirnta **kaṉi*** (fall ^{PE2} ripe-fruit) "**fruit** which fell" [AN 380-4/5]
- *utirum ... **malar*** (fall^{PE1}flower) "**flowers** which fall" [AN 199-2/3][8]
- *aṭṭa **mallar*** (kill^{PE2} warriors) "**warriors** which killed" [KT 34-5]

I shall refer to such examples as instances of $\{V^{PE1}N\}$ or $\{V^{PE2}N\}$ constructions. These examples, which are "real" examples extracted from the CTamil corpus, are simply the most elementary ones, inside a class of examples where we most often find more than one nominal argument, as in the following two examples, which are instances of, respectively, $\{N_2V^{PE2}N_1\}$ and $\{N_3...N_2V^{PE1}N_1\}$ constructions.[9]

Revisiting a 2007 article: M-syntax vs. U-syntax

constructions. These examples, which are "real" examples extracted from the CTamil corpus, are simply the most elementary ones, inside a class of examples where we most often find more than one nominal argument, as in the following two examples, which are instances of, respectively, $\{N_2V^{PE2}N_1\}$ and $\{N_3...N_2V^{PE1}N_1\}$ constructions.[9]

- *kāy*$_{N2}$ *utirnta [paiṅ kulai]*$_{N1}$ (unripe-fruit fall^{PE2} green bunch) "**green bunch** from which unripe fruits have fallen" [AN 21-20][10]
- *[puli ēṟṟai]* $_{N3}$ ... *kaḷiṟu* $_{N2}$ *aṭūum* ... ***veṟpu***$_{N1}$ (tiger male elephant-bull kill^{PE1} mountain) "**the mountain** where the male tiger kills the elephant-bull" [NA 36-1/4]

Revisiting a 2007 article: M-syntax vs. U-syntax

It should also be noted here that, although the inflectional paradigm of the noun is primarily described in terms of case by traditional CTamil grammarians, the multiple nominal arguments N_1, N_2, N_3 that appear in these examples are not morphologically marked and this is quite typical of the language of that corpus, where case morphemes are not at all frequent.[11]

This information being given, I shall now give parallel examples for the *viṉait tokai* "verbal compound", which results, as already noted, from the combination of a verbal root V^0 and of a noun. In the case of the verbs *utirtal* and *aṭutal*, the V^0 are *utir* and *aṭu* and we can see them used in the following examples, combined with a noun,

Revisiting a 2007 article: M-syntax vs. U-syntax

- *aṭu pōr* (kill^{V0} war) "deadly battle" (NA 39-9)
- *aṭu piṇaṉ* (kill^{V0} corpse) "killed corpse" (NA 329-2)
- *utir vī* (fall^{V0} flower) "flowers that drop" (AN 99-7)

Just like the V^{PE1} and V^{PE2}, the V^0 (or plain verbal root) is often found with two arguments, like what we have in the following example, which I shall describe as an instance of $\{N_2V^0N_1\}$:

- *ikal aṭu munpiṉ* (enmity^{N} kill^{V0} strength$^{N\text{-}obl.}$) "[tusks having the] **strength** to kill enemies" (AN 172-9)

Revisiting a 2007 article: M-syntax vs. U-syntax

Analysis of a two-lines sample text

In order to explain more precisely what I refer to when talking about a dual syntax, I shall start with the analysis of a 2-lines passage extracted from a 15-lines poem in the AN anthology. This will also serve as an illustration of the difficulties inherent in dealing with classical poetry, and this is why I shall first give it in its metrical form, accompanied by an unpublished English gloss going back to the late V. M. Subramanya Ayyar (henceforth VMS).

mutumarat tuṟaiyu muravuvāy mutuputṭ
katumeṉṉak kuḷaṟuṅ kaḷutuvaḷaṅ karainā̤ḷ [AN 260, 12-13]

"At midnight when the demons start to roam about, the old rock-horned owl which has a broken beak and is always residing in the old tree, ..., suddenly cries." (VMS)

Revisiting a 2007 article: M-syntax vs. U-syntax

If we separate the metrical feet into words and additionally dissolve the sandhi, while keeping as an information the fact that there was originally a sandhi change[12], we obtain a text, which is the following, in which I have isolated 14 strings.

mutu marattu{S^{aut}} *uṟaiyum*{S^{aut}} *muravu vāy mutu puḷ*{S$^{?sign}$}
katum-eṉa{S^{aut}} *kuḻaṟum*{S^{aut}} *kaḻutu vaḻaṅku*{S^{aut}} *arai-nāḷ*

---

[12] This information must not be lost, because it can have consequences on how one must interpret the syntactical relations. However, many sandhi rules can be described as "automatic", and they shall be indicated here by {S^{aut}} but these "automatic" sandhi rules are most of the time not applied by modern editors, which might give the impression to the reader that they are optional. The significant sandhis shall be indicated by {S^{sign}} or by {S$^{?sign}$} (when there is a doubt).

Revisiting a 2007 article: M-syntax vs. U-syntax

Using as exponents grammatical tags, intended as short-cuts[13] into the morphology of Tamil and soon to be explained, and giving the meanings, the 14 strings reduce to 12 items, because *katum-ena* can be analysed as an ideophonic expression, containing the particle *ena* as a constituent, (see Chevillard 2004: 416) and because *arai-nāḷ* "midnight" is best treated as an unanalysable idiom, although *arai* "half" and *nāḷ* "day" could exist as stand-alone nouns. The 12 items are as follows:

old^{A0} tree$^{N\text{-}OBL}$ reside$^{V\text{-}PE2}$ broken^{D} beak^{N} old^{A0} bird^{N}
suddenly$^{ideoph\text{-}ena}$ cry$^{V\text{-}PE2}$ demons^{N} roam^{V0} midnight^{N}

Revisiting a 2007 article: M-syntax vs. U-syntax

old^{A0} tree$^{N\text{-}OBL}$ reside$^{V\text{-}PE2}$ broken^{D} beak^{N} old^{A0} bird^{N}
suddenly$^{ideoph\text{-}ena}$ cry$^{V\text{-}PE2}$ demons^{N} roam^{V0} midnight^{N}

For reasons which are ultimately rooted in the traditional grammatical analysis of CTamil "compounds", three phrases inside the text have been underlined. They constitute what I designate as the domain of U-syntax (unmarked syntax), and I shall call them "U-components". A further step in their analysis, which shall shortly be explained, can be accomplished by adding some punctuation marks, adapted from Coulson [1989]:[14]

- (old^{A0}:TREE) "old tree"
- ((broken^{D}:beak^{N})-(old^{A0}:BIRD^{N})) "old bird with a broken beak"
- ((demon^{N}-roam^{V0}):MIDNIGHT^{N}) "at midnight, when the demons roam about"

Link towards the rest of the article

https://gokulam.net/my_articles/2007/