

Saying It in Other Words: Towards a History of the Tamil-Sanskrit Poetic Lexicon

Eva Wilden

For some two thousand years, Tamil and Sanskrit have had a common literary history in the South of India. Naturally this interaction has left traces in the lexicon. The current article will not focus on ideology – rampant today when the *Tamil Lexicon* is under revision and in the process of being purged of all its Sanskrit etymologies¹ – nor on the literary manifestations of that interaction, on which are concentrated the recent analytical efforts of Ciotti & McCann 2021, who, first of all, bring to our attention the fact that Tamil and Sanskrit are not the only players in the field. Instead I would like to start from a simple, but puzzling observation on the language use not of one of the great poetic works, but of a humble, anonymous commentary on one such work. The commentator of the *Akanāṅṅūru*, one of the celebrated Caṅkam anthologies (on which more below), when glossing difficult words, not infrequently offers a Sanskrit word in explanation of a Tamil word. Why would he do that?

Part of the answer is very simple. He made use of lexicographical resources in which the lexical items from the two oldest Indian literary languages appear freely mixed. In those Tamil works, for many though not for all *sūtra*-s roughly 50% of the vocabulary come from Sanskrit or, to a slightly lesser extent, from Prākṛit – reason enough to describe them as crypto-bilingual, which is arguably one of the reasons why they appear to have gone radically out of fashion during the 20th century. What are these works and what does a literary commentator do with them?

The medieval commentaries (in the majority produced perhaps between the 10th and 15th centuries CE) constitute the link between the literary works and the language-oriented theoretical disciplines of grammar, poetics and lexicography. Tamil traditional lexicography goes by the name of *nikaṅṅu*, easily recognised as Sanskrit *nighaṅṅu*-, and comprises a huge number of works

¹ See the new *Tamil Lexicon*, University of Madras, vol. I “a”, 2012.

(James 2000). The domain is rooted in the oldest surviving grammatical treatise, the *Tolkāppiyam*, where we find the first lists of words with rare, poetic meanings (Chevallard 2019a+b). The later *Nikaṇṭu*-s, however, are not simple word lists or glossaries, but thematic collections of semantic fields of the Thesaurus style that at the same time have something in common with cultural encyclopaedias, in verse form for easy memorisation. Put simply, they do not only enumerate the names applicable to various deities, or the many words for “lotus” a poet might be looking for in a particular metrical context, but also, say, the number of jewels, nine, both according to Tamil sources as well as to Sanskrit ones. In other words, there is ordsa Pan-Indian lore of cultural information which is shared by the various lexicographical traditions of the subcontinent, and this type of knowledge is evidently sought after by poets, who need the matter to write poetry on, but also by the commentators, whose task it is to explain poetry. The two earliest Tamil *Nikaṇṭu*-s are the *Tivākaram* (ca. 9th c. CE) and the *Piṅkalam* (ca. 10th c. CE)², and one of the things I want to demonstrate here is that they inform the glosses given by our commentary.

Before going into the sources, it will be necessary to settle another matter first, namely the issue of representation. How to depict two languages with a huge divergence of phonemes and only a partly overlap? The simple answer is, one has to develop two scripts, Tamil and Grantha, with an overlap for the letters both have in common and with sets of special characters on both sides for the phonemes that are not shared. Practically speaking, Sanskrit texts were copied in Grantha script in the South of India, at least in the historical period when texts were still copied by hand and in manuscripts. The problems start when one wants to integrate Sanskrit words into a Tamil text. The following tables of letters illustrate the extent of the difficulties encountered:³

-
- 2 Quite obviously both of them also bear Sanskrit names, namely *divākara*-, “the day maker”, i.e. the sun, and *piṅgala*- and adjective for “yellow” or “golden”, perhaps to be seen as a noun for “gold”. Poetically speaking it would make perfect sense to call a good lexicon the sun or the gold of a poet, but that leaves aside traditional thoughts on the subject which stipulate both titles to be connected to the names of the authors.
 - 3 The matter is a topic of concern also for the grammatical tradition of the early 2nd millennium: cf. Chevallard 2011 on its treatment in the *Nannūl* and its commentaries.

Tamil				Sanskrit						
<i>Vowels</i>				<i>Vowels</i>						
a	ā	i	ī	u	ū	e	ai	o	ō	au
<i>Consonants</i>				<i>Consonants</i>						
guttural:	k	(g)	ñ	guttural:	k	kh	g	gh	ñ	
palatal:	c	(j)	ñ	palatal:	c	ch	j	jh	ñ	
retroflex:	ṭ	(d)	ṇ	retroflex:	ṭ	ṭh	ḍ	ḍh	ṇ	
dental:	t	(d)	n	dental:	t	th	d	dh	n	
labial:	p	(b)	m	labial:	p	ph	b	bh	m	
semi-vowel:	y	r	l	semi-vowel:	y	r	l	v		
alveolar:	ṛ	ṛ̣		sibilant:	ś	ṣ	s			
retroflex proximant:	ḷ			aspirate:	h					

To sum up, the vowels overlap to a large degree, except that Tamil also has short versions of *e* and *o*, while Sanskrit has the additional *ɽ/ṛ̣* vowels as well as the basically defunct *ḷ*.

For the consonants, Tamil has four extra ones that do not have a counterpart in Sanskrit, namely *ḷ*, *ṛ̣*, *ṇ* and *ḷ*. Tamil script is undermarked in that it phonetically distinguishes voiced and unvoiced but does not have the corresponding letters; the matter is settled by position: an intervocalic single consonant of the guttural, retroflex, dental, palatal and labial series is pronounced as voiced. Sanskrit has special letters for the corresponding elements plus an aspirated counterpart for everyone of them, thus bringing up a set of twenty letters to Tamil's five. And finally, Sanskrit has the three sibilants *ś*, *ṣ* and *s* plus the freestanding aspirate *h*. An additional fundamental difference is that Sanskrit allows consonant clusters realised in writing via complicated ligatures, where Tamil allows only double consonants (intervocalic double consonant corresponding to unvoiced) and the combination of a class-nasal with its consonant.

So far the technical constraints, but what about the practical choices? In order to answer such a question, it is necessary to step back behind the print sources, that barrier that with it unifying tendencies and shifting political agendas has distorted so much of the earlier Indian literary cultures. The practical choices made by Tamil scribes can be depicted in a continuum of five stages of representation, the last two of which make loan words (or quotations) invisible since they do no longer make use of any "foreign" letters:

1. Sanskrit words in Grantha
2. Sanskrit letters not found in the Tamil alphabet in Grantha
3. only the sibilants + h in Grantha
4. completely Tamilised
5. Retranslated (calque)

The first option is to write Sanskrit words in a Tamil text fully in Grantha. This is not normally done, at least not in the type of manuscript we are looking at. The second option is to write only those Sanskrit letters in Grantha which have no counterpart in the Tamil alphabet, as well as the consonant clusters which cannot be depicted in Tamil. This is a practice that can actually be observed, especially in the generic portions of a manuscript where there are many Sanskrit elements, namely in the paratexts such as colophons, benedictions and blessings. Here we may observe a tendency to use Grantha letters even for Tamil words or part of them. As such it is possible to find the ubiquitous Śaiva blessing *civamayam*, usually spelt in Tamil as சிவமயம், with a Grantha ś at the beginning – no surprise here – but even with a Grantha final m at the end as ஸிவமயம்.

The third option is the one officially condoned in modern Tamil script and the corresponding fonts, namely extra Grantha letters for only the four missing consonants ś, ṣ, ṣ and h (often plus one extra ligature for the frequent combination *kṣa*). This is the utmost that may remain in print of a manuscript with a more promiscuous usage. Nowadays even the sibilants and h are often eliminated as in option four where no outward sign of anything non-Tamil remains. The advantages from a poetic point of view are obvious: this technique of writing multiplies the number of homophons that may be poetically exploited. As a simple illustration of this point, the Tamil word *caṅkam* may correspond to Sanskrit *sanga*- “union”, *sangha*- “congregation”, or *śaṅkha*- “conch”. The fifth and final option concerns mostly technical vocabulary, for example in the realm of theology. Established terms may be calqued, as in Skt. *veda*-, which may be *vētam* in Tamil, but its well-known calque is *marai*.

This continuum of choices at least partly seems to mirror the cultural process of absorption. On the one end there is total separation of the two languages, on the other end Sanskrit is assimilated to a degree that many (most?) speakers of Tamil are no longer aware of the presence of loan words. What interests me here is some point in between – a stage when Tamil made use of Sanskrit words as evident and frequent elements in the lexic and poetic repertoire, while at the same time their distinctness was marked by the use of

Grantha letters. That Sanskrit words were so perceived is laid down already in the early grammatical tradition. In *Tolkāppiyam Collatikāram* 391i we read:

*iyarcol tiricol ticaiccol vatacol enru
aṇaittē ceyyuḷ iṭṭac collē.*

“**Natural words, refined words, regional words** [and] **Northern words**,
such are the words to be brought together in poetry.”

Here “Northern words” refers to Sanskrit (without distinguishing very clearly between Sanskrit and Prākṛit). And this is, by all appearances, what informed the agenda of the Tamil *nikaṇṭu-s*. As dictionaries they are not necessarily useful, since they are not concerned with the ordinary meanings of words. What interests them is extraordinary meanings, polysemy and homophony. The extent to which Northern elements are found in them will be attested by the *sūtra-s* quoted below. The *nikaṇṭu-s* themselves, however, betray no awareness of linguistic variance; as far as their authors and/or compilers are concerned, they deal with poetic vocabulary. In the current print versions, the word material is moreover completely Tamilised, although it is possible that manuscripts existed (or even exist) that make use of Grantha letters.

Verifying that in detail will be a daunting future task, and for this reason I am making a modest start here by looking at a commentary for which I happen to know the sources since it was part of a critical edition of a poetic text I am engaged in (= Wilden 2018). This poetic text was the *Akanāṇūru*, one of the Caṅkam anthologies, the classical Tamil corpus par excellence, the beginnings of which date back perhaps to the early centuries of the Christian era. It is a collection of four hundred love poems distributed over three books, and for the first ninety of them there is an old anonymous commentary, traditionally attributed to the 12th c. CE. It survives in a single manuscript of the U.V. Swaminathaiyar Library in Cennai (UVSL 297), undated, but as old as they come, which means, in South India, with some luck the 18th century. It has been published first with the *editio princeps* of the first book in 1918, but is more readily available in the reprint of the whole work in 1933.⁴ After that it

4 Volume 3 of Wilden 2018 contains a critical re-edition of the commentary based on the manuscript and the edition, documenting in detail the changes among other things in the choice of script made by the editor Rākavaiyankār. For a more detailed analysis of commentarial strategies, see Wilden 2020.

was supplanted by modern commentaries on the full work, though its glosses naturally left traces in the subsequent works.

Now, an intriguing issue that has to be gotten out of the way before going into the commentary itself is the policy of spelling and glossing the Indo-Aryan words that are already part of the *mūlam* itself. To summarise the general situation briefly, already Vaidyanathan 1971 published a book listing all such words appearing in the Caṅkam corpus and the *Cilappatikāram* (the earliest epic poem or *peruṅkāppiyam/mahākāvya*).⁵ The drawback of that work is that it simply lists the words without going into any detail as to distribution, number of occurrences and derivational forms. The first precise statistics for a few texts come with the recent critical editions (Wilden 2010, 2012, 2018). The situation can be summarised as follows: the Caṅkam corpus already has a share of about 10% Indo-Aryan words in its vocabulary, but only in the very late texts more than just lexical items and proper nouns are taken over, namely wholesale compounds (detectable by Sanskrit *sandhi*), numerals and verb forms. The early loans are never spelt with Grantha letters and for the most part do not even seem to be perceived as specifically poetic vocabulary in that they normally do not trigger glosses on the part of commentators. The earlier type of commentary does not include a paraphrase but only gives glosses for words that are perceived as difficult, either because they are rare, because they occur with an extended meaning that deviates from the expected one, or because they have homophons and are in need of disambiguation.

The glosses found in the particular commentary we are looking at often can be traced back to the two early *Nikaṅṭu-s* already mentioned, the *Tivākaram* (Tivā) and the *Piṅkalam* (Piṅk) – the only two that were available in the stipulated period of the commentator –, just as grammatical explanations usually can be traced back to the *sūtra-s* of the grammatical treatises, especially to the *Tolkāppiyam*. In what follows I am going to present five cases of such glosses and the probable authorities they lean on.⁶

5 To date there is no comprehensive dictionary for Indo-Aryan loan words in Tamil; the *Dravidian Borrowings from Indo-Aryan* are only a slim supplementary volume to the *Dravidian Etymological Dictionary* (which excludes non-Dravidian items) and does hardly more than scratch the surface.

6 All the entries from the *Tivākaram* or the *Piṅkalam* are based on the recent Cānti Cātaṅā edition which provides word splits and a list of headwords. Critical editions of the most important *Nikaṅṭu-s* are among the really urgent desiderata of Tamil studies.

Case 1: a Tamil-Tamil gloss

The first example is not yet a case in point, but a testimony to the fact that commentators (like modern lexicographers still), are aware of the lexicographical resources found in the *Tolkāppiyam* and make use of them. For *Akanāṅṅūru* (AN) 14.16, the commentator glosses a Tamil word with a Tamil word:⁷

AN 14.16 *paiyuḷ* – *nōy*, “pain”

The not-so-frequent noun *paiyuḷ* is glossed with a single second noun, namely *nōy*. Now, in the *Uriyiyal* (“the chapter on appropriate [words]”)⁸ of the *Tolkāppiyam Collatikāram* (TC; the second “book on words”, mostly dealing with morphology and syntax) we find the following word explanation:

TC 335i⁹ *paiyuḷum ciṟumaiyum nōyiṅ poruḷa*.
 “*paiyuḷ* and *ciṟumai* have the meaning of pain.”

The grammarian explains two poetic words by an ordinary one, namely *nōy*, “pain”. While *paiyuḷ* presumably is included here just because it is a poetic word, not used in ordinary language – and I say presumably because in fact we do not have access to the ordinary language of the time, except by what grammarians tell us about it –, *ciṟumai* is an ordinary word meaning “smallness”, but has a secondary poetic usage which is elucidated here. From the point of view of early poetry, the noun *paiyuḷ* is not so rare as to be absolutely in need of a gloss. But the commentator takes the opportunity¹⁰ to demonstrate his awareness of the tradition, perhaps both as a proof of the erudition that is

7 The Tamil Lexicon, while giving for the meaning the expected reference to TC 335i, seems uncertain about the etymology and compares to Pkt. *payyāula-*, which would correspond to Skt. *pariyākula-* (information kindly provided by Harunaga Isaacson). There is, however, a plausible Tamil etymology for this word of moderate frequency, namely a lost verbal root *pai*, “to suffer” (still surviving in the isolated but frequent verbal noun *paital*, “suffering”) in combination with the rarer nominal suffix *-ul*, which is found also, for example with *ceyyuḷ*, “verse”. whatever may be the historical fact, clear is that *paiyuḷ* is part of the long-established poetic vocabulary that is not perceived as borrowed.

8 For an explanation of the title, see Chevillard 2019a.

9 TC 335i: the siglum “i” refers to the counting of the *sūtra*-s according to the commentator ḷampūraṇar.

10 The occasion to do so is fairly rare since the nouns mentioned in the *Uriyiyal* are as yet only about 120 (see Chevillards 2010 for details and precise figures).

expected in one of his trade and as a sign of affiliation to the oldest school of grammar (which is not surprising for someone in the 12th century writing a commentary on what already then was ancient poetry).

Interestingly, also the *Tivākaram*, the oldest of the *nikaṇṭu-s*, takes it from here. Of course the arsenal of poetic vocabulary has multiplied considerably by the 9th century when this first work is devoted to nothing but listing such words. The head noun is still the one used in the *Uriyiyal*, *nōy*, but the number of possible cognates has gone up to 10, including *paiyu!* (though no longer *cīrumai*):

Tivā 8.149 ***nōy*** “pain”:
 maṭaṅkal, nīlāl, maṭi, aṅar, aḷuṅkal,
muṭaṅku, *kuttu*, *piṇi*, ***paiyu!***, *cūr* – ***nōy***.
 “*maṭaṅkal* (bending over?), *nīlāl* (ailment?), *maṭi* (being prostrate?), *aṅar* (grief), *aḷuṅkal* (affliction)
muṭaṅku (crippled condition?), *kuttu* (ache?), *piṇi* (sickness),
paiyu! (sorrow), *cūr* (suffering) [are] ***nōy***.”

This is not the place to go through the list in detail.¹¹ Of course not all of these words simply mean “pain”; they all have semantic shades of their own. Thus it is not reasonable to explain this as a list of synonyms; it rather corresponds to a semantic field where under certain contextual circumstances one word can stand in for another. Note that this *sūtra* does not contain a single Sanskrit equivalent, which is not generally the case, as we shall see in what follows.

Case 2a: a Tamil-Sanskrit gloss found in both old *nikaṇṭu-s*

The second example is a perfectly ordinary Tamil combination of two nouns which is glossed by a single Sanskrit noun:

11 One of the constant problems when dealing with the testimony of the *nikaṇṭu-s* is that many meanings do not seem to be attested in literature, which means the corresponding entry in the *Tamil Lexicon* is based on the thesaurus itself. In the case of *nōy*, there is the additional complication that the old meaning of “pain” was at some time overlaid with a new meaning “disease”. Since the *Tamil Lexicon* was made by modern Tamil speakers, the meaning adduced to all the really rare words here consequently is “disease” or something similar. In my suggestions above I have been guided, where possible, by etymology (with an added question mark), unless the word is known to me from its use in literature.

AN 73.13 *vaya-māṇ* – *siṃham* [சிஓஹ], < *siṃha*- “lion”

Here the Tamil is in need of disambiguation because *vaya-māṇ* literally simply means “strong animal”, which, depending on the context, may refer to several beasts, hence the presence of a gloss. The commentator’s choice may have fallen on *siṃham* as a gloss because it is the simplest unambiguous cognate in a group of words¹² he could have chosen from, or because *ciṅkam* (with the totally Tamilised spelling of the edition) is the headword chosen for the corresponding *sūtra* in the *Tivākaram*:

Tivā 3.1 *ciṅkattiṅ peyar*, “the nouns for lion”:
 āḷi, maṭaṅkal, ari, kaṅṭiravam,
kēcari, mirukāpati, vayappōttu,
cīyam, vayappuli, pañcāṇanam, vayamā¹³,
 (mēviya) ciṅkam (eṇa viḷampinarē).

“āḷi, maṭaṅkal, ari (< Skt. hari-), kaṅṭirava (< Skt. kaṅṭhīra va-), kēcari (< Skt. kesarin-), mirukāpati (< Skt. mṛgapati-), vayappōttu, cīyam (< Pkt. sīha-), vayappuli, pañcāṇanam (< Skt. pañcānana-), vayamā, are suitable for “ciṅkam” (lion), they (~ the scholars) have said.

In this list, I have underlined the Indo-Aryan elements and put in brackets those words that form the backbone of the syntagma. Five of ten nouns are of Sanskrit or Prākṛit origin – note the double implication of *siṃha*- in its Sanskrit form and its Prākṛit form *sīya*-. Only these two are actual words for “lion”; all the rest are either rare, polysemic or metaphorical, in short, poetic words.¹⁴ The choice of spelling in the manuscript – whether by the commentator

12 Presumably Tamil does not have a straightforward word for “lion” since the lion is, in the first place, a poetic animal, not a biological reality.

13 The difference between the *vayamāṇ* of the poem and the *vayamā* of the *nikaṅṭu* is merely that of a morphological variation that does not affect the meaning.

14 Among the Tamil nouns, *āḷi* is a rare word referring either to a mythological beast of prey or later to a lion; *maṭaṅkal* is an extremely polysemic word with no less than 19 meanings attested in the Tamil lexicon; *vayappōttu*, *vayappuli* and *vayamā* are three variants of a noun combined with *vayam*, “strength” (in Tamil a noun in attributive position assumes the function of an adjective, hence “strong”), with *pōttu* “male animal”, *puli* “tiger” (or another big

or simply by the copyist we have no means of knowing – is purely Grantha, albeit combined with the Tamil nominal suffix *-m* (still spelled in Grantha), which arguably points to the object being perceived as a Sanskrit word.

I’m adding here the corresponding *sūtra* from the second *nikanṭu*, the *Piṅkalam*, where the same correspondence of a headword *ciṅkam* with *vay-amā* as one of the cognates can be found, only the list has become even longer in the intervening century:

Piṅk 8.99 *ciṅkam* “lion”:
 āḷi, vāḷari, ari, mirukāpati,
 kōḷari, kantīravam, pañcānanam,
cīyam, maṭaṅkal, kēcari, muṭaṅkuḷai, arimā,
 vayappōttu, pañcanakāyutam,
vayamā, *vayappuli*, *ciṅkap* (*peyarē*)

So here the count has gone up from ten to sixteen, with the same rational of Tamil-Sanskrit distribution, with the difference that here we see two Tamil-Sanskrit compounds, both with *ari* (Skt. *hari*- “the yellow one”), presumably possible because *ari* is part of the oldest stock of perfectly assimilated loan words: *kōḷ-ari* “the murderous yellow one” and *ari-mā* “the yellow animal”.

Case 2b: a Tamil-Sanskrit gloss found only in the 2nd *nikanṭu*

The third example concerns a perfectly ordinary Tamil word which has a rarer poetic meaning, which the commentator glosses with a Sanskrit equivalent:

AN 61.3 *tāḷ* – *utsāha* [உத்ஸாஹ] < *utsāha*- “effort”

In the Caṅkam corpus, the “normal” meaning of *tāḷ* would be “foot”, but the context of this poem requires the poetic meaning “effort”, which is why the commentator feels the need to provide a gloss. Now here his choice is interesting. In the *Tivākaram* we find a *sūtra* which lists *tāḷ* as one of the cognates under the headword *muyarci*, “effort”, incidentally another of the *sūtra*-s where the entries are all pure Tamil:

cat like the “leopard”), and *mā* “animal”. The Sanskrit words are *ari*- “the yellow one”, *kaṅṭhīrava*- “whose sound comes from the throat”, *kesariṅ*- “the hairy one”, *pañcānana*- “the five-faced ~ the frightening one”.

Tivā 9.71 *muyar̥ci* “effort”:
 uñar̥ru, ulappu, tāḷ, āṇmai, āḷviṇai, ūkkam,
iṇar̥ral (*eṇru ivai*) *muyar̥ci* (*ākum*).
 “*uñar̥ru* (zeal), *ulappu* (exertion), *tāḷ* (effort), *āḷviṇai*
 (manly enterprise), *ūkkam* (zeal),
iyar̥ral (endeavour) – these will be *muyar̥ci*.”

According to the model of *arimā-siṃham* one might expect the AN commentator to chose the headnoun *muyar̥ci* as a gloss, but that is not the case, which rules out the *Tivākaram* as a source text. However, a look at the *Piṅkalam* is more rewarding. As to be expected, the number of equivalents has grown, only here it is distributed over two entries:

Piṅk 7.33 *muyar̥ci* “effort”:
muyar̥ci, *ūkkam*, *iyar̥ri*, *āḷviṇai*
uñar̥ru tāḷ tuppū ur̥cākamē.

Piṅk 7.34 *ur̥cākam* “effort”:
uñar̥ru tāḷviṇai tuppū ūkkam ur̥cākam.

Both verses together bring the *Tivākaram*’s count of seven words for *muyar̥ci* up to eleven. But first of all, here we have to correct a blatant mistake in the editorial framing. The headword given for the first *sūtra* by subsequent editors from 1890 onwards is *muyar̥ci*, as in the corresponding *sūtra* in the *Tivākaram*, while syntax – end position followed by sentence-final particle *-ē* – unequivocally demands *ur̥cākam* as the predicate and headnoun, as is done for the extension *sūtra* that follows. Here the editors too see *ur̥cākam* as the headnoun. Since already the first *sūtra* contains *tāḷ* as a possible equivalent, we find here the source for our commentary gloss, because *ur̥cākam* is the completely Tamilised version of Sanskrit *utsāha-*, which our commentator spells in Grantha and just so, without even the obligatory Tamil final *-m*.

Case 3: a Tamil-Sanskrit gloss beyond the *nikaṇṭu*

The fourth example brings us into the range of extreme polysemy and to another type of *sūtra*, that is, here we are not dealing with a list of cognates that share the semantic field of the headword, but with a set of radically different meanings which can all be expressed by the headword. In this case I will start with the *nikaṇṭu sūtra*-s in order to show the semantic range of the noun in question and will then go on to the gloss provided by our Akam commentator.

Tivā 11.255 *ēmam*:
cēmam kāval iṅpam iravu poṅ
ēmāppu mayakkam ēmam (ākum).
“Protection, guard, pleasure, night, gold,
safety, confusion, [these] will be *ēmam*.”

Piṅk 10.201 *ēmam*:
iṅpamum, nīrum, poṅṅum, kāvalum,
iravum, cēmamum, ēmāppum, mayakkum
(eṅpaṇa) ēmap (peyaṅa ākum).
“Pleasure, ash, gold, guard,
night, protection, security, confusion
– [all] these will be the noun *ēmam*.”

The polysemic words illustrate nicely the fact that the *nikaṇṭu* authors did not distinguish between polysemy and homophony (which is irrelevant from a poetic point of view): *ēmam* on the one hand is a well-attested old Tamil word with two meanings which may look contradictory but which could be interpreted as the same thing viewed from a slightly different angle, namely “pleasure” and “confusion” (present in the *sūtra*-s as *iṅpam* and *mayakkam*). Several consecutive borrowings from Sanskrit added three homophons to this. The first borrowing is *kṣema*-, which in the earliest Tamil corpus appears as *ēmam*, while in post-*Caṅkam* times the standard form is *cēmam*. This accounts for the meanings connected with security and protection (here *cēmam*, *kāval* and *ēmāppu*).

The next borrowing is the noun *hema-*, “gold” into Tamil, again with the form *ēmam*, and accounted for by Tamil *poṇ*. The last in the row is *yāma-*, “a night watch”, which again may be represented in Tamil as *ēmam*, interpreted here as a noun for “night” (*iravu*).¹⁵ The odd one out is *nīru*, “ash”, and for the time being I cannot offer either an explanation or at least a literary quotation for that meaning.

Now, when the *Akanāṇṭūru* uses the word *ēmam*, the commentator does not do what is expected, namely to chose the contextually appropriate meaning from the *sūtra*-s (be it *Tivākaram* or *Piṅkalam*) but he comes up with a gloss of his own (or from an unidentified source), and again he proposes a Sanskrit word:

AN 84.5 *ēmam* – *rakṣai* [ரக்சை], < rakṣā- “protection”

Again he uses the Grantha spelling, very overtly un-Tamilian, because in Tamil *r* is not found in word-initial position (the fully Tamilised version would be *iraṭṭai*, attested late and rarely). To me this suggests somebody who does not only have a training in Tamil *nikaṇṭu* with its high share of Sanskrit elements, but in the language and its literature, in other words, a multilingual scholar.

Case 4: Tamil-Sanskrit gloss not from the *nikaṇṭu*

This conclusion is confirmed by the fifth example:

AN 62.6 *maṛai* – *rahasyam* [ரஹஸ்யம்], < Skt. *rahasya-* “secret”

From the point of view of early Tamil poetry the word *maṛai* is not really polysemic: it is the well-attested root noun of a frequent verb *maṛaital*, “to be hidden”, and its meaning is “secret”. This presumbaly is the reason why as such it has not found entry into either the *Tivākaram* or the *Piṅkalam*.

However, with the advent of Northern religion it becomes the calque for the Sanskrit word *veda-*, referring to the holy scriptures beginning with the *Rig-veda*, and the *Piṅkalam* has a *sūtra* where *maṛai* is listed as one of the words

¹⁵ Yet another line of investigation would be the semantic development Sanskrit words undergo in the process of being borrowed.

for *vētam*.¹⁶ For the commentator in the early second millennium this may have been the far more natural and frequent application of the word *maṛai*, hence the need to bring forward a gloss. Since the *nikaṇṭu*-s did not oblige here, he had to find an equivalent for himself, and he chose the Sanskrit word *rahasya*-, spelt in Grantha, but with an *anusvāra* at the end instead of the Tamil final -m. Again this is something phonetically difficult to represent in Tamil – the *Tamil Lexicon* has an entry for *irakaciyam*, without, however, being able to provide a single quotation. In my view this confirms the conclusion that either the commentator was a bilingual scholar, or the 18th-century scribe of our manuscripts was, in which case one may wonder of course whether he added glosses of his own to the existing commentary. There is no direct evidence that he did, though, since all the glosses mentioned that were mentioned above are part of the main body of the commentary whereas usually additions seem to be added at the end of the commentary on a particular poem (for details see Wilden 2020).

Preliminary Conclusions

Early Tamil grammar shows an awareness of Sanskrit (*vaṭamoli*, “the Northern language”) as a provider of poetic words, but early lexicography, interested in explaining difficult words, at least overtly is now aware of such linguistic distinction: the first poetic dictionaries are of the Thesaurus type that put word fields into verse for easier memorisation. Their interest is not semantic but poetic (finding the right word to fit a metrical scheme).

Already the first of them, the *Tivākaram* (9th c. CE), appears to be highly Sanskritised. The most likely explanation for that is that the concept of the Thesaurus was taken over from the North, as the Tamil designation (*nikaṇṭu* < Skt. *nighaṇṭu*-) suggests. As the genre developed, the goal appears to have been a quest for more synonyms: accordingly the entry verses become longer, regardless of the language of origin for the single words.

The early commentaries on works from the Caṅkam corpus can be shown to have made use of these Thesauri; in some cases it is possible to locate the exact passage where a gloss came from. The anonymous commentator on the *Akanāṇṭuru* (ca. 12th c. CE), from whose work the present examples have been collected, often chooses a Sanskrit equivalent for his gloss on a Tamil word, also in cases where it would have been easy to find a Tamil counterpart.

16 Piṅk 7.301 *irukku, curuti, maṛai, eḷutāk kiḷavi, | ātinūl, cākai, āraṇam – vētam*.

At least the scribe of the single surviving manuscript (UVSL 297) appears to have made a difference between the two languages, because he often uses Grantha to represent Sanskrit letters, although by no means always. Epigraphic evidence seems to corroborate this practice. Moreover, his use of Sanskrit clearly goes beyond the *nikaṇṭu*, since in two of the cases discussed he chooses glosses that cannot be traced back to the lexicographical works.

However, it is difficult to decide for which period the evidence from the manuscript may be valid, the period of the commentator or the period of the scribe (perhaps the 18th century). It is perfectly possible that successive scribes have been faithfully copying not only the transmitted wording of the commentary, but also its spelling. Nevertheless, they may also have interfered with either wording or spelling or both, and a much broader analysis of the surviving evidence, that is, the manuscripts both for the classical commentaries and for the *nikaṇṭu*-s would be necessary in order to gain a clearer idea of the issue.

Very clear is, by contrast, that with the Tamil renaissance of the late 19th century, when the poetic texts were prepared for print and new commentaries emulating the medieval ones were written, Sanskrit began to be expurgated, first by getting rid of the script, then of the few permitted special letters, and finally by exchanging words. Our dictionaries belong to this phase. Currently there are efforts in Tamilnadu to rewrite the classics themselves into pure Tamil.

Bibliography

Primary

- Cēntaṅ Tivākaram*. Ed. by Tāṅṭavarāyamutaliyar, Manōṅmaṇivilāca Accuk-kūṭṭam, Ceṅṅnai 1917.
- Cēntaṅ Tivākaram Piṅkalam Cuṭāmaṇi*. Cānti Cātaṅā, Ceṅṅnai 2004.
- Neṅuntokai ākum Akanāṅūru mūlamum paḷaiya uraiyum* [old comm. for KV-90, modern comm. for 91–160]. Ed. by Vē. Rā. Rākavaiyankār, Kamparvilāsam Vē. Rājakōpālāryaṅ, Kampar Pustakālayam, Mayilāppūr, Ceṅṅnai, Śrīmuka v° (= 1933/34).
- Piṅkala Nikaṇṭu*. Ed. by Ti. Civaṅpillai and Ti. Ka. Cupparāyacetṭiyār, Intu Tiyalājikal Yantiracālai, Ceṅṅnai 1890.
- Tolkāppiyam, Eḷuttatikāram, Collatikāram, Poruḷatikāram with all the commentaries*. 14 vols. Edited by T.V. Gopal Iyer, Tamiḷmaṅ Patippakam, Ceṅṅnai 2003.

Secondary

- Chevillard, Jean-Luc 2010. “‘Rare words’ in classical Tamil literature: from the Uriyiyal to the Tivākaram”, in: *Acta Orientalia Academiae Scientiarum Hung.*, Volume 63, Number 3/September 2010, pp. 301–317.
- 2011. “*urappiyum eṭuttum kaṇaittum*: Contrastive Phonetics or How to Describe one Language with the Help of Another One”, in: *The Tamils: From the Past to the Present. Celebratory Volume in Honour of Professor Ālvāpiḷḷai Vēluppiḷḷai*. Ed. by Peter Schalk & Ruth van Nahl, Kumaran. Colombo/Ceṇṇai: Book House, pp. 105–117.
- 2016. “On a 1968 incarnation of the Piṅkalam, a traditional tamil thesaurus”, in: *Tamiḷ tanta paricu: The Collection of Articles in honor of Alexander M. Dubyanskiy, (édité par) O. Vecherina, N. Gordiychuk, T. Dubyanskaya*. Moscou: Russian State University for the Humanities. Collection *Orientalia et Classica*. Papers of the Institute of Oriental and Classical Studies, pp. 5–28.
- 2019a. “Tamil,” in: *The Cambridge World History of Lexicography*, in: John Considine (ed.): Cambridge & New York: Cambridge University Press, 80–83.
- 2019b. “The Development of the Poetical/Classical Tamil Lexicographical Tradition,” in: John Considine (ed.): *The Cambridge World History of Lexicography*. Cambridge & New York: Cambridge University Press, 148–158.
- Ciotti, Giovanni & Erin McCann (eds.) 2021. *Linguistic and Textual Aspects of Multilingualism in South India and Sri Lanka*, EFEO-IFP: collection indologie 147, NETamil series 8, Pondicherry.
- Dhamodharan, A. 1978. *Tamil Dictionaries. A Bibliography*. Beiträge Zur Südasiensforschung Südasiens-Institut Universität Heidelberg, Band 50, Wiesbaden: Franz Steiner Verlag.
- Dravidian Borrowings from Indo-Aryan*. Ed. by T. Burrow, M.B. Emeneau. Berkeley & Los Angeles: University of California Press 1962.
- Dravidian Etymological Dictionary*. Ed. by T. Burrow, M.B. Emeneau. 1st ed. London 1961, 2nd ed. Oxford 1984.
- Dravidian Etymological Dictionary*. Supplement. Ed. by T. Burrow, M.B. Emeneau. Oxford: Clarendon Press 1968.

- James, Gregory 2000. *Colporuḷ: A History of Tamil Dictionaries*. Ceṇṇai: Cre-A.
- Tamil Lexicon*, University of Madras, 7 vols., rep. 1982.
- Tamil Lexicon*, University of Madras, vol. I “a”, 2012.
- Tamiḷ ilakkiyaḷ pēraḷarāṭi*, [Tamil-Tamil dictionary of classical Tamil literature in 5 vol.], Cānti Cāṭaṇā Ceṇṇai 2001.
- Vaidyanathan, S. 1971. *Indo-Aryan Loanwords in Old Tamil*. Madras: Rajan Publishers.
- Wilden, Eva 2008. *Critical Edition and Annotated Translation of the Naṟṟiṇai + Glossary*. 3 volumes, EFEO/Tamilmann Patippakam, Critical Texts of Caṅkam Literature 1.1–1.3, Chennai.
- 2010. *Kuṟuntokai*. Critical Edition, Translation, Notes + Word Index. 3 vol., Critical Texts of Caṅkam Literature 2.1, 2.2, 2.3, Ceṇṇai: Tamiḷmaṇ Patippakam.
- 2018. *Akaṇāṇūru* + Old Comm. (1–90). *A Critical Edition and an Annotated Translation with Glossary of the Akaṇāṇūru; Volume 1–3 – Kaḷiṟṟiyāṇainirai*. Collection indologie 134.1–3 – NETamil series 1.1–3 – Critical Texts of Cankam Literature 4.1, 4.2, 4.3, Pondicherry: EFEO-IFP. (<<https://hal.archives-ouvertes.fr/NETAMIL/halshs-01892291>>, <<https://hal.archives-ouvertes.fr/NETAMIL/halshs-01892273>>, <<https://hal.archives-ouvertes.fr/NETAMIL/halshs-01892108>>, accessed 30.03.2022).
- 2020. “*Akaṇāṇūru paḷaiyavurai*: The Gradual Accumulation of a Commentary”, in: Suganya Anandakichenin & Victor D’Avella (eds.): *The Commentary Idioms of the Tamil Learned Traditions*, EFEO-IFP: collection indologie 141, NETamil series 5, Pondicherry, pp. 167–208.

