BEI 34 (2021) : XX-XX

Jean-Luc CHEVILLARD

Travel guides for European students of Tamil across the centuries: the Paris, Rome and Lisboa stations

ஐரோப்பியத் தமிழ் இரட்டை வழக்கு ஆற்றுப்படை

Hommage à Élisabeth: உலகம் சுற்றும் மன்னுயிர்

No one acts in complete solitude, who wishes the result of one's actions to survive one, although, in what is basically a collective quest, there are always moments when a person has to make a step forward and to walk alone into the unknown, before often progressively coming to terms with the realization that that "unknown" is not unknown to everyone, and has already been explored by others —*nihil sub sole novum*— who are part of the same human network which we collectively build with our lives and our interactions. In the realm of Language studies, there is also the additional paradox that what is someone's object of discovery is another person's inner identity, as a spontaneous mother tongue, or as a long-cultivated cultural universe. My colleague and friend Élisabeth Sethupathy [1952-2018] (henceforth ÉS) and myself [b.1956] have been walking on such multilingual paths roughly at the same period during our stays on this earth: I first knew her as Élisabeth Barnoud and first met with her thanks to François Gros in the eighties when she was a student at Madurai Kamaraj University. At some point in time, it appears useful to perform a mental recollection, seeing things from a greater distance, as if one was

composing an ஆற்றுப்படை, or a “Guide de Voyage”, telling some of the things one has encountered and which one’s “contemporains” have also seen, presenting unfinished ongoing tasks, and explaining to those who will come after how and where to find clues about them. And since such an account has to be by nature limited, I shall here concentrate on two topics which are a natural part of the intersection between a long-running task (started 2,381 days ago in 2013) in which I am currently trying to make progress and the present memorial, in which I shall evoke some of the tasks in which I knew ÉS to be engaged during her lifetime. Those two topics are (1) the Tamil diglossia, our main topic here, and (2) the transmission of the *Tēvāram*, which will be alluded to in the final section.

From Tamil Nadu to France

This presentation being a “hommage” to ÉS, I shall start with a brief quotation from her *Parlons Tamoul*, a book which she published in 2015:

> (CD) Dans cette collection « Parlons... » de l’Harmattan, le tamoul ne pouvait pas ne pas figurer, mais, en même temps, si des millions de Tamouls parlent cette langue tous les jours, ceux d’entre eux qui ont aussi connaissance de la langue écrite, n’ont, pour beaucoup d’entre eux, guère de considération pour la langue parlée qu’il est traditionnel de considérer comme vulgaire et non raffinée. Nous avons choisi ici de ne pas adopter un tel point de vue, mais une vision plus moderne et dynamique de la langue, réfléchissant une évolution récente en milieu tamoul qui rend à la langue parlée toute son importance de langue vivante et en constante évolution (ÉS, 2015 : 14).

In what follows, I shall not use the same terminology as ÉS, preferring, in conformity with some of my earlier writings,[1] to talk about the

[1] See for instance CHEVILLARD (2018).

distinction between “formal Tamil”, which she refers to as “langue écrite”, and “vernacular spontaneously spoken Tamil”, which she refers to as “langue parlée”. The primary reason for my terminological preference is that formal Tamil (and also Classical Poetical Tamil[2]) can be oralized, and is indeed used in oral performances such as for instance making public speeches, giving lectures, singing songs, chanting hymns,[3] by people who are certainly not reading from a book. However, if we suggest that “la langue écrite” refers in fact to “a language which can be dictated and written down VERBATIM without difficulty using the existing writing system”, whereas “la langue parlée” is something which cannot always be written down easily because of the absence of standard conventions for reproducing it in writing, unless one makes a phonetic transcription using the IPA, then both ÉS’s terminology and mine can be reconciled in their intention, and I shall not rest further on that point.

Illustrating ÉS’s declaration in (1), I shall now provide in a chart, immediately below, a small sample containing a few pairs of forms extracted from the glossary which is found on p. 368-409, at the end of her 2015 book *Parlons Tamoul*, and some of those forms will be used as **anchors**, or relating elements, through which connections or parallels can be established between ÉS’s data harvesting, brought from Tamil Nadu to France for the benefit of French speakers, and some earlier occurrences of data harvesting, which occurred as early as the 16th century, and whose traces are found in several libraries, around the world.

---

[2] As I have argued in CHEVILLARD (2018) and in CHEVILLARD (2015b), it is not enough to distinguish two levels, as in the diglossia, but there is also a third level, which is Poetical Tamil, rooted in Classical Tamil.

[3] Singing and chanting concern the *Tēvāram* and the *Nālāyira Tivviya Pirapantam*.

#	Formal Tamil	Vernacular Tamil	Entry
1	வரவேற்று	வரவேத்து	accueillir (p. 368)
2	வயது	வயசு	âge (p. 368)
3	அப்போது,அப்பொழுது	அப்போ, அப்ப	alors (p. 369)
4	ஆண்டு, வருடம்	வருஷம்	année (p. 369)
5	கற்றுக் கொள்	கத்துக்கொ, கத்துக்க, கத்துக்கிட்டு	apprendre (p. 369)
6	அன்று, அன்றைக்கு	அன்னெக்கு	ce jour-là, alors (p.373)
7	ஐந்து	அஞ்சு	cinq (p. 374)
8	குழந்தை, பிள்ளை	கொழந்தெ, புள்ள	enfant (p. 381)
9	கண்	கண்ணு	œil (p. 394)
10	கன்று	கன்னு	veau (p. 407)

Chart 1. Diglossic Contrastive Sample taken from Parlons Tamoul (ÉS, 2015)

From Ambalacatta to Rome

We shall now move back in time, from year 2015 to year 1679. That date is an important milestone in the exploration of the Tamil speaking sphere by Europeans. In that year, a book whose posthumous author was Antam de Proença (1625-1666) was published in a place referred to on the title page as “Ambalacatta”.[4] That book remains today as the earliest

[4] That place is currently referenced on Google maps as Ambazhakad (in Kerala) and is not very far from Kodungallur. See: https://www.google.fr/maps/@10.26256,76.289781,19z

printed Tamil "dictionary" —it is however called *Vocabulario*— through which was established the first important real linguistic bridge between Tamil India and Europe. The title of that book is *Vocabulario Tamulico com a Significaçam Portuguesa* (henceforth VTCSP) and a single copy is currently known to exist in the Vatican library in Rome. Fortunately a 87pecializ edition was published in 1966 in Kuala Lumpur by Xavier S. Thani Nayagam [1913-1980], and therefore everyone who makes the effort can try to read it without going to Rome, although it is not easy to find one's way among the 16210 entries which are printed on the 1016 columns contained in the 508 pages of that book, one of the reasons being that those entries are ordered following an arrangement which differs completely from the one normally found in Tamil dictionaries nowadays, although that normal order was known to the VTCSP compiler (see Fig. 3 and Fig. 4). The VTCSP arrangement is explained in the two parallel figures given below, in which the NON-STANDARD scheme of 29 word-initial TAMIL letter subdivisions visible in Figure 2, is an expanded disguise or refinement based on the underlying ORDERLY scheme of 18 possible word-initial LATIN letters visible in Figure 1. Those Latin initials are: ABC DEG ILM NOP QRS TVX, although the number of entries per section differs greatly, which is why they are not printed in the same point size in Figure 1.[5]

[5] I have tried to illustrate graphically the magnitudes of the 18 subdivisions in Figure 1, by giving each Latin letter a point size which is proportional to the square root of the number of entries in the section, which is provided below each letter, starting with A, for which there are 1748 entries. The same is done in Figure 2. More details concerning the counts and the arrangement can be found in CHEVILLARD (2017). I must add here, for the sake of completion, that some of the figures contained in my earlier publications were extrapolated on the basis of a partial entering and have been very



Fig. 1: Latin ordering of the 18 sections of the 1679 VTCSP (with entry counts)



Fig. 2: corresponding NON-STANDARD Tamil SEQUENCE for the 1679 VTCSP 29 subsections (with entry counts).

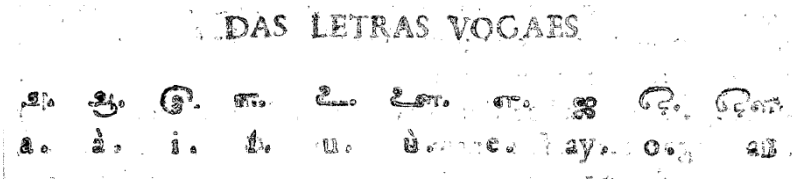


Fig. 3: STANDARD arrangement for Tamil vowels, as known to the 1679 VTCSP, with proposed Latin transcription.

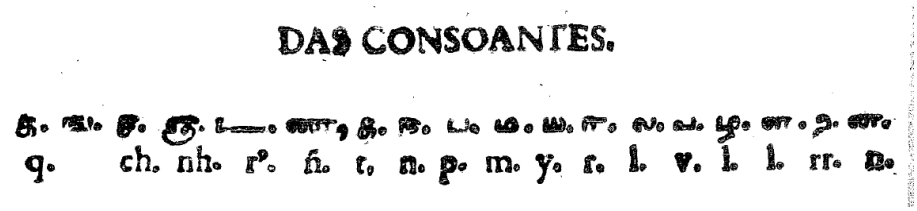


Fig. 4: STANDARD arrangement for Tamil consonants, as known to the 1679 VTCSP, with proposed Latin transcription

Linking the 17th c. Rome and the 21st c. Paris data harvests

We shall now try to connect Chart 1 and Figure 2. More precisely, we shall attempt to find out which elements of information a paradoxical

---

slightly adjusted here. My current figure (in October 2020) for the total number of entries is 16,217. My ultimate goal is to produce an electronic XML-encoded edition of the VTCSP, hopefully to be completed in 2021.

virtual user (henceforth PVU) of ÉS's 2015 *Parlons Tamoul* who would not know French well enough could find in the 1679 VTCSP, provided that PVU could comfortably read 17th century Portuguese. And we shall start with entry 9 in Chart 1, which is said to mean "eye" (French "œil") —but our PVU does not know what the French word "œil" means— and for which we have the competing Tamil forms கண் and கண்ணு. We must credit our PVU with the capacity to read the introductory sections in Portuguese, and especially the one for which the title is DAORDEM, QVE SE GVARDA NESTE VOCABVLARIO,[6] which means "Order followed in this vocabulary".[7] When reading it, the PVU will be told among many things that the 18 consonants of Tamil are க ,ட ,ஞ ,ச ,ங ,... (see Fig. 3, above) and that க frequently corresponds to Q and sometimes to G. Armed with that knowledge, the PVU can search the VTCSP for the word கண், or rather for the form கண, because no புள்ளி is used in the VTCSP above pure consonants. Nothing is found in the small G-section, which goes from p. 77 to p. 80 and which contains 70 entries (as indicated in Fig. 1), but the PVU is successful in his search inside the big Q-section, which goes from p. 252 to p. 323, contains 2251 entries and has 44 subdivisions, namely: "Q.ab.ch. &d.", "Qay.", "Qal.", "Qal. ள", "Qal. ழ","Qam.", "Qan. ந. & ன.", "Qan. ண", "Qan. ஞ. ங.", "Q.ap. & q.", ... As should be clear from this incomplete enumeration, the target of the search is found inside the 8th subdivision ("Qan. ண"), which is spread on p. 262-264. More precisely,

---

[6] I reproduce this title VERBATIM for my readers to understand that capital U was not distinguished from capital V in the VTCSP, although lowercase v was sometimes distinguished from lowercase u. This will also make clear to them why section V in Figure 1 corresponds to the THREE subsections உ ,ஊ & வ in Fig. 2.

[7] This is the translation given by Edgar C. Knowlton Jr. and Xavier S. Thani Nayagam (1966: 13).

the target is on p. 262 and occupies the last three lines of the L(eft) column, as well as the first five lines in the R(ight) column. I shall refer to it by a coordinate which is 262_L_k, the logic being that the ten other items preceding it inside column 262_L(eft) receive coordinates which go from 262_L_a to 262_L_j.[8] However, if we examine the content of the entry, we find a little more than what we were looking for, because the entry starts with:

(CD) க ண. Olho, **item** malha da // rede. [...]

"eye, moreover mesh of a net. [...]" (VTCSP, entry 262_L_k) (see also Chart 2)

It is, of course, not surprising that a word has more than one meaning. When the *Madras Tamil Lexicon* (henceforth MTL) was compiled in the first half of the 20th century, nineteen meanings were enumerated for கண் as a noun, to which were added three meanings for கண் as a particle (see MTL: 683). Interestingly, the meaning "malha da rede" ("mesh of a net") is not among the nineteen enumerated and is only found in the explicit entry வலைக்கண் on p. 3539 of the MTL.

If we now try to locate ÉS's item கண்ணு which was reproduced in Chart 1 at the intersection of row 9 and column 2, we do find something, at entry 263_R_k, but it is not what we expected, being as follows:

(2) க ண ணு. Bezarinho nouo.

"new calf" (VTCSP, entry 263_R_k)

[8] See CHEVILLARD (2017) for more details on the coordinates system.

This indicates that from the point of view of the VTCSP, entry 263_R_k is not associated in the first place with entry 262_L_k,[9] but with another entry found on the previous page, namely 262_L_j, which is:

(3) க ன று. Quòd க ண ணு. Bezerri- // =nho, itẽ aruoresinha, ou plan- // =ta tenra.

“Same as க ண ணு. Small calf, moreover tree sapling, or tender young plant (VTCSP, entry 262_L_j)

This seems to mean that among the Tamil speakers who were the informants of the compilers[10] of the VTCSP, the vernacular equivalent of கன்று was கண்ணு and not the form கன்னு which we find in ÉS’s glossary (see Chart 1, line 10). Such a phonetic divergence is of course not at all surprising: dialectal variation is something existing in most languages, unless some strong grammatical “91pecial” (or “barrage”) is erected by some speakers in order to try to pretend otherwise. What we see in ÉS’s *Parlons Tamoul* is based on one particular dialect, possibly being promoted nowadays as a standard, whereas the VTCSP is based on other dialects which may still be spoken or which may have disappeared.

The VTCSP seen in its dynamic temporal context

Although a unique printed copy of the VTCSP is known to exist in Rome, this does not mean that this is our only source of access to its content, seen as an “échantillon”, or “coupe sagittale”, extracted from

---

[9] The meaning “eye” for கண்ணு is probably overlooked in VTCSP 263_R_k, because Beschi explicitly says that கண்ணு corresponds both to கன்று and to கண். See citations from Beschi in CHEVILLARD (2017 : 122-123).

[10] Proença refers to predecessors, and it is difficult to compare their testimonies with his because the earliest MSS have been lost. We can nevertheless compare the 1679 VTCSP with earlier MSS preserved in Goa. See JAMES (2000 and 2007).

centuries-long campaigns of linguistic data harvesting. As demonstrated by Gregory James (2007), the VTCSP is an expurgated version of earlier MSS preserved in Goa. Additionally, MS copies of the VTCSP are found in several libraries around the world, including the BnF in Paris, and I am fortunate enough to be currently a participant in an ANR-DFG project called TST thanks to which some of those MS copies of the VTCSP have been digitized[11] and are now accessible online to everyone who wants to examine them.[12] In this section, I shall briefly provide visual evidence illustrating what happens when a MS becomes the source of a book, as well as what happens when a book or a MS is the source for a MS. I shall also explain the conventions which I follow when transcribing those MSS *verbatim*. The sample chosen is the succession of two entries having the coordinates 262_L_j and 262_L_k inside the VTCSP: Figure 5 contains an extract from a GOA MS (bearing the indication M-34 at the bottom of every page) which I have obtained through the kind help of Cristina Muru (University of Viterbo, Italy) and Hugo Cardoso (University of Lisboa, Portugal); Figure 6 contains the (barely readable) corresponding passage

---

[11] TST means “Texts Surrounding Texts”. See https://tst.hypotheses.org/the-tst-project. The principal investigators are Emmanuel Francis (CNRS) and Eva Wilden (University of Hamburg). To this should be added that the digitizing of some of the BnF MSS has been performed thanks to the support of another Research Project, called NETamil, financed during the period 2014-2019 by the European Research Council (ERC) and for which the principal investigator was Eva Wilden (formerly EFEO and now University of Hamburg).

[12] The direct links for BnF Indien 221 and BnF Indien 222 are:

- https://archivesetmanuscrits.bnf.fr/ark:/12148/cc103757b
- https://archivesetmanuscrits.bnf.fr/ark:/12148/cc103758m/cN1005A

in the VTCSP; Figure 7 and 8 contain the corresponding passage in two BnF MSS preserved in Paris, namely “Indien 221” and “Indien 222”.

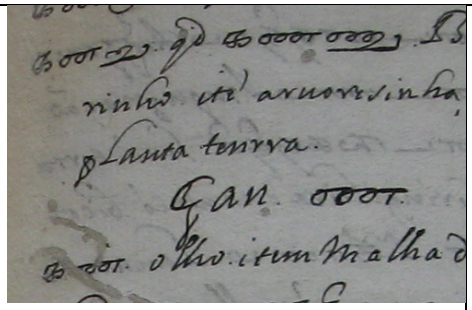


Fig. 5: GOA MS (predates the VTCSP)

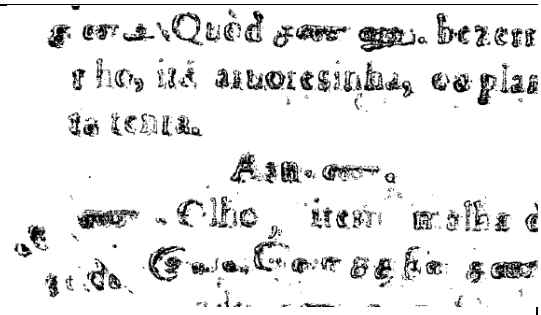


Fig. 6: VTCSP entries 262_L_j and 262_L_k

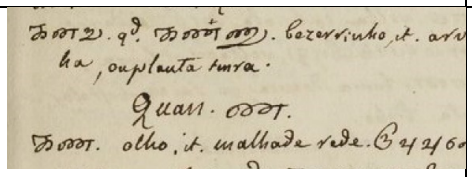


Fig. 7: BnF Paris, Indien 221 (later than the VTCSP)

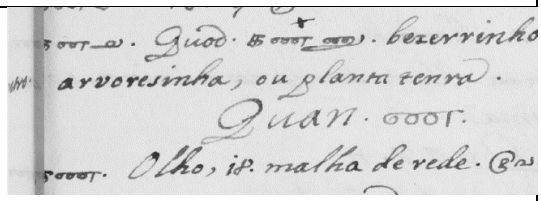


Fig. 8: BnF Paris, Indien 222 (later than the VTCSP)

Chart 2: Comparing the VTCSP with earlier and with later MSS

Remarks on the strict transcription system used here

Since the data contained in Chart 2 has already been presented in transcription inside (1) and (3) above, but without explanations, it seems appropriate that I should now provide here the reader with some technical comments (see Chart 3, below) concerning the strict transcription system which I have used in those examples and which will also be seen in later examples. The conventions used by me for transcribing are as follows:

- no *puḷḷi* (or dot above) is used above the consonants which would have it in the modern spelling.
- a space is inserted after every non-breakable written symbol (alias MS-glyph[13]), even though that may split what is considered nowadays as a single "letter".
- the "@" sign stands for the "ெ◌" sign, which is not yet distinguished from the "ே◌" sign at that period.[14]
- the "#" signs stands either for "ர", or for "ர்", or for the "◌ா" sign (indicating long ā) because they are not distinguished in the writing system used in the 1679 VTCSP. This does not of course apply to றா, னா and ணா which are non-breakable MS-glyphs.
- the "ை◌" sign is represented by "@@" (considered as a single sign). This also does not apply to லை, ளை, னை and ணை which are non-breakable MS-glyphs.
- in the cases where conjuncts have been used in the 94peciali, "{" and "}" are used as delimitors as can be seen below for the conjunct which is analytically represented as {ஸ்தி} inside example 8b.
- change of line is represented by "//"

---

[13] My criterion is practice based: what I call here individual "MS-glyph" is a symbol which could be separated from the previous symbol or from the following symbol by a line break on a MS. In the modern writing system, "கெ" is a non-breakable single symbol, BUT in the ancient practice of manuscripts, which I represent by the sequence "@ க", it consisted of two MS-glyphs, because the detached "horn" (called ***kompu***) which is on the left side of this digraph could be at the end of a line while the க would be found at the beginning of the following line.

[14] Although it initially makes it more difficult for the reader of the transcription, I see at least three arguments for doing this: (1) no preference between "ெ◌" and "ே◌" is conveyed by the use of "@"; (2) one does not have to use a non-standard unicode font (with a special character in the Unicode PUA), as would be the case if one wanted to use the real shape (as I have done in some earlier publication); (3) this makes it easier to represent the practice of MSS where the kompu can be at the end of a line while the associated consonant sign is on the next line (see previous footnote).

Chart 3. Conventions followed inside the strict VERBATIM transcriptions made from the VTCSP and from MSS

The coexistence of “Standard” and “Non-Standard” Tamil in the VTCSP

The most striking feature concerning the data contained in the previous section is that it gives us to see in print (or in writing) words which would never be printed as such in a modern Tamil dictionary. Since an isolated example, such as கணணு in (2), may not have a sufficient power of conviction, I shall now provide more examples, starting with a number of example pairs or triplets, 4a-4c to 9a-9b, where the VTCSP provides us with two concurrent forms connected by the Latin word “quod”,[15] in order to indicate the equivalence, and continuing with examples 10 to 15 where a non-standard form only is registered without being linked with the form nowadays considered as standard. Although I have in fact collected several hundreds of such examples,[16] I shall limit myself to a small (hopefully) representative subset, which is as follows (the approximate English translations from Portuguese are by me):

(4a) பீ டு. Quòd பூ டு. Erua [“grass”]. (Entry 237_R_k)

(4b) பூ ண டு. Quòd பூ டு. (VTCSP, Entry 251_L_c)

(4c) பூ டு Eruasinha. [“small grass”] @ வ ள @ வ ங க # ய ப பூ டு, l [“or”], பூ ண டு, l [“or”], க ட @@ ட. Cabeça de alhos. [“head of garlic”] (VTCSP, Entry 252_L_d)

(5a) அ ய ம ப து. Quod அ ன ப து. ſm- // =couenta. [“fifty”] (VTCSP, Entry 8_L_p)

[15] Similar to what we had in (3): க ன று .Quòd க ண ணு.

[16] At the time of this writing (in december 2020), my current count for those links is 1261.

(5b) அ ன ப து. Sincoenta. [“fifty”] (VTCSP, Entry 18_R_f)

(6a) அ ய ய ம – அ ச ச ம. Medo [“fear”] (VTCSP, Entry 7_R_n)

(6b) அ ச ச ம. Medo. [“fear”] (VTCSP, Entry 3_R_k)

(7a) அ ழி ச ச ல. Quod அ ழி த த ல. (VTCSP, Entry 12_R_a)

(7b) அ ழி த த ல. Idem [“Same”][17] (VTCSP, Entry 12_R_p)

(8a) அ ற றி வ # # ம. Vid. [“see”] அ {ஸ்தி} வ # # ம // Alicese. [“foundations”] (VTCSP, Entry 32_L_l)

(8b) அ {ஸ்தி} வ # # ம. Aliſſeſe [“foundations”] (VTCSP, Entry 38_L_h)

(9a) ப ய ம. Bhayam [transliterated Sanskrit spelling], medo [“fear”], vid. Im p. [“See in the P section”] (VTCSP, Entry 56_R_i [inside B-section])

(9b) ப ய ம. Medo, que se tem [“fear, which one has”]. (VTCSP, Entry 188_R_b [inside P-section])

(10) இ சி க கி ற து. Puxar. [“to pull”] (VTCSP, Entry 96_R_n)

(11) ம த ளை ச @ ச # ல. Palaura mal // pronunciada, como de meni- // =no, ou primeira falla dos meni- // =nos. [Imperfectly uttered speech, like that of children, OR first prattling/babbling of children] (VTCSP, Entry 108_L_h)

(12) ம ணி ப பி ற. Huma casta de // pomba. [“a type of dove”] (VTCSP, Entry 115_L_k)

(13) பி ற. Pomba, ou rola. [“dove, or turtledove”] க # ட டு ப பி ற. Propriamente rola. [“properly turtledove”] (VTCSP, Entry 231_R_m)

(14) க ல # சு. Marinheiro [“sailor”], nas prayas [“coastal dialect”][18] (VTCSP, Entry 255_L_n).

(15) த # ட சி. Humildade. (VTCSP, Entry 354_R_b) [“Humility”]

---

[17] This refers to the preceding entry 12_R_o: அ ழி க கு @@ க. Desfazimento &c. [“dissolving, undoing”]

[18] Literally “on the beaches”.

Briefly commenting on all those entries, we can say that, if we take for instance the Cre-A dictionary (henceforth CAD) as a reference concerning formal modern Tamil, those entries can be classified into FIVE subgroups:

A. those entries which are present (with a modernized spelling) in the CAD with the same meaning. This includes for instance (4b) பூண்டு "garlic", (5b) ஐம்பது "fifty", (6b) அச்சம் "fear", (7b) அழித்தல் "undoing, destroying", அஸ்திவாரம்

B. those entries which modern speakers would describe as phonetically modified versions of words found in the CAD. This includes for instance (5b) அன்பது [non-standard equivalent of 5a], (6a) அய்யம் [non-standard equivalent of 6b], (7a) அழிச்சல் [non-standard equivalent of 7b], (8a) அற்றிவாரம் [non-standard equivalent of 8b], (10) இசிக்கிறது [non-standard equivalent of இழுக்கிறது], (11) மதளைச்சொல் [non-standard equivalent of மழளைச்சொல் "prattling of children"], (13) பிறா [non-standard equivalent of புறா "dove, pigeon"], (15) தாட்சி [non-standard equivalent of தாழ்ச்சி "humility, submissiveness"]

C. those entries such as 9a and 9b which would not be distinguished at all, because initial voicing cannot be represented in the Tamil script, whereas an arrangement such as the one used in the VTCSP distinguishes the small groups of words perceived by Portuguese speakers as having initial B, D and G from the much bigger groups said to have initial P, T and Q (see Figure 1). There are many such pairs inside the VTCSP.

D. those entries which would puzzle modern speakers because, not understanding the Portuguese gloss, they would hesitate between

recognizing an archaic rare poetical word and identifying a "deformed" modern word. This would be the case with (4a) பீடு and (4c) பூடு. The first one could be confused with the literary item பீடு "proud bearing" (attested in the CAD), but is here declared to be a variant of பூடு, which itself is not found in the CAD but would be familiar to those who have listened to the well-known சிவபுராணம்[19] which is the first component inside the திருவாசகம்.

E. those entries such as 14 (in which we see the notation « nas prayas ["coastal dialect"] »), which would be described as "dialectal"[20]

How to decide what to include in a Tamil dictionary? Or How Tamil dictionaries were sometimes/progressively expurgated (or simplified)

In the preceding section, when enumerating items which have disappeared from later dictionaries, I have put the emphasis on phonetic differences, because this is the dimension for which it is the easiest to provide evidence. However, those partly forgotten forms, possibly considered nowadays as phonetically "corrupt" are not the only ones to which a historian of the description of Tamil should pay attention. There are also expressions such as the one exemplified in 16, below, for which we do have early attestations but which are not included in the modern CAD or in the LIFCO பேரகராதி. The example chosen occurs inside the VTCSP as follows:

(16) வ # க க # ட டு கி ற து. Mandar // enuiar carta, prezente &c. ["to dispatch, to send a letter, a gift"] (VTCSP, Entry 412_R_d)

---

[19] See the line 26 inside the சிவபுராணம், which is: "புல்லாகிப் பூடாய்ப் புழுவாய் மரமாகி."

[20] See the notation வ .வ = .வட்டார வழக்கு in CAD, p. xxv.

This VTCSP item, of which we also have a mention in the MTL, and in the Santi Sadhana epigraphical glossary, was already known, as we shall see in (17), to Henrique Henriques [1520-1600] (henceforth HH), who is author of the oldest grammar of Tamil made by a Westerner, a hundred years before the VTCSP, in the 16th c. The MS of HH's grammar (see fragment in Fig. 9) is preserved in the Lisboa BnP, where it was discovered and identified by X. S. Thani Nayagam. It has been critically edited in 1982 by Hans J. Vermeer [1930-2010], and an English translation, by Jeanne Hein [1919-2013] and V.S. Rajam [b. 1942] was published in 2013 as volume 76 in the Harvard Oriental Series.[21] The occurrence of the item which we are discussing is found inside a paragraph illustrating what Tamil grammarians call பெயர் எச்சம், which is frequently referred to nowadays in English as a "relative participle" but for which HH did not yet have a specific terminology, which is why it is drily classified as **"preterit tense, subtense J"** in the English translation by Hein & Rajam (2013: 93). The example in which HH makes use of this item is:

(17) a olla q~ me mãdaſtes ni vara cáttina olæi etc. (HH, BnP, MS 3141, fol. 49r) (see figure 9, line 6) "The letter which you sent me '**ni varacáttina**[244] **olæi**, etc" [Footnote 244: The same form வரக்காட்டின occurs in *Flos Sanctorum* (அடியார் வரலாறு, p. 75) (HEIN & RAJAM 2013: 93).

Because this example comes from a much older layer, I reproduce here, in figure 9, a small excerpt from the Lisboa MS in which it is found, for the readers to see with their own eyes the very real difficulty, which every PIONEER in linguistic encounters has to face.

---

[21] See Review in MURU (2014).

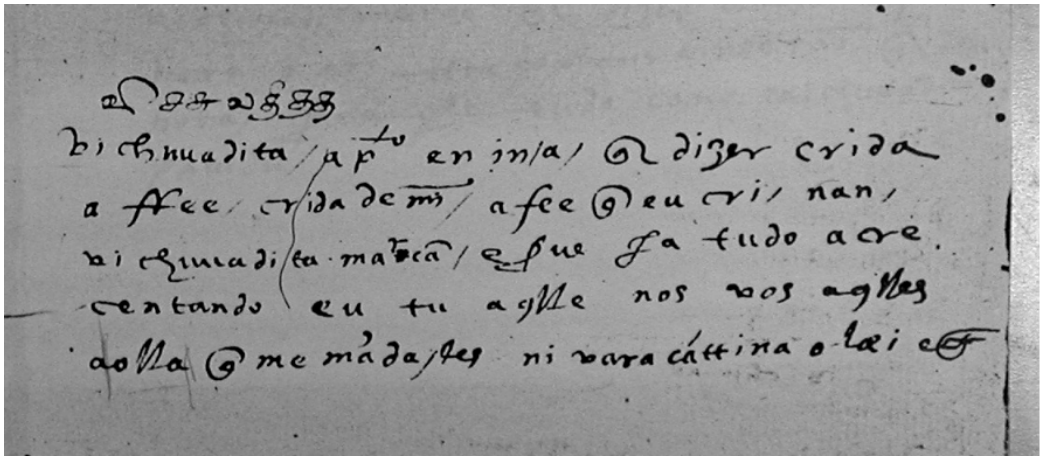


Fig. 9: extract from folio 49r. in HH's Tamil grammar (BNP, Lisboa)

Regarding, however, our main topic, citation (17) includes a lexical remark, by Hein and Rajam, concerning the item வரக்காட்டின, transcribed as **varacáttina** in the original MS (see Fig. 9, line 6). To those informations can be added the fact that the MTL contains an entry for the "same" item, although the MTL makes use of a different citation form, which is:

(19) வரக்காட்டு-தல் vara-k-kāṭṭu- , v. tr. < id. +. To send, as by post or courier; அனுப்புதல். (யாழ். அக.) [MTL, p. 3506]

Importantly, there is inside this MTL entry a reference to the 1842 யாழ்பாணத்து மாணிப்பாயகராதி printed in 1842 in Jaffna. This possibly indicates that the compilers of the MTL considered that this item was not part of standard Tamil usage in their time. We can also note that it is also not present in Winslow's Comprehensive Tamil and English dictionary, printed in 1862 in Chennai. However, that such an expression word has been used in Tamil Nadu at various places and at various times is confirmed by the presence of four dated entries, வரக் காட்ட (=அனுப்ப) (dated 919), வரக் காட்டி (=அனுப்பி) (dated 1444), வரக்காட்டின (=அனுப்பின) (dated 970) and வரக் காட்டுக (=அனுப்புக) (dated 998) inside the *Tamiḻk Kalveṭṭuc Collakarāti* (Vol. 2, p. 537) compiled by

Y. Subbarayalu and published in 2002 by Santi Sadhana. Not being a native speaker of Tamil myself, I leave it to those students of ÉS and of her successors, who have a first-hand knowledge of some dialects of Tamil to tell us whether the item வரக்காட்டுகிறது found at coordinate 412_R_d inside the VTCSP is still nowadays used spontaneously in some parts of the Tamil-speaking world.

Does poetry have practical uses? Beyond diglossia

I shall now briefly touch upon a topic which I have already partly explored elsewhere, but from a very different angle. We have so far distinguished “formal Tamil” and “vernacular Tamil”, on the basis of a statement made by ÉS (see 1), and have tried to see whether the distinctions made in the 21st c. were already there in the 17th c. We cannot take it for granted that the compilers of the VTCSP thought in terms of diglossia, although citation 21 below reveals a high sensitivity for register distinctions. There is indeed an apparent dissymmetry in statements like 4a, 4b, 5a, ... which all fall under the general pattern “A quod B”, which we can roughly translate by “A: is the same as B”. This dissymmetry may have meant that B was more current than A, but it might also simply be the unwanted consequence of an attempt to save printing space and not be the expression of a preference.

The strongest argument for concluding that there was indeed frequently a hierarchy between a (preferred) formal standard and a (deprecated) vernacular, and that Western descriptors progressively adopted that hierarchy from their Tamil teachers, is seen in the PROGRESSIVE EPURATION which we see in successive dictionaries, and also in the (often negative) evaluations made of their predecessors by several descriptors, like when we read inside the book *A Grammar of the*

Tamil Language, published in 1836 by C.T.E. Rhenius [1790-1838], the following statement:

> (20) It is not the object of the above observation to detract any thing from the valuable works of Ziegenbalg, Beschius and others. They did in their days what they could in Tamil literature, and we are greatly indebted to them for the degree of knowledge they have given us of the Tamil language. But they have all failed in giving us pure Tamil; they have mixed vulgarisms with grammatical nicities [sic], and left us in want of a regularly digested Syntax (RHENIUS 1836: ii).

Amusingly, Beschi himself seems to have thought almost the same concerning the VTCSP.[22] However, what is the most interesting in those points of view, as far as I am concerned, who cannot have personal preferences and sees myself as an agnostic historian of all languages descriptions attempts, is how much it differs from the one which was perceptible inside the VTCSP's preface, from which I shall now give a long extract (in English translation).

> (21) Regarding the purity and dignity of the words (*pureza e alteza de palavras*), I well understand that it would be better to include both high and low, pure and corrupt, in accordance with the usage of the various classes of people, so as thus to meet better the range of the language, with particular indication of words used by humble and rural people, and those current only on the Coast* ***{{*Editor's Footnote: Coast is the Pearly Fishery Coast, or the South Eastern Coast of South India. Jaffnapatam is Jaffna, Ceylon}}***[23] and in the Kingdom of Jaffanapatam,

---

[22] See CHEVILLARD (2017), p. 122-123, citations 10a and 10b.

[23] In the original Portuguese, the clause "those current only on the Coast" appears as "das que correm ſó nas prayas". We have already met with the expression "nas prayas" inside example (14). I have so far identified 46 occurrences of the expression "nas prayas" inside the Portuguese glosses given for the 16,217 VTCSP entries, and I

> especially since this vocabulary is intended for the use of priests in that Kingdom and on the Coast. But in as much as to include words current only in those places would require a few years of residence there, I leave this research for another occasion. Nevertheless, I shall include those I can collect from the *Flos Sanctorum* printed on the Fishery Coast and from the vocabulary of F. Ignaccio Bruno prepared in Jaffnapatam, and finally those which are current in this Kingdom of Madura among the lower classes, as far as I have been able to collect them. I shall leave out also poetic words (of which the vocabulary of Fr. Ignacio Bruno is full) because they are useless for practical purposes or for prose, and the Tamil poets have their own vocabularies for them, in which those who are interested and might wish to compose verses may find them (Proença, translated by Edgar C. Knowlton, Jr. & Xavier S. Thani Nayagam (1966: 12-13).

Although this extract from the preface of the VTCSP shows that, unlike Ignacio Bruno, Proença had decided not to include “Poetical words”, because of their being “useless for practical purposes or for prose”, nevertheless, a few of them escaped the consequences of that decision, as can be seen in the following entries:

(22a) அ ம பு லி. Mundo. Pal: de uerso. [“world. Poetical expression”] (VTCSP, Entry 15_L_m)

(22b) அ # ங கு. Hum poyal pequeno // nas coſinhas pera porem pane- // =las tigelas &c. [“a small *poyal* in the kitchens for putting pots and bowls”] **item** repartimẽ- // =to na caza, recamara [“**moreover** subdivision inside the house, bedroom [inner room]. **Item** a- // =juntamento, de letrados, [“**moreover** gathering of the learned”] **item** // ilhota no rio: palaura

---

expect to discover a few more, in the process of finalizing the electronic edition of the VTCSP.

de uerſo. [“**moreover** small island in a river: Poetical expression” (VTCSP, Entry 26_R_j)

(22c) ப # ணி. Agoa, palaura de uerso [“water, Poetical expression”] (VTCSP, Entry 211_R_d)

(22d) @ ப ண டூ #. Molher, honori- // =fice, hè palaura de uerso. [“woman, honorific, this is a Poetical expression”] (VTCSP, Entry 217_R_m)

(22e) @ ச ன னி. Cabeça de uerso [“head Poetical”] (VTCSP, Entry 478_L_i)

Of course, as we now know, staying away from poetical language is not what C.J. Beschi did in the following century. And his compiling of the சதுரகராதி, centered on Poetical Tamil, is certainly what allowed him to acquire everlasting fame in Tamil Nadu, under the name of வீரமாமுனிவர். But that is a different story.[24]

The porous border between the literary and the dialectal, and the role of the திசைச் சொல் grammatical category

I would now like to return to the small sample taken from ÉS’s 2015 *Parlons Tamoul*. We have in fact so far discussed only the items found on lines 9 and 10 of Chart 1. We should try at least, for the sake of completion to say something about a few additional items. This may however be a tall order, if we want to do it in depth, as shall be seen presently when we examine in this section the items found on line 6, and see how much time it takes to even start making a serious comparison. That line 6 contains three forms, namely அன்று, அன்றைக்கு and அன்னைக்கு. Turning to the VTCSP, we find only one exact match, namely for அன்று, in entry 18_R_n

[24] See CHEVILLARD (2019).

(see 23a), which is however accompanied, although not necessarily on the same page, by related forms, which I also reproduce here, as follows:

(23a) அ ன று. Aquelle dia. [“that day”] (VTCSP, entry 18_R_n)

(23b) அ ற @@ ற வ @@ # க கு ம. Athe aque- // =le dia. [“Until that day”] (VTCSP, entry 32_L_j)

(23c) அ ன ற ன று. Cadadia. [“every day”] (VTCSP, entry 18_R_j)

(23d) அ ன ற ன று ள ள. C. de cada dia. [“daily”] (VTCSP, entry 18_R_k)

(23e) அ ற @@ ற க க ற @@ ற. Cada dia. [“Daily, from day to day”] (VTCSP, entry 32_L_i)

As already explained, we cannot declare with absolute certitude what the exact status of those forms was with respect to the diglossia, if we postulate that it already existed, because we do not have an explicit statement inside the VTCSP telling us what the difference is between two items in an explicit pair such as the one examined earlier in (2) and in (3). Unlike the modern period, where we potentially have access to native speakers, and can distinguish between their formal and their spontaneous manners of speaking, when the 17th century is concerned we only have written testimony which can only be analysed in a speculative manner. What we can do however is to enlarge the specimen base, drawing from more sources. We shall start with the 2008 edition of the *Cre-A Dictionary of Contemporary Tamil*, which is the best current representative of Formal Tamil. We find in it the following entries, which are relevant for the current examination, and which I partially reproduce:

(24a) அன்று[1] [...] that day; a segment in the past. (Cre-A Dictionary, 2008, p. 86, Right col.)

(24b) அன்று[2] [...] on that (specified) day (Cre-A Dictionary, 2008, p. 86, Right col.)

(24c) அன்றாடம் [...] daily; every day. (Cre-A Dictionary, 2008, p. 86, Left col.)

This can be compared with a 106pecialized source dedicated to Modern Tamil, the 2012 *Spoken Tamil Dictionary*, by T. Parasuraman, which was printed in Pondicherry. It contains the following relevant entries.

(25a) அன்னெக்கி annekki *N* that day (Parasuraman, p. 12)

(25b) அன்னெக்கி annekki *Adv* on that day (Parasuraman, p. 12)

(25c) அன்னாட)ம்(annaaDa(m) *Adv* daily (Parasuraman, p. 12)

This can also be compared with another 106pecialized source, representing a **very different dialect**, **spoken in Sri Lanka**, namely the 2005 book called *An Introduction to Spoken Tamil*, by James W. Gair, S. Suseendirarajah and W.S. Karunatilaka, which was published in Colombo (Sri Lanka). It contains the following entries:

(26a) அண்டைக்கு [...] on that day (GAIR et alii: 360)

(26b) அண்டைக்கு ஒரு நாள் [...] on that day or other (GAIR et alii: 360)

We should also at this stage quote from the MTL itself, limiting ourselves to a few entries:

(27a) அற்றை a<u>rr</u>ai, < அன்று .Adv. On that day; அன்றைத்தினத்தில். அற்றை வெஞ்சமரில் (பாரத. பதினேழா. 2). — adj. 1. Of that day; அந்நாட் குரிய. அற்றைத் திங்கள் (புறநா. 112). 2. Daily; அன்றன்றைக்குரிய. 3. Little, trifling; அற்பமான. அற்றைக்காரியம். (J.) (MTL: 173)

(27b) அன்றை a<u>nr</u>ai, n. < அன்று[1]. That day; அந்நாள். (பாரத. ஆறாம். 24.) (MTL: 184)

(27c) அன்றைக்கு a<u>nr</u>aikku, ind. < id. See அன்றைத்தினம். (MTL: 184]

(27d) அன்றைக்கன்று aṉṟaikkaṉṟu, adv. < id. + அன்று. Daily, from day to day; அன்றன்று. அன்றைக்கன் றிருமடங்கா (திருவிளை. மெய்க்கா. 12) (MTL: 184)

(27e) அன்றன்று aṉṟaṉṟu, adv. < அன்று + அன்று. Daily, from day to day. (MTL: 184)

(27f) அன்று¹ aṉṟu, n. < அ. [K. andu.] That day, then, any time but the present; அந் நாள். (நாலடி. 23.) (MTL: 184)

As we can see, this quest, which was started as an attempt to find out where the border lies for various types of Tamil speakers between the Formal and the Colloquial, ends up somehow inside the Literary, where that which under other circumstances might have been considered as a DEPRECATED[25] form, because of being colloquial, alias vernacular, alias "vulgar", ends up in an enumeration of forms and of meanings which is partly based on the occurrences of those forms inside ancient Tamil literature, accessible to us via the resources of Philology. And in this new context some of those forms might even receive a new collective designation, namely திசைச் சொல், which the *Tolkāppiyam* recognizes as one of the four types of words acceptable in Poetry.[26]

Are there திசைச் சொல் inside the தேவாரம்?

As announced in the introductory section, this exploration of a too vast domain cannot be concluded with very precise answers or with definitive statements. It was rather my intention to end up with some open

---

[25] See ÉS's remark in (1): « ceux d'entre eux qui ont aussi connaissance de la langue écrite, n'ont, pour beaucoup d'entre eux, guère de considération pour la langue parlée qu'il est traditionnel de considérer comme vulgaire et non raffinée. »

[26] See CHEVILLARD (2008).

questions and to link them with a possible intellectual battlefield where those who come after us can try to answer them. I had also announced that I would establish a link between the questions which have been examined here and a collection of texts which has taken up a lot of the intellectual energy which ÉS had to share with others, and also taken a lot of mine. The name which that collection text has acquired in the course of its history is தேவாரம். ÉS has written her Ph.D. thesis, submitted in 1994, on the singing and on the transmission of that collection of texts, and I have tried to make good use of some of the information contained in her thesis while preparing an electronic edition of the *Tēvāram* containing the English rendering by V.M. Subramanya Ayyar, as well as a concordance, which came out in 2007 as a CD-ROM under the name *Digital Tēvāram*.

I shall now provide the readers with TWO citations from the *Tēvāram*. The FIRST one will be a continuation of the exploration started in the preceding section, where we were trying to find out what were the historical antecedents of the diglossic pair, அன்றைக்கு/அன்னெக்கி which appears on line 6 of Chart 1. The SECOND will be connected with a similar question concerning the diglossic pair, ஐந்து/அஞ்சு, which is on the line 7 of Chart 1. Concerning the FIRST, we do find inside the *Tēvāram*, on the 3rd line of the 7th stanza of hymn 2-63, composed by Campantar, the following clause:

(28) அலந்த அடியான் அற்றைக்கு அன்று ஓர் காசு எய்தி “the devotee who suffered privation having received, from day to day, a coin” (English rendering by V.M. Subramanya Ayyar) (see *Digital Tēvāram*: 2-63_(7)-

Regarding the SECOND, we find inside the first half of the 5th stanza of Hymn 6-54, which was composed by Appar, a numeric progression which is:

(29) மின் உருவை ["God who is the brightness in lighting"]; விண்ணகத்தில் ஒன்று ஆய் ["being the rudimentary element of sound in space"], மிக்கு வீசும் கால் தன் அகத்தில் இரண்டு ஆய் ["being the two rudimentary elements of sound and sense of touch in the wind blowing violently."], செந்தீத்- // தன்உருவில் மூன்றுஆய் ["being the three rudimentary elements of sound, sense of touch and light in the form of the red fire."], தாழ் புனலில் நான்குஆய் ["being the four rudimentary elements of sound, sense of touch, light, and taste in water which has the tendency to flow downwards"], // தரணிதலத்து அஞ்சுஆகி ["being the five rudimentary elements of sound, sense of touch, light, taste and odour in the earth."], [...] (English rendering by V.M. Subramanya Ayyar) (see Digital Tēvāram: 6-54_(5))

The question which I cannot answer but which I nevertheless think should be asked is the following: What is the status of forms such as "அற்றைக்கு அன்று" in (28) and of அஞ்சு in (29), knowing that at many other places in their hymns the authors have also used a form such as ஐந்து? Should we believe that they would have used அஞ்சு while speaking spontaneously, reverting to ஐந்து when being formal, or that both forms were equally literary for them?

Abbreviations

BnF : Bibliothèque nationale de Portugal

BNP : Biblioteca Nacional de Portugal

CAD: see Cre-A (bibliography)

ÉS : Élisabeth Sethupathy.

HH: Henrique Henriques (*Arte da Lingua Malabar*.); See Vermeer, 1982, and Hein & Rajam, 2013.

MTL: Madras Tamil Lexicon.

VTCSP: Vocabulario Tamulico Com a Significaçam Portugueza; see Thani Nayagam.

[யாழ். அக.] யாழ்பாணத்து மாணிப்பாயகராதி. தொகுத்தோர்: சந்திரசேகரப் புலவர். American Mission Press, Jaffna, 1842.

Bibliography

CHEVILLARD Jean-Luc, 2008, "The concept of ticaiccol in Tamil grammatical literature and the regional diversity of Tamil classical literature", *in* KANNAN M. (ed.), *Streams of Language: Dialects in Tamil,* French Institute of Pondicherry, Pondicherry, p. 21-50. (A corrected version (2009/12/18) is available on HALSHS (Open Archives): http://halshs.archives-ouvertes.fr/halshs-00442188/fr/).

CHEVILLARD Jean-Luc, 2015a, "The challenge of bi-directional translation as experienced by the first European missionary grammarians and lexicographers of Tamil", *in* AUSSANT Émilie (dir.), *La Traduction dans l'Histoire des Idées Linguistiques*, Librairie Orientaliste Paul Geuthner, Paris, p. 107-126.

CHEVILLARD Jean-Luc, 2015b, "The Discovery of the Tamil Linguistic Complex by Europeans and other Westerners, & the exploration of 'Triglossia'", *ஐரோப்பியத் தமிழ் ஆய்வியல் மாநாடு. சிறப்பு மலர்*, Octobre 10 & 11, 2015, Paris, p. 333-341.

CHEVILLARD Jean-Luc, 2017, "How Tamil was described once again: towards an XML-encoding of the Grammatici Tamulici", *Histoire Epistémologie Langage* (ISSN (8069-0750, vol. 2/39, SHESL/EDP Sciences. P. 127-103.

CHEVILLARD Jean-Luc, 2018. "The Tamil grammatical tradition: a long commute between theory and practice", *Indian Linguistics. Journal of the Linguistic Society of India*, Linguistic Society of India. Deccan College, Pune, 2018, 79 (1-2), p. 25-38.

CHEVILLARD Jean-Luc, 2019. "The Development of the Poetical/Classical Tamil Lexicographical Tradition", *in* John CONSIDINE (ed.), *The Cambridge World History of Lexicography*, Cambridge and New York: Cambridge University Press, p. 148-58.

Cre-A Dictionary, 1992: *kriyāviṉ taṟkālat tamiḻ akarāti. Tamiḻ-tamiḻ-āṅkilam. Cre-A Dictionary of Contemporary Tamil (Tamil – Tamil – English)*, 1st edition, 1992. Chennai. India.

Cre-A Dictionary, 2008: *kriyāviṉ taṟkālat tamiḻ akarāti. Tamiḻ-tamiḻ-āṅkilam. Cre-A Dictionary of Contemporary Tamil (Tamil – Tamil – English)*, 2nd edition, 2008. Chennai. India.

JAMES Gregory, 2000, *Colporuḷ, A History of Tamil Dictionaries*, Cre-A, Chennai, India.

JAMES Gregory, 2007, "The manuscript and printed versions of Antão de Proença's Vocabulario tamulico (1679)", *no:kku*, vol. 1, n.

GAIR James W., SUSEENDIRARAJAH, S. KARUNATILAKA, W.S., 2005 (Second edition), *An introduction to Spoken Tamil*, S. Godage and brothers, Colombo.

HEIN Jeanne and RAJAM V. S., 2013, *The Earliest Missionary Grammar of Tamil. Fr. Henriques'* Arte da Lingua Malabar*: Translation, History and Analysis,* Harvard Oriental Series (v. 76), Harvard University Press, Cambridge, Massachusetts and London, England.

LIFCO: 2008, லிப்கோ தமிழ்-தமிழ்-ஆங்கிலப் பேரகராதி, லிப்கோ ப்ளிஷர்ஸ், Chennai.

MURU Cristina, 2014, "Review of Hein & Rajam [2013]", *Histoire Épistémologie Langage*, Tome 36, Fascicule 2, p. 184-188.

PARASURAMAN T., 2012, *Spoken Tamil Dictionary*, 'Muthu', 35, Veterinary Hospital Street, Karuvadikuppam, Puducherry, India.

RHENIUS C.T.E., 1836, *A Grammar of the Tamil Language,* Church Mission Press, Madras.

SETHUPATHY Élisabeth, 2015, *Parlons Tamoul. Inde*, L'Harmattan. Paris.

SETHUPATHY Élisabeth, 1994, *Le Chant du Tevaram dans les temples du pays tamoul. Au confluent de la bhakti śivaïte et de la musique tamoule*, Thèse de doctorat non publiée, EPHE, Paris.

SUBBARAYALU Y., 2002, *Tamiḻk Kalveṭṭuc Collakarāti* (2 vol.) (compiled by —), published by Santi Sadhana, Chennai.

THANI NAYAGAM Xavier S., 1966, *Antaõ de Proença's Tamil-Portuguese Dictionary, A.D. 1679*, facsimile edition, Kuala Lumpur: University of Malaya.

VERMEER Hans J., 1982. *The first European Tamil Grammar*, A Critical edition by Hans Vermeer, English version by Angelika Morath, Julius Groos Verlag, Heidelberg.

VINSON Julien, 1910, “Le Premier Vocabulaire Tamoul Imprimé”, *Revue de Linguistique et de Philologie Comparée*, Tome 43, p. 74-82.

Winslow’s Comprehensive Tamil and English Dictionary printed in 1862 in Chennai.

CD-ROM

Digital Tēvāram, 2007, with the complete English gloss of the late V. M. Subramanya Ayyar (IFP) accompanied by 6 hours of MP3 audio recordings, various maps and other related material. Co-edition. IFP/EFEO, Pondicherry, 2007. Edited by Jean-Luc Chevillard and S.A.S. Sarma.

MSS

BNF Indien 221 (cf. figure 7)

BNF Indien 222 (cf. figure 8)

BNP MS (cf. figure 9)

GOA MS (cf. figure 5) (bearing the indication M-34 at the bottom of every page)

Travel guides for European Students of Tamil across the centuries

RÉSUMÉ

L'article explore la perception que les Européens qui ont décrit et enseigné le tamoul depuis le seizième siècle ont eu de la diglossie tamoule, en centrant principalement, mais pas uniquement, son examen sur une comparaison des états d'un texte en devenir, à savoir le *Vocabulario Tamulico Com a Significaçam Portugueza* (VTCSP), dont la version de référence est un livre imprimé en 1679, mais qui est aussi connu par des états manuscrits antérieurs et postérieurs. Cette succession d'états illustre les difficultés que l'on rencontre quand on veut rendre compte d'une langue parlée où des dialectes coexistent.