Metaphony, or height harmony, in written Tamil and Malayalam

Herman Tieken

The first thing anyone familiar with modern written Tamil as well as with classical Tamil will notice when he or she starts to learn spoken Tamil is that under certain circumstances the root vowels i and u are lowered to e and o. This change takes place before the derivational suffix -ai and when the radical vowels are followed by a single consonant only. Thus, for instance, nilam “earth” is pronounced nelam and kutai “parasol” is pronounced koṭai.

The same situation is found in written and spoken Malayalam. As a result of this development, spoken Tamil and Malayalam seem to go back to the situation which existed in Proto-South-Dravidian (PSDr). That is to say, in Proto-South-Dravidian Proto-Dravidian (PDr) /uCa(i) had become /eOCa(i). In classical Tamil, and consequently in written Tamil as well as written Malayalam, /eOCa(i) changed back into /uCa(i). As a result we find PDr it- > PSDr et-aţi “stumble”, Kannada edaalu, edapu, Tulu etepi but Tamil itaţi and Malayalam itaŋka and PDr pur- > PSDr pur-ay “layer”, Kannada pore, Telugu pora but Tamil purai. In modern spoken Tamil and Malayalam this /uCa(i) thus changed back into /eOCa(i) yet again, as a result of which the situation of Proto-South-Dravidian was, as it were, restored.

At first it was assumed (or claimed) that with /uCa(i) written Tamil and Malayalam represented the pre-Proto-South-Dravidian situation (For the earlier literature, see Krishnamurti 1958: 461, esp. note 7). Subsequently, however, Krishnamurti has convincingly shown that we are dealing with an

1 In Tamil and Dravidian studies the phenomenon discussed in this article is known as metaphony. The term height harmony has been borrowed from Germanic philology.
1. The evidence from literary texts and inscriptions

As indicated, in classical and modern written Tamil e and o do not occur before Ca(i). Beside the long vowels ā (ātai “dress”), ī (ikai “gift”), ū (ūkai “owl”), ē (ēval “command”) and ō (ōlai “palm-leaf”) in this context only i and u are found, as well as a (akam “interior”), though this is not relevant here. However, there are a few exceptions to this in classical Tamil, in particular in Cankam poetry (usually dated some time between the second century BC and the fourth century AD). Interestingly, most of the words concerned are found in modern written Tamil as well. The words in question may be divided into the following two categories. The first category consists of verbal nouns ending in –ai or –al, like koṭai “the act of giving” from koṭu “to give” and ceyal “the act of doing or making” from cey “to do or make”. Other words belonging to this category are etai “weight” (not found in Cankam poetry), kolai “the act of killing”, noṭai “price” and poṭai “forbearance”. A word of a slightly different make-up which might be included here is celavu “expenses” from cel “to go away”. Another word to be included here is the verb tolaɪ “to waste”. This verb has probably been based on the verbal noun tolaɪ from toli “old, ancient” as in the title of Tol-kāppiyam “the old poem”. The second category consists of the following words: peyar “name”, peyar “to go”, poṭai “female of animal”, koṭar “to bring”, toṭanku “to begin” and toṭar “to follow”.

Instances of e/oCa in Cankam Tamil and modern written Tamil

1) Verbal nouns
   - ceyal, celavu, (etai), koṭai, ceyal, kolai, noṭai, poṭai, tolaɪ etc.

2) A limited group of words
   - peyar, poṭai, koṭar, toṭanku, toṭar.

The survival of these two groups of exceptions in modern written Tamil has not been explicitly discussed anywhere so far. Apparently, it is something that is taken for granted. Modern written Tamil is indeed an artificial language, of which classical Tamil was one of the main sources. However, if in the strictly literary tradition, for instance, toṭanku “to begin” has remained unchanged from the period of Cankam poetry down to modern written Tamil, this is not the case in inscriptions. In inscriptions, as we shall see, toṭanku is occasionally found as well.

The earliest Tamil inscriptions are the so-called Tamil-Brahmi inscriptions (Mahadevan 2003, Ticken forthcoming). These are short texts, mostly
independent sound change in which \( e/oCa(i) \) from \( i/uCa(i) \) become \( i/uCa(i) \) again (Krishnamurti 1958). According to Bright (1966: 312) it is not impossible that the mid vowels in the modern spoken languages are actually survivals from the Proto-South-Dravidian period. However, Bright restricts the phenomenon to only some spoken varieties of Tamil, while it is actually widespread in Tamilnadu (Tamil) as well as in Kerala (Malayalam). Generally, the state of affairs in the modern spoken languages is presented as the outcome of a series of discrete, consecutive steps (Krishnamurti 2003: 101-105), which may be presented in the following schematic way:

```
Proto-Dravidian
  i/uCa(i)
      ↓
Proto-South-Dravidian
  e/oCa(i)
  ↓
Classical/Written Tamil
  and Malayalam
  i/uCa(i)
      ↓
Modern Spoken Tamil
  and Malayalam
  e/oCa(i)
```

What calls for particular comment in this going back and forth from \( e/o \) to \( i/u \) to \( e/o \) again is the fact that classical and written Tamil and Malayalam separate, like a wedge, modern spoken Tamil and Malayalam from Proto-South-Dravidian. In this connection it should also be noted that as far as written Tamil is concerned only part of the available material has been taken into account so far. Thus, the evidence provided by inscriptions has not been considered at all. In fact, the picture emerging from the inscriptions appears to be less regular than that from the literary texts which have been used so far. I will therefore take a closer look at the situation in the written language, classical as well as modern. In so doing I will focus, in the first place on Tamil, which was the first of the two languages concerned to develop a written variety. Only at the end will I come back to Malayalam. In what follows I intend to show that with the forms with the raised vowels we might well have to do with learned forms introduced in the literary language, which are supposed to correct a presumably careless pronunciation in the spoken language.

1. The evidence

As indicated, in certain cases before \( Ca(i) \) becomes \( k \) (kākai “owl”), \( ē \) becomes \( o \), only \( i \) and \( u \) are relevant here. However, in particular in Classical/South Dravidian Tamil and modern spoken Tamil and Malayalam, the change \( e/oCa(i) \) to \( i/uCa(i) \) may be divided into two stages consisting of verbal nouns ending in -\( i \) “to give” and ceyal “to come”. Other words belong to “forbearance”. A few other words belong to “forbearance” in the same meaning. Words like ceval, ceval, ceyal and ceval.

Instances of ceval, ceval, ceval:
1) Verbal
   ceval, ceval
2) A limiting
   ceval, ceval

The survival of ceval, ceval, ceval in the literary language has not only something to do with classical articular languages, but also with the fact that these languages have remained written. However, in modern written Tamil, ceval, ceval, ceval, ceval, ceval is not used. The earliest inscriptions (Ma...
found outside or inside natural caves in Tamilnadu which served as retreats for monks. A majority of the inscriptions mention simply the monks’ patron who was responsible for making the cave fit for habitation during the rainy season. Often this involved cutting a rim in the overhanging rock which served as a dripping ledge, thus preventing the water from flowing down the rock into the cave. In addition, beds were cut in the floor of some of the caves for the monks to rest on. The Tamil-Brahmi inscriptions are dated between the second century BC and the fourth century AD and are thus approximately contemporaneous with Caṅkam poetry. The linguistic material of the inscriptions is restricted, consisting mainly of names and titles. Interestingly, as far as i/uCa- and e/oCa- is concerned, there are traces of what may be interpreted as hesitation on the part of the scribes or engravers about the orthography of the words in question. Thus, while in some inscriptions one may find ila- “young” (in iḻankāṅṅkō “name of a prince of the Irumporai dynasty”, iḻankāyipan “name of a Jain monk”, iḻankō “prince, heir apparent”, iḻañcaṭkaṭ “name of the son of Caṅkō” and iḷañāṭ “name of a donor”) and iḷayar “young person, title of a member of a royal clan” (Mahadevan 2003: 645, 662), in others we find ela (in elacanṭan, elamanṭan and elā-a aṭan). The equation of elai in elai-y-ūr with itai “middle or the herdsman caste” is uncertain. Mahadevan 2003: 663). Similarly we find pokalβ (“fame” beside pukalβ (Mahadevan 2003: 668).

The vocabulary of the inscriptions includes one of the words of the above list of exceptions, namely poṭai in the royal title perumpoṭai. However, beside perumpoṭai in the Tamil-Brahmi inscriptions from Pugalur and on some copper coins of the same dynasty, perumpuṭai is found as well (Mahadevan 2003: 669).

The picture emerging from inscriptions dating from the sixth to eighth centuries, mainly inscriptions of the Pallava dynasty, is similar. Thus, beside peyar “name” occasionally piyar is found, beside cēlava “expense” cilavu, beside kōpaṟntu “having brought” kūpaṟntu and beside kōλai “killing”, in kōlaikkalpu “murderous elephant”, kulai (Velu Pillai 1976: 3, 6). In inscriptions from the ninth century piyar, inakkō (“for me”), tuṭanki(y), kūpa(r)ntu and kūpaṟntu are found alongside the forms with the mid vowels (Velu Pillai 1976: 140, 143); for the tenth century piyar, cilavu, inakkō, kūpa(r)nten, tuṭankiṇa, tuṭara are found alongside peyar etc. (Velu Pillai 1976: 358, 361-362); for the eleventh century piyar, cilavu, tuṭanki(ya), tuṭar and tuṭai (“thigh”) alongside peyar etc. (Velu Pillai 1976: 594, 599-601); and for the twelfth century, beside most of the words already quoted above, kulaitara (kōlaitara “to disperse”), tulaiyāṭa (tōlaiyāṭa “which is not coming to an end”) and tuṭa 857). A study of 1 yielded the follow

Shanmugam 1970:

Instances of written Tamil have

6th–8th century piyar, cēlava
9th century piyar, cēlava
10th century piyar, cēlava
11th century piyar, cēlava
12th century piyar, cēlava
1150 to 1200 cēlava, in

It is to be

Metaphor

The inscription with cēlava and from which su curious, as the it wrote literary to take into accounts available in the instance, 1st M:
served as retreats the monks’ patron
during the rainy
ning rock which
owing down the
or some of the
ptions are dated
AD and are thus
able, while in
òl “name of a
of a Vaina monk”,
ion of Catíkam” and
le of a member of a
we find ęla (in
ai in ęla-y-űr with
adevan 2003: 663).
defined by ędevan 2003: 668).
dards of the above list
. However, beside
ralur and on some
as well (Mahadevan
n the sixth to eighth
ticular. Thus, beside
“expense” ćilavu,
: kolai “killing”; in
1976: 3, 6). In
or me”), ćutanki(y),
ith the mid vowels
, ćilavu, ćinaku,
r etc. (Velu Pillai
u, ćutankitya), ćutar
976: 594, 599-601);
ready quoted above,
which is not coming
to an end") and ćutáí (ćotai “continuity") (Velu Pillai 1976: 850, 852, 856-
857). A study of the Tamil inscriptions from the period 1250 to 1350 has
yielded the following instances: ćilavu/ćelavu, ćinaku/ćinaku, ćia/ćia,
ćukay/ćokay, ćutanku/ćotanku and ćutar/ćotar (Agesthialingom and

Instances of ćuCa in the inscriptions where Catíkam and modern
written Tamil have ćoCa

6th to 8th century (Velu Pillai 1976: 3, 6)

ćiyar, ćilavu, ćuparntu, ćukalakkalų

9th century (Velu Pillai 1976: 140, 143)

ćiyar, ćelavu, ćinaku, ćupa(r)ntu, ćupa(r)nta, ćutanki(y)

10th century (Velu Pillai 1976: 358, 361-362)

ćiyar, ćilavu, ćinaku, ćuparntu, ćutankina, ćutara

11th century (Velu Pillai 1976: 594, 599-601)

ćiyar, ćilavu, ćutankiya, ćutanki, ćutar, ćutarvōm, ćutai

12th century (Velu Pillai 1976: 850, 852, 856-857)

ćiyar, ćilavu, ćinaku, ćiat, ćupa(r)ntu, ćutara, ćutarntu, ćutanki, ćutai,
ćukalaitara, ćukalaitā

1250 to 1350 (Agesthialingom and Shanmugam 1970: 25-26)

ćilavu, ćinaku, ćiat, ćukay, ćutanku, ćutar

It is to be noted that in inscriptions instances of the reverse development,
namely of ćo for ću, are also found. Thus, beside ęla and pokal(l) in
the Tamil-Brahmi inscriptions (for which, see above), ēptatu (for ēptatu) has
been attested, as well as ćutan (ćutan) and ponaI (ponaI) (Velu Pillai 1976:
140, 143). The latter three forms are found in inscriptions from the sixth to
eighth centuries.

The inscriptions, in which ćilavu and ćutanku are found side by side
with ćelavu and ćotanku, differ from this respect from the literary tradition,
from which such variation seems to be almost completely absent. This is
curious, as the inscriptions were most likely written by the same scribes who
wrote literary texts (Tieken 2006). In this connection, however, we should
take into account the nature of the various sources. Thus, an inscription is
available in the form in which it was written down by scribe X on, for
instance, 1st May, 876. By contrast, literary texts have come down to us
through an endless series of copies of copies. Each new copy provided ample occasion for corrections or for adaptation to what was considered the norm within the literary tradition. At the same time it should be noted that the situation in the literary tradition might not have been as uniform as all that. For instance, in later texts, like Kampan’s Irāmōvatāram, one may come across tuṅkku beside toṅkku. Unfortunately, the situation in medieval Tamil texts has not as yet been studied in this respect. The same applies to inscriptions which date from after the middle of the fourteenth century.

2. Reconsidering the sequence of developments

Krishnamurti (1958: 465-466) considered the so-called exceptions found in Cankam poetry as fossils from Proto-South Dravidian. However, this explanation does not seem to account for the occurrence in the same poetry of the verb tuṭkkku “to tie, bind, bring together”. This verb most likely goes back, through tuṭkku, to the verb toṭar, found in Cankam poetry as well. In addition, Krishnamurti’s explanation is based on the dating of Cankam poetry between the second century BC and the fourth century AD. However, in recent publications, I have questioned this early date and tried to show that Cankam poetry most probably originated only from the eighth or ninth century AD (Tieken 2001, 2003, 2004 and forthc.1). Old Tamil Cankam poetry is generally dated in the period described in the poems, which is an early bardic society. I have tried to show that Cankam poetry should be dated not in the period it describes but in a period after that, the poems evoking a society which at the time already belonged to the past. My dating of the poems in the eighth or ninth century is based on the use of the Tamil language in the poems for the depiction of, among other things, local history as against all-Indian mythological history. A similar use of Tamil is also met with in the inscriptions of the Pandyas of the eighth or ninth century, and only in the inscriptions of that particular dynasty. Thus, these inscriptions contain two genealogical accounts of the Pandyas, one in Sanskrit tracing them back to a mythological past, the other in Tamil which by contrast provides concrete, local history. If Cankam poetry does indeed belong only to the eighth or ninth century, this would mean that after a period of variation in the inscriptions between, for instance, the spellings peyar and peyyar and toṭar and tuṭar in Cankam poetry, the matter was settled in favour of the former spellings. Subsequently such “exceptions” as peyar and toṭar have come to be considered as typical of the literary idiom.

In section role in the optio Cankam poetry material descr: sources the ten before Ca(t). N. variant with ti Tamil-Brahmi inscriptions (V, would seem that in a fairly stabil high (e.g. utavi side. In any case itu in ituCa(y) during which earlier, to the r of information inscriptions of tempting to cor the situation hi which we find

3. An attempt

Rather than wi developments domain each o be noted that i dialects as th concerned hac unlikely as di different regi forms with th sense either. beginning of Dravidian) or present-day si written langu spoken langu
In section 3, I will come back to the ideas which might have played a role in the option of, for instance, *kotai* “the act of giving” instead of *kutai* in Cankam poetry. First I would like to discuss another matter. From the material described above (section 1) it becomes evident that in the written sources the tendency was clearly and strongly towards the raising of *e/o* before *Ca(i)*. Nevertheless, over a period of more than a thousand years, the variants with the mid vowels kept on appearing, such as *pokal(j)* in the Tamil-Brahmi inscriptions, *etattu, ota* and *ponal* in eighth-century inscriptions (Velu Pillai 1976: 138-139) (and *totar* in Cankam poetry). It would seem that in the latter cases we are dealing with isolated interferences in a fairly stable situation. All this time the two forms, namely those with the high (e.g. *utavi* “help”) and mid vowels (*otavi*), seem to have existed side by side. In any case, the scribes appear to have been conscious of the fact that *i/u* in *i/uCa(i)* could represent *e/o* and *vice versa*. The length of the period during which this situation applied, from the beginning of our era if not earlier, to the middle of the fourteenth century, is striking. Despite the lack of information on the interchange of *i/uCa(i)* and *eOCa(i)* in texts and inscriptions of the period from the fourteenth century to the present time, it is tempting to conclude that from the beginning of the written tradition onwards the situation has never been very different from that of modern Tamil, in which we find *utavi* in the written language and *otavi* in the spoken variety.

3. An attempt at a reconstruction of the situation

Rather than with two consecutive situations we seem to be dealing with two developments existing side by side. The question then arises in which domain each of the two variants has to be situated. In this connection it is to be noted that it is not very likely that we are dealing with forms of different dialects as this would assume that the relation between the dialects concerned had remained the same for more than a millennium. This is unlikely as during this period the political situation changed repeatedly, different regions becoming prominent with each change. To ascribe the forms with the vowels *i/u* to the spoken language would not really make sense either. In any case, the forms with *i/u* did not belong there in the beginning of the development (during the formation of Proto-South Dravidian) or at the end (modern spoken Tamil). In fact, in the light of the present-day situation it is tempting to take the raised forms as typical of the written language and the ones with the low vowels as interferences from the spoken language.
The forms with the high vowels may have been introduced as learned formations in the fabrication of which etymology or derivation could have played a role. Thus, *i*tam “place” rather than *etam as it was derived from *i*ti “to place” and, vice versa, *e*tai “weight” and not *i*ti as it was derived from *e*ti “to take up”. In this way, forms like *etam could have come to be considered as the careless pronunciations of illiterate speakers unfamiliar with grammatical concepts which, in such forms, were not considered fit to be used in the written language.

As to the supposed role of etymology or derivation in the fabrication of the written forms such as *itam, it should be noted that the beginning of written Tamil was accompanied by sophisticated linguistic speculation. This becomes clear from the orthography, which takes into account the rule in Tamil known as the “convertibility of surds and sonants” (Burrow 1939, Krishnamurti 2003: 144-150). This rule implies that for the Tamil plosives the distinction between voiced (sonants) and voiceless (surds) is not phonemic. The voiced and voiceless plosives are in complementary distribution. Thus, /k/ occurs initially (kātā “love”) and doubled (akkā “elder sister, elderly female”); /g/ occurs only after a nasal (tangā “younger sister”) and between vowels, when it was actually pronounced /k/. Typically, when in the second century BC the North-Indian Brāhmi alphabet was adopted for Tamil, the letters for the voiced plosives (e.g.  in gūthu “householder”) were ignored. Thus, (tanigai) was written as tankai and (a xam) “interior” as akam. While, only recently, this situation in the Tamil-Brāhmi inscriptions was taken as an indication that there were no voiced plosives in Old Tamil (Mahadevan 2003: 247-251), comparative evidence clearly suggests that there were. If so, the orthographic practice of Tamil testifies to an advanced knowledge of the grammar of Tamil on the part of the first scribes who developed the written form of that language. We may therefore assume that these scribes were also familiar with a process like derivation and knew how this worked in the case of Tamil.

While in the case of words like koτai “the act of giving” instead of koτai it might not be hard to imagine the role of derivation, it is admittedly less clear in the other category of words, namely petai, peyar, totar, totanku and kopar. The influence of derivation is in retrospect difficult to establish without a certain degree of arbitrariness. Thus, in the case of petai “female of animal” (modern Tamil petṭai) its relation with peṇ “female, daughter” might have played a role, which would have been lost with the spelling pitṭai.

The word peyar “narr maintain a connecti meaning, compare S individual”) and the move”. However, in not have *veyar inst that both words are with, begin” might supposed derivation

4. Malayalam

The discussion has a situation that is fou e/oCa(i) in the spok Malayalam are offis two languages may. Written Malayalam. Vāzappaḷḷi inscripti approximately the agreement of wriutt adoption of i/uCa developed writing east-coast culture α in inscriptions and inscription by the v the endorsement of only after they has custom of the land is Tamilnādu (Tiel Akapporuḷ an acc Composed by Na handed down from town in Kerala (Tiel is the adoption by had proved her eff
The word *peyar “name” may have been retained in the written language to maintain a connection with *pēr “person” (for the development of the meaning, compare Sanskrit nāma “name” in nāmarūpa “name and form, individual”) and the verb peyar “to go” to maintain a connection with *pēr “to move”. However, in that case it is not clear why next to vēr “sweat” we do not have *veyar instead of viyar. In the case of toṭar and toṭanku the idea that both words are somehow related to toṭu “to touch, come into contact with, begin” might also have played a role and in the case of koŋpar the supposed derivation of the word from the combination koŋtu + var.

4. Malayalam

The discussion has so far been restricted to Tamil. As noted above, the same situation that is found in Tamil, with i/uCa(i) in the written language and e/oCa(i) in the spoken language, is found in Malayalam as well. Tamil and Malayalam are offshoots of Proto-Tamil-Malayalam and the division of the two languages may go back to prehistoric times (Govindankutty 1972). Written Malayalam, however, goes back only to the ninth century AD (the Vāzappal̄i inscription). The first literary work, the Rāmacaritaṁ, dates from approximately the twelfth century (Krishnamurti 2003: 22). As to the agreement of written classical Malayalam with Tamil on the point of the adoption of i/uCa(i) instead of e/oCa(i) it should be noted that Kerala developed writing and a written literature only under the influence of the east-coast culture of Tamilnadu. This process of acculturation is referred to in inscriptions and is the topic of literary legend and literary texts. Thus, an inscription by the west-coast Ay king Karunatakkkan (ninth century) refers to the endowment of a college for Vedic students. The students were admitted only after they had passed an entrance test in subjects like the “laws and custom of the land of the three kingdoms”. The “land of the three kingdoms” is Tamilnadu (Tieken 2001: 134), in a commentary on the poetical treatise Akapporuḻ an account is given of the transmission of the commentary. Composed by Nakkirā in Madurai in Tamilnadu the commentary was handed down from teacher to pupil, to arrive finally in Muciri, a harbour town in Kerala (Tieken 2001: 202); and the topic of the epic Cilappatikāram is the adoption by the west-coast rulers of the cult of Pattini, a goddess who had proved her efficacy in Tamilnadu (Tieken 2001: 202-208).
5. Conclusion

The above is an attempt to account for the curious state of affairs in written Tamil and Malayalam. In the written varieties the development of /ucCa(i) to /oCa(i), which took place in Proto-South-Dravidian and which seems to have survived in the modern spoken languages has been, as it were, reversed. So far, the various forms have been placed one after the other as a result of subsequent, historical developments: Proto-Dravidian /ucCa(i) > Proto-South-Dravidian /oCa(i) > (classical) written Tamil/Malayalam /ucCa(i) > Modern Spoken Tamil/Malayalam /oCa(i). However, this explanation does not properly account for the existence during almost the entire period of the two forms side by side, /ucCa(i) and /oCa(i). By way of alternative it has been argued that the forms with the high vowels present an artificial spelling introduced into the written language. Etymological speculation may have played a role in their fabrication in combination with notions of correct, careful pronunciation. In this connection it should be noted that in written Tamil (and Malayalam) had to vie with Sanskrit, a learned and in many ways artificial language, which had become completely detached from the spoken languages and was carefully kept apart from them. Sanskrit was the language par excellence of the textual tradition in India. It was only at the end of the first millennium that the so-called vernaculars came to be used as written languages as well. For this they had to be fitted out with grammars and ancient literary traditions, just like Sanskrit. A case in point of such an ancient literary tradition is Cankam poetry, which evokes a society in which wandering bards sang the kings' praises in Tamil. Subsequently these new literary languages were preserved just as carefully as Sanskrit, and kept apart from the spoken languages. This may account for the conservative appearance of modern written Tamil with which, leaving aside the problems of technical vocabulary, it is indeed possible to read eighth-century inscriptions.

References


Mahadevan Irvatham, 2003, Early Tamil Epigraphy From the Earliest Times to the Sixth Century A.D., Chennai and Cambridge, Massachusetts.


