Orthographic knowledge, reading and spelling development in Tamil: the first three years

Sonali Nag and Balambigai Narayanan

Abstract

Tamil (tamil) is a South Dravidian language with a contained orthography in contrast to the other more extensive akshara orthographies of South Asia. In this chapter we briefly describe the Tamil orthography, introducing the rich vocabulary available in the language to describe the symbol set. Similar descriptions are given for Tamil phonology and morphology. We also focus on Tamil teaching before presenting research findings on early literacy development. Our survey data suggest that important milestones in the journey to mastery of word reading and spelling include learning about the native and *grantha* consonants and unconnected diacritic markers, especially those with ambiguous visual elements such as the *thunaikaal*. A discussion on writing routines, diglossia and children's spontaneous writing is also included.

Key words: Tamil, reading, spelling, orthographic knowledge, diglossia, writing routines, contained orthography, literacy instruction

Author details

Sonali Nag,

Associate Professor of Education and the Developing Child, Department of Education, University of Oxford, 15 Norham Gardens, OX2 6PY, UK

and

Honorary Associate Director, The Promise Foundation, 231, Cozy Home Layout, Sarjapura, Bangalore, India

sonalinag@t-p-f.org

sonali.nag@education.ox.ac.uk

Balambigai Narayanan,

Lecturer, Department of Speech, Hearing & Communication, National Institute for Empowerment of Persons with Multiple Disabilities(Divangjan) (NIEPMD), East Coast Road, Muttukadu, Chennai, India

balambigai.n@gmail.com

Acknowledgement:

This book chapter is funded by a seed grant on early literacy research to Balambigai Narayanan from The Promise Foundation, and a research grant to Sonali Nag from Brasenose College, University of Oxford.

Introduction

Tamil (tamiḻ) belongs to the South Dravidian language family. Tamil has the largest number of speakers compared to the other literary South Dravidian languages of Telugu, Malayalam, and Kannada. Tamil is spoken in parts of India, Malaysia, Singapore, Sri Lanka, the Fiji Islands, and an extensive diaspora elsewhere. The southern state of Tamil Nadu in India has the largest concentration of Tamil speakers which with a population of 96 million (Census India, 2015) is larger than most countries in the world.

Tamil has an uninterrupted literary history of more than two thousand years with some of the earliest cave and potshard inscriptions dated to the second century BCE. The history of Tamil is typically seen as spanning works belonging to Old Tamil, Middle Tamil and Modern Tamil. The earliest treatise on Tamil grammar, orthography and theoretical linguistics is entitled *Tolkaappiyam*. While the date and authorship of this ancient text is somewhat disputed, one school of thought is that *Tolkaappiyam* is perhaps a series of anthologies compiled over several centuries.

The following sections will cover descriptions of Tamil orthography, phonology, and morphology interspersed with our preliminary research findings in the area of early reading and spelling development. We also focus on Tamil teaching practices and describe in particular detail the first year of literacy instruction. The formal and written form of Tamil or the 'high' literary Tamil is called centamiḻ ('pure Tamil') whereas the vernacular, 'low' variety is called koṭuntamiḻ (the harsh or 'crooked Tamil'). As with other languages imbued with diglossia—Bengali and Hindi for example—very little information is available about the impact of diglossia on Tamil language and literacy acquisition. This chapter will touch briefly on issues of Tamil diglossia and literacy acquisition and flag this area as a gap that needs systematic research attention.

Tamil orthography

The Tamil orthography is an alphasyllabary belonging to the Brahmi group of scripts. Tamil-Brahmi, or Damili, predates the Prakrith-Brahmi scripts and includes unique symbols to capture the sounds of the Tamil language (Rajan &Yatheeskumar, 2013). Similar to other Brahmi-derived orthographies, the symbol units in Tamil are written in symbol blocks. The *akshara* is minimally a vowel (/V/) or consonant (/C/), and a consonant-vowel combination.

The mid-low vowel /ʌ/[1] is found inherent in the consonant, and for all other vowels a diacritic marks the vowel in a CV akshara. Unlike other Brahmi-based systems, conjoint symbols to represent consonant clusters are rare in Tamil. The Tamil orthography is considered mostly transparent although both historic and current influences have introduced several instances of opacity in the system.

Diachronic and synchronic features

The ancient script of Tamil Brahmi shares features with Sinhala Brahmi and predates Asokan Brahmi of North India. The features of Tamil Brahmi seem to have been best suited for the phonological characteristics of Old Tamil. Over time the symbol pairs [ɳ] and [n], [ɾ] and [r], and [ɻ] and [l] found in formal Tamil and in spelling have merged in spoken, informal Tamil. Another more recent influence to the orthography is the changing of the visually complex symbol shapes, perhaps led by two forces: One is typological, linked to the advent of the printing press but crystallised as font decisions for computer technology, and the other of orthography reform led, most notably, by the Government of Tamil Nadu. Figure 1 gives examples of these shifts. A curious possibility we will consider towards the end of this chapter is whether these trends to simplify and linearise the symbol set have inadvertently increased the learning demand for the young learner.

Figure 1: Recent shifts in the form of select Tamil akshara[1]

Note: [1]Examples of akshara from Old Tamil are in the first line. In the second line are the same akshara written using changes initiated by the orthography reform (contemporary Tamil). The last line gives the sound value of each akshara.

[1] We represent Tamil phonemes and akshara sound values using the alphabet of the International Phonetic Association (IPA), 1999.

The symbol set and a vocabulary to describe it

Tamil orthography (tamil̲ariccuvaṭi[2]) consists of ten vowels and two diphthongs (together called uyirel̲uttu), and eighteen consonants (meyyel̲uttu). The vowels can be further divided into the five long and short vowels (the neṭil and kur̲il). Consonants are divided into hard plosives (vallin̲am) which includes /k, s, ʈ, t̪, p, r/, the soft nasals (/n, n̪, ŋ, ɲ, ɳ, m/, mellin̲am), and medium consonants (iṭaiyin̲am) which include /j, ɾ, ʋ, l, ɭ, ɻ/. The symbol ஃ (called āyutael̲uttu) is mainly used as a pre-fix to native symbols to accommodate sounds in loan words (ஃ + ப̇ = f). In addition to this are the *grantha* symbols (e.g., /ʤ, s, ʃ, h/) to accommodate the phonology of Sanskrit which was a dominant source for loan words into Old Tamil. As is evident from the above description, Tamil has a rich vocabulary for sets of symbols in the orthography. Similarly, labels are available for intra-symbol elements. Examples of labels for the Tamil symbol unit include puḷḷi (◌̇, 'dot'), tun̲aikkāl (◌ா 'extra leg'); ottaikkompu (ொ 'one horn'); reṭṭaikkompu (ோ, 'two horns'); reṭṭaicul̲i 'n' (ன, '2 loops n') and mūn̲r̲ucul̲i 'n', (ண, '3 loops n'). An important point to note for emergent orthographic knowledge is that these labels for visual elements are generously used in early instruction.

Nag (2007, 2017) describes the world's orthographies by their orthographic breadth. Within this formulation, orthographies may be placed along a continuum that is contained to extensive in symbol repertoire: Cyrillic, Hebrew, Arabic and the Roman-derived alphabets with small symbol sets fall on the contained end of the continuum, while Chinese, with thousands of symbols, falls in the extensive end. The akshara orthographies fall somewhere in between and amongst them some have a smaller symbol set. Tamil may be considered a *relatively more contained orthography* when compared to the other more extensive akshara orthographies (usually symbol sets of 400 to 700). Ten vowel and two diphthong symbols combine with eighteen consonants to form 216 CV combinations. Added to this is the limited set of *grantha* akshara and the symbol ஃ described above. As mentioned earlier, conjoint consonant clusters are rare in Tamil, and the exceptions are ஸ்ரீ and க்ஷ (/ʃɪiː/ and /kʂʌ/, respectively). For other instances of consonant clusters, found mainly in loan words, the first

[2] We use the Tamil Lexicon Notation (1939) collated by the University of Madras to represent Tamil linguistic terms. See Appendix for the linguistic terms written in Tamil along with their IPA transcription and a simple definition.

consonant gives up the inherent vowel by taking on the schwa suppressing marker (puḷḷi, ◌்), making the number of completely new looking akshara beyond the basic set less common in Tamil. The linearised arrangement of the CCV akshara block has a second implication: Tamil is unlike neighbouring Kannada and Telugu which have a rich non-linear arrangement with the first and second consonants joined-up or ligatured to each other.

Surface features of vowel diacritics

The description in the research literature on symbol blocks within the akshara writing system has been dominated to date by those akshara orthographies that have a preference for joined-up or ligatured markers (e.g., Bengali, Hindi, Gujarati, Kannada and Marathi). These forms of representation are arguably later inventions of Brahmi. The surface architecture of earlier forms of Brahmi—the Tamil Brahmi and Sinhala Brahmi from fifth and sixth century BCE for example—had many more disconnected elements and a less expansive use of the visuospatial space. Modern Tamil has retained this preference for unconnected elements and this defining surface feature is seen most prominently in the vowel diacritic system of the orthography (see Figure 1).

The primary form of vowels is of full size and written in-line (e.g., உ, /ʊ/; ஒ, /o/). The secondary form of the same vowels is a diacritic and is either a fraction of the primary form (இ - ◌ி, /i/) or a visually distinct unit (ஆ - ◌ா, /aː/). Vowel diacritics of /aː, e, eː, o, oː/ and diphthongs /ʌj/ and /ʌʊ/ sit in-line and without a connected point with the base consonant (e.g., பெ, பா and பொ; /pe/, /paː/, and /po/, respectively) while the vowel markers /pu/ and /puː/ are off-the-line but ligatured to the base (பு and பூ, respectively). Some vowel diacritics have one and others have two unconnected elements, such as in the examples above with /e/, /aː/ and /o/. The learning task here is therefore quite different from the learning task in other akshara systems where diacritics are ligatured or connected to a point in a base consonant. In Tamil, novice readers should learn to gather up the right number of unconnected elements and not treat each as separate akshara.

The location and shape of a vowel marker within the akshara symbol block is entirely predictable except for diacritic markers for /u/ and /uː/. The diacritical marker for vowel /u/ always begins from the 3rd quadrant but the shape depends on the base consonant (e.g., கு, சு, and னு; /ku/, /su/, and /nu/, respectively). More variable in form is the vowel /uː/ (e.g.,

சூ and நூ; /suː/ and /n̪uː/). Taken together, the vowel diacritic system in the Tamil orthography is mostly predictable in location within a symbol unit.

Opaque features

Tamil orthography should be considered partially opaque for several reasons. At the level of the vowel, the few instances where ligaturing rules are needed are not consistent (see examples for the vowel diacritic u and u: above). Secondly, some vowels are visually confusable with some consonants (diacritic vowel form: ை and ன, /ʌj/ - /nʌ/; primary vowel form ஐ and ஜ, /ʌj/ - /ʤʌ/). Third, two vowel diacritics share visual elements with <aː>, <e>, and <eː> (ா, ெ, and ே); these are <o> (ொ) and <oː> (ோ). Added to this, some unconnected elements become more complex because when read as a concatenating whole they signify one sound value but if read in isolation each carries a different sound value (e.g., கௌ is /kʌʊ/, but க seperate from ௌ are /ke/ and /l̪ /). Readers probably use the context to decide which unconnected elements must be gathered up to form the akshara unit. Finally, even though the Tamil orthography is less non-linear in visuo-spatial arrangement than many other akshara systems, the non-linearity that does exist is because some vowel diacritics occur before, above and below the consonant, or on both the sides of the consonant (e.g., பெ, பி, பு, and பொ; /p/ + /e/, /i/, /u/, and /o/). Diacritics written off-the-line include the extended line below the consonant /u/ in /pu/ (பு), /ju/ (யு), and /ʋu/ (வு), the all-encompassing curl below the consonant /uː/ in /puː/ (பூ), /juː/ (யூ), and /ʋuː/ (வூ), and the minor curl above or in-line for /iː/ in /kiː/ (கீ), /siː/ (சீ), /liː/ (லீ) and /ɻiː/ (ழீ). These features of the orthography could arguably make akshara processing more demanding.

Drop in transparency at the level of the consonant is because the Tamil orthography does not differentiate between voiced and voiceless stops. The symbol க represents /kʌ/, /gʌ/ or /hʌ/ and the symbol ச represents /sʌ/, /t͡ʃʌ/ or /ʃʌ/; in other words, the affricate and fricative are represented by the same symbol. A novice reader must learn to use context to decode these akshara. Going beyond the single symbol level, a further point of ambiguity is with phonological changes in sound values within words. For example, when the consonant /r/ (ற்) geminates it is pronounced as /ttr/ as in /ʋettri/ (வெற்றி) but when it follows a nasal consonant it becomes a /dr/ as in /sendraːn/ (சென்றான்). And finally, spellings within

words may be opaque due to diglossia. For example, the written form of /kaːleːlʌ/ (‘in the morning’) is <kaːlʌjjil> (காலையில்), and <ɾeɳɖu> represents the spoken word /iɾʌɳɖu/ (‘two’, இரண்டு) (more details related to diglossia are given in the next section).

In summary, Tamil’s surface organization is more linearised than most other akshara orthographies. There is reduced stacking of markers compared to orthographies like Kannada, and no subscripts. The defining feature of the akshara block in Tamil is markers mainly placed to the right and left of the base, and use of the upper and lower quadrant of the visuo-spatial space.

Tamil phonology

A description of the phonological system of Tamil must begin with a note on whether the description is about phonology as transcribed in written Tamil or phonology as found in the multiple spoken varieties of the language. This distinction is important because Tamil is a strongly diglossic language and apparent differences across scholars in their account of the phonology of the language are primarily because of the form described; and till recently, the focus was mainly on written Tamil. We closely follow the diglossia by first describing Tamil phonology as captured in the orthography followed by accounts of the spoken varieties. Our description is brief and hence selective; for a more complete description see Annamalai and Steever, 1998, Christdas, 1988, and Ramasamy, 2010.

Written Tamil has eighteen consonants which include five plosives (/k, s, ʈ, t̪, p/), six nasals (/n, n̪, ŋ, ɲ, ɳ, m/), two lateral approximants (/l, ɭ/), two trills/taps (/r, ɾ/) and three central approximants (/ɻ, j, ʋ /) (c.f., Yang, 2011). In addition, written Tamil includes five consonants that can be traced back to early borrowings from Sanskrit (/ʤ, ʂ, s, h, kʂ/). Written Tamil has four front vowels (/i, iː, e, eː/), two central vowels (/ʌ, aː/), four back vowels (/u, uː, o, oː/) and the two diphthongs, /ʌj/ and /ʌʋ/. All vowels in Tamil can occur in all positions except the diphthong /ʌʋ/ which does not occur in the word-final position. Among the consonants, /k, t̪, n̪, p, m, s, ʋ, ɲ, j, ɾ, l / can occur in the word-initial position; /n, ɳ, m, j, ɾ, l, ɭ , ɻ / can be found in the word-final position and all eighteen consonants can appear in the word-medial position. Alveolar and the retroflex sounds do not occur in the word-initial position (Christdas, 1988) although Annamalai and Steever (1998) show occurrence of initial retroflexes in what they describe as native onomatopoeia. Consonant sequences are found in the word-medial position and these are typically nasals with stops /ŋg , ɲ$\widehat{ʤ}$, ɳɖ, nr, n̪t̪, mp/ in

words such as <maːŋgaːj>, <pʌ**mp**ʌrʌm>, and <kʌ**ɳɖ**ʌm> (‘raw mango’, ‘spinning top’, and ‘continent’, respectively).Turning to consonant clusters, all consonants in the native Tamil repertoire except for /r/ and /l/ can geminate as in words like /pʌkkʌm/, /pʌʈʈʌm/, and /t͡ʃɛʊʊaːj/ (‘near’, ‘kite’, and ‘Tuesday’, respectively). Some consonants are realised both in short and long duration (e.g., /k/ and /kk/). The long consonant is called a geminate. Three consonantal clusters are combinations of geminates /kk, t͡ʃt͡ʃ, ʈʈ , pp, nn, ɳɳ , mm, n̪n̪/ preceded by /r/, /j/ or /ɻ/. Native Tamil phonology does not allow clusters in the word-initial and word-final position, but loan words bring the exception to this rule (e.g., ‘**gl**obal’- க்ளோபல், ‘**gl**ass’ – க்ளாஸ், ‘ba**nk**’ – பாங்க், and ‘pa**rk**’ – பார்க்).

Tamil is an agglutinating language, and an important aspect of word-level phonology is around the morpheme boundary in inflected words (e.g., root + past tense marker in verbs) and compound words (e.g., *Chennai* + *pattinam* = *Chennaippattinam*, Chennai city). Adjacent morphemes trigger morpho-phonological processes with a strong influence of a rich set of phonological principles called *sandhi.* One abiding influence is to maintain sonority. When the morpheme ends in a vowel and the next word begins with a consonant, then there is a geminate (see example above of *Chennai<u>pp</u>atinam*). Another combination at the boundary is a morpheme-final vowel followed by a morpheme-initial vowel. In such instances of consecutive vowels the consonant /j/ or /ʋ/ is inserted with the choice of consonant constrained by the value of the final vowel (e.g., *paravai* (‘bird’) + *ai* (accusative) is *paravai**y**ai* but *puu* (‘flower’) + *ai* (accusative) is *puu**v**ai*). Turning next to the occurrence of consonant sequences across morphemic boundaries, the examples are several: with nasals such as /ɳk, ɳp, nk, np/ in words like *thira**nk**al, ka**nk**aanippu* and *paya**np**aduthu* (‘skills’, ‘tracking’, and ‘apply’ respectively), retroflexes like /ʈk, ʈs, ʈp/ in words like *u**tk**ondu*, *naa**tk**al* and *u**tp**attu* (‘feed’, ‘days’, and ‘subject’, respectively), and trills in words like *aasiriya**rk**al,u**rp**aththi* and *pa**rk**kal* (‘teachers’, ‘manufacture’, and ‘teeth’, respectively). Similarly, there are several instances involving approximants at the morphemic boundary such as /lk, lp, l ʋ, ɭk, ɭʋ, ɻʋ/ in words like *kappa**lk**al, seya**lp**aadu, and na**lv**aravu* (‘ships’, ‘function’, and ‘welcome’, respectively).

The Aspirated-Voiceless and Unaspirated-Voiced pattern

The Tamil phonological system comprises unaspirated and voiceless sounds and the orthography has dedicated akshara for these sounds. However, since aspirated and voiced sounds do occur in loan words, these are accommodated in the orthography by doubling the

use of the available akshara set. Thus, akshara for unaspirated consonants are used to also represent the corresponding aspirated consonant and the akshara for voiceless consonants for the voiced counterpart (e.g., akshara க for /kʌ/, /gʌ/). This opaque feature of the orthography does not necessarily hinder decoding of words because, for the most part, voicing changes do not bring a change in meaning, although there are exceptions because of loan words (e.g., /baːʊʌm/ - /paːʊʌm/, 'expression' – 'sin', where *baavam* is a borrowing).

There is considerable debate on the characterization of voicing in Tamil phonology. While the classical text, Tolkaappiyam, does not differentiate between voiced and voiceless consonants and has named both types under the valli<u>n</u>am consonant set, others propose that voiced consonants are specific to loan words. What is undisputed however is the number of positional constraints in Tamil phonology. For example, stop consonants in the initial position and geminates in the medial position should be voiceless (e.g., *saththam*, sound) but the consonant is voiced if in the medial position and followed by a homorganic nasal (e.g., *kaagam*, crow). The voicing rule for the stop consonants in the medial position is not straight forward. While the rule could be mostly true for the retroflex (/ʈ/), it is inconsistent for other stop consonants (an example of the lack of clarity is in the name *Prakalad* where the pronunciation of the akshara <ka> is also rendered as *Pra<u>g</u>a<u>l</u>ad* and the Sanskrit borrowing, *Pra<u>ha</u>lad*).

Spoken varieties

The description of Tamil phonology thus far has been with the formal, literary variety. In the spoken varieties, the phoneme inventory and positional constraints change. In Spoken Malaysian Tamil for example, the number of central vowels increases dramatically with use of not just the low central /a/ and /aː/ but also high central /ʉ/ and mid central /ʌ/ and /ə/. Voiced consonants also appear, as do other variations in the consonant repertoire (Ramasamy, 2010). Similarly, within the southern state of Tamil Nadu in India, the phoneme count varies across the dialects of the region. Keane (2004) reports 21 phonemes in the spoken Tamil varieties near Madras (the current Chennai area). Within this Madras Tamil variety are included fricatives, affricates, and voiced counterparts for plosives and an ongoing reduction in distinction on the /r/ and /ɾ/, and the nasals /n̪/ and /n/. Annamalai and Steever (1998) report 22 consonants including allophones, while other counts are more extensive. For example, Christdas (1988) demonstrated 35 phonemes in the Kanyakumari Tamil variety. A further characteristic of the spoken varieties across the region is that the distinction between phonologically-close consonants (such as l, ɭ, ɻ) is often blurred.

The ambiguity with voicing in Tamil phonology appears to make a difference to the pattern of acquisition of literacy. For example, in a comparison of simple frequencies of voicing in primary school children, Malini (1993) found variability in oral language sampled during casual speech, careful speech, word list reading and passage reading. In the word-initial position, use of voicing was more frequent than voiceless consonants in casual speech, but the opposite was true for passage reading. In other words, the rules of written Tamil were followed more carefully for word-initial phonology when children dealt with text. In the intervocalic word-medial position, voiceless velar fricatives were more common than voiced velar plosives in both causal speech and passage reading. It appears that children were over applying the rule by devoicing to word internal plosives. One important trend from the above data is that the mapping of voiced and voiceless plosives to orthography is a pressure point in Tamil reading for some children. We found some evidence of this in a reading survey of 120 children in Grades 2 and 3 in three schools in Karaikudi (Tamil Nadu, India). Three words—<uːɾɖi>, <kaːɳbi>, <sʌŋgʌɖi>, 'vehicle', 'show', and 'topic', respectively—were chosen for analysis because all had the potential for devoicing errors on sounds appearing in the non-initial position in the word. We found 23.33% of children reading <uːɾɖi> as /uːɾʈi/, 9.16% reading <kaːɳbi> as /kaːɳpi/, and 5% reading <sʌŋgʌɖi> as /sʌŋgʌʈi/. These devoicing errors are made by a relatively small percentage of children, but the numbers are significant enough to suggest that fine-grained representation of voicing-devoicing is perhaps an important aspect of phonological knowledge in Tamil, and in turn of Tamil orthographic knowledge. It could be that children unfamiliar with the target words sounded out the written string using the default voiceless option for the ambiguous akshara (i.e., <d> read as /t/, <b> as /p/, <g> as /k/).

The trends from these exploratory reading studies suggest that it is reasonable to expect that phonological characteristics of the spoken varieties will also be a potential source of error in spelling words in standard literary Tamil. But it is also possible that classroom instruction may reduce its effects. In a Grade 3-4 narrative writing programme in Kannada for example, an (unintended) outcome was a reduction of diglossia effects on children's spelling (Nag, 2013). The purpose of communicating for an audience seemed to perhaps increase use of the literacy variety in children's writing. This pattern appears in the one other study we found on spelling performance in Tamil by Aaron and Joshi (2005): Most errors were on the retroflex and secondary vowel diacritic markers and characteristics of casual, conversational speech such as palatalization and cluster reduction were not seen; children appeared to always use

the literary variety to inform their spelling. These interesting trends in contexts of Tamil diglossia clearly need research attention.

Tamil morphology

Tamil is a morphologically rich language. Nouns and verbs in Tamil can be simple or compound and differ based on what they stand for. Adjectives and adverbs may be derived and there are several instances of compounding in the language. This section will describe features of noun and verb morphology to illustrate the morphological, morphophonological and morphosyntactic features of Tamil (for a comprehensive review see Krishnamurti 2003; for a comparison with other Dravidian languages see Amritavalli, 2017).

Nouns

Nouns (called *peyarccol* in Tamil) are classified as rational and irrational, as in other Dravidian languages. Rational nouns include humans and the Gods, whereas irrational nouns are everything else, living and non-living. Rational nouns can be further classified based on gender and number (masculine, feminine, and neuter; singular and plural). Nouns in Tamil are inflected for number, gender, and case. The basic form of the plural marker is /–kʌɭ/, which is geminated to /–kkʌɭ/ if following a long vowel. The singular forms are unmarked. Examples of gender markers are /–ka: ɾʌn/ and /–ka: ɾi/ in words such as *samayal.kaaran* and *samayal.kaari* (male cook and female cook).

Case markers are bound morphemes suffixed to the root noun. Excluding the benefactive and vocative case described by linguists, the marking for the other seven cases are illustrated with the Tamil word for fruit, *pazham* (/pʌɻʌm/):

- The nominative case has no marker. Thus while the noun suffixed with other case markers will have varied final syllables (e.g., /pʌɻʌ̹̈ttil/ and /pʌɻʌ̹̈ttoːɖu/; ‘in the fruit’ and ‘with the fruit’), the nominative case is the noun stem /pʌɻʌm/, ‘fruit’.
- The accusative is marked by /-ʌj/. In the case of a rational and definite object noun phrase, the accusative marker is obligatory, but in the case of the irrational and indefinite noun phrase, it is optional (e.g., /pʌɻʌ̹̈ttʌj/).
- The instrumental and sociative case has multiple markers. The suffix /–aːl/ indicates causation and thus is instrumental (e.g., /pʌɻʌ̹̈ttaːl/, ‘with the fruit’) and suffixes /–oɖu/, /–oːɖu/, and /–uɖʌn/ indicate connection and thus are sociative markers (e.g., /pʌɻʌ̹̈ttoːɖu/, ‘with the fruit’).

- The dative case is indicated by /–ku/, /–kku/, and /–ukku/. While the suffix –kku occurs after the vowels /i, iː, ʌj, u/ and /–ku/ occurs on nouns ending with euphonic increments /in/ and /an/, /–ukku/ follows all the other nouns (e.g., /pʌɻʌt̪t̪ukku/, 'for the fruit').
- The ablative case of separation or motion is indicated by /–iɾun̪d̪u/. Inanimate forms take the marker /–il/ whereas animate nouns take the marker /–iɖʌm/ followed by /–iɾun̪d̪u/ (e.g., /pʌɻʌt̪t̪iliɾun̪d̪u/, 'from the fruit').
- A variety of suffixes communicate genitive case. Suffixes /–ʌd̪u/ and /–uɖʌjjʌ/occur with nouns to make a complex noun phrase; suffix /–in/ occurs with nouns which function as noun modifier, its functions are no different from euphoric increments (e.g., /pʌɻʌt̪t̪inuɖʌjjʌ/, 'of the fruit').
- Locative case markings are most commonly /–il/ occurring with both inanimate and animate nouns, and /-iɖʌm/ with only animate nouns (e.g., /pʌɻʌt̪t̪il/, 'in the fruit').

A sub-class of nouns is pronouns referring to the first, second and third person. Interrogative pronouns are formed by the addition of the vowel 'e' to nouns while demonstrative pronouns prefix /ʌ–/ (/ʌʋʌn/ and /ʌʋʌɭ/, 'that' he and she), and /i–/ (/iʋʌn/ and /iʋʌɭ/, 'this' he and she). The first, second and third person forms are inflected for number but markers for gender are only carried in the third person. A feature that is shared with other languages in the Dravidian language family is that the third person plural marker differentiates between human and non-human but not gender or number. The 'we' in Tamil could be inclusive (/n̪aːm/) or exclusive (/n̪aːŋgʌɭ/).

Adjectives in Tamil (called uriccol) can occur naturally (e.g., /n̪ʌllʌ/ and /pʌɻʌjja/, 'good' and 'old') or are derived from nouns by the addition of /–aːnʌ/ and /–uɭɭʌ/ (e.g., /ʌɻʌgaːnʌ/, /iɾʌkkʌmuɭɭʌ/, 'beautiful' and 'sympathetic'). Similarly adverbs can either occur naturally or are derived typically from nouns.

Verbs

Verbs (called vin̠aiccol) are of two types: the finite verb with tense and person-number-gender (PNG) markers and a non-finite verb which does not mark the PNG. PNG markers attached to verbs always agree with the nouns, and examples of these morpho-syntax dependencies are:

a) *avalpaarththaal* ('She saw', /ʌʊʌɭ + paːɾ +t̪ + t̪ + aːɭ/; ʌʊʌɭ - she, noun; paːɾ - see, verb; t̪ - sandhi; t̪ - past tense; aːɭ - 3rd person singular female).
b) *avanpaarththaan* ('He saw', /ʌʊʌn + paːɾ + t̪ + t̪ + aːn/; ʌʊʌn - he, noun; paːɾ - see, verb; t̪ - sandhi; t̪ - past tense; aːn - 3rd person singular male).
c) *avarkalpaarththaarkal* ('They saw', /ʌʊʌɾkʌɭ + paːɾ + t̪ + t̪ + aːɾkʌɭ /; ʌʊʌɾkʌɭ - they, noun; paːɾ - see, verb; t̪ - sandhi; t̪ - past tense; aːɾkʌɭ - 3rd person plural)

Verbs in Tamil, similar to other agglutinating Dravidian languages, are inflectionally packed. Verbs inflect on tense, mood, and aspect. The present tense markers are /–kir, –kkir,/ (e.g., /peːsukiraːn/, 'he is talking'), past tense markers are /–t̪ , –tt̪ , –in, –n̪d̪, -i/ (/peːsinaːn/, 'he talked'), and the future tense markers are /-p, –pp and –v/,(/peːsuʊaːn/, 'he will talk'). Imperative verbs occur only with the second person. Singular is unmarked and hence is the same as the verb stem; negation, plural and plural negative are marked by /–aːt̪/, /–uŋkʌɭ/, /–iːɾkʌɭ/, respectively. The indicatives also take a positive and negative form.

Returning to our focus on orthographic knowledge, the multiple sandhi rules and morphophonological processes will both be important to examine, particularly since diglossia permeates these aspects of Tamil as well. For example, in our survey in Karaikudi, we recorded a distinction between the Brahmin dialect and the working class dialect, in a variety of multi-morphemic words ('to lie down': *taaccu* versus *paduththu;* 'didn't you pluck?': *killiniyono* versus *killinella*; 'didn't I tell?': *sonneenoonoo* versus *sonnenla*). These and other instances of diglossia around the phonological realisation of inflections and the morphemic boundaries most probably influence development of orthographic knowledge. While preliminary evidence of difficulty on morphemic boundaries is reported by Aaron and Joshi (2005) we were unable to find any research report examining whether inflections are pressure points in Tamil literacy acquisition.

Tamil literacy instruction

Our discussion of Tamil literacy instruction is specific to the practices recorded in the Indian state of Tamil Nadu. This state has a strong history of universalist social policies as well as trust in public institutions (Drèze & Sen, 2013). The public funded school system is administered with comparative efficiency and program continuity when compared to other states in India. Importantly, in the context of literacy instruction, successive governments have avoided fast-track measures that have gained in popularity in many low and middle income countries (e.g., low paid, part-time teachers, cash vouchers to private schools).

Instead, the early years curriculum has been strengthened through multi-party partnerships (Neisz, Krishnamurthy & Mahalingam, 2012). The early grades program in this state is called Activity Based Learning, or ABL, referring to a commitment to a child-centered, activity-filled approach to learning. First introduced in 2003, the ABL method is one of India's celebrated stories of education reform. Reports on the efficacy of ABL in Tamil Nadu are limited to thick descriptions (e.g., Anandalakshmy, 2007), commentaries (e.g., Krishnamurthy, 2011) and secondary data analysis of children's outcomes (Aslam, 2016). No controlled trials are available.

Schools in the ABL programme use government prescribed material. Tamil textbooks are supplemented with multiple teaching-learning resources. These include a set of activity cards, a two-lined book for Tamil writing practice, several story books, and locally sourced materials. There are also homework cards and evaluation cards. The ABL programme began with a focus on activity cards but transitioned into a hybrid programme with a prescribed textbook and activity cards that mirror the textbook and additional material for reinforcement. The sequence of activities in the instruction scheme is illustrated as a fold-out paper ladder prominently displayed, within the child's reach. Learning steps for Grades 1 and 2 are together on one ladder and there is another ladder covering learning targets for Grades 3 and 4. This simple illustration captures the philosophy that children can move at their own pace; children in Grade 2, for example, can continue consolidating lower level skills, while those in Grade 1 can pull ahead if they have attained a particular milestone. This pedagogical commitment to child-paced instruction is compromised in the first instance by a more recent policy change that insists that all children must move up to the next grade at the end of each academic year; no child should fail and repeat an academic year. A second level of compromise is in the individual interpretations of the ABL programme within different schools (more details below).

Activities for each skill level (or steps on the ladder) may be seen to roughly correspond to a chapter in a traditional textbook. Each grade has multiple milestones, and in each milestone, there are multiple activities. Activity icons cover discussion, storytelling, singing, drawing and coloring, homework, and term tests. There are also subject-specific logos such as reading a poem, writing practice for akshara and word games. Activity cards are colour coded, with the cards for the Tamil language in blue (mathematics is maroon, science green, social science orange, and English pink). Finally, when a milestone is reached, a star logo on

the ladder indicates that it is time for evaluation. In addition the logo of a book with a bouquet interspersed throughout the programme signals that the child can take a break from the textbook and activity cards to read books in the class library.

The first introduction to the akshara set is through visually similar but phonologically distinct akshara (e.g., ட, ப, ம; /ʈʌ, pʌ, mʌ/). These are taught as singletons, embedded in mono- or bisyllabic words, tagged to pictures, and through copying practice (see Table 1). The prescribed scheme for Grade 1 focuses on consonants with the inherent vowel /ʌ/, consonants with the vowel diacritic /aː/ and consonants as a phonemic unit. The introduction of the phonemic consonant in the first year is unique because akshara instruction regimes typically postpone this to the middle school years with consequences on development of both orthographic knowledge and phonological processing (e.g., Kannada: Nag, 2007). Examples of phonemic consonants (i.e. the akshara with the dot above) in Grade 1 are seen in these lists of rhyming words பட்டம், மட்டம், சட்டம், வட்டம், கட்டம் (/pʌʈ̪ʈ̪ʌm, mʌʈ̪ʈ̪ʌm, sʌʈ̪ʈ̪ʌm, ʋʌʈ̪ʈ̪ʌm, kʌʈ̪ʈ̪ʌm/; 'kite', 'level', 'law', 'circle', and 'grid', respectively), and words with the vowel diacritic /aː/ சாயம், காயம், தாயம்,மாயம் (/saːjʌm , kaːjʌm , ʈ̪aːjʌm , maːjʌm/; 'dye', 'injury', 'dice', and 'magic', respectively). This early introduction of the phonemic consonant is useful to initiate reading of authentic texts because several early acquired words in Tamil end in closed syllables which can only be written using a phonemic akshara. In addition, the presence of such lists suggests a focus on phonological awareness at the level of syllable and body-coda processing. Finally, a range of activities point to an oral language and print-rich curriculum for literacy learning (for a summary of the Grade1 curriculum, see Table 1).

Table 1: A summary of the textbook for Grade 1

Term	***Oral language and print related activities***	***Singleton akshara***
1	Use of pictures: • Name and describe pictures with a focus on target sounds • Match pictures to given word labels • Label pictures Sound awareness: • Listen to monosyllabic and bisyllabic words with target sounds whose akshara will be or have been introduced • Identify words with target long or short vowels • Identify rhyming words Akshara and word reading: • Identify familiar akshara in a word, identify unfamiliar akshara in a word • Read words with target akshara, combine target akshara to form words • Read words which also have geminates and clusters of taught akshara • Write one akshara of the given words in each tile (e.g. /goːpuɾʌm/ 'ornate temple gate' - 4 tiles for <goː>, <pu>, <ɾʌ>, <m>, கோ. பு. ர. ம்) • Sequence words by the standard vowel order (e.g., *aṇil*, *aama*i, ... ('squirrel', 'tortoise',…) • Sequence pictures with target akshara in word initial position by the standard consonant-vowel order (e.g., *kaṇ, kaagam*, *kiḷi*, *kiiri*; 'eye', 'crow', 'parrot', 'mangoose') Word building and writing: • Copy akshara or words with the target akshara • Fill in akshara missing from all word positions, or in words with more than one akshara missing • Complete crosswords • Combine words to make new words • Arrange jumbled-up akshara to form words • Form new words hidden in a longer word • Substitute the medial akshara to form new words • Write short sentences of 2-3 words length with a focus on target akshara Spelling to dictation (Term 3): • Write the dictated word Reading connected texts (Term 3): • Read short phrases • Write the missing word in a sentence	The order of akshara introduction is with the inherent vowel and the vowel /aː/ • ட, ப, ம, ர, ழ (/ʈʌ ,pʌ ,mʌ ,ɾʌ ,ɻʌ/) • பா, மா, ய(/paː, maː, ja/) • க, ச, த, ந (/kʌ, sʌ, t̪ʌ ,n̪ʌ/) • ங, ற (/ŋʌ, rʌ/) • ல, வ, ன, ண (/lʌ, ʋʌ, nʌ, ɳʌ/) • ள, ஞ (/l̪ʌ, ɲʌ/) These are introduced in parallel in words (see adjacent column). The above consonants are also introduced without inherent vowel (/C/, phonemic consonants) but these are always encountered in a word context
2		• அ, ஆ (/ʌ, aː/ – the primary form) • இ, ஈ (/i iː/) and diacritics ி and ீ with the taught consonant set, e.g., கி, கீ (/ki, kiː/) • உ, ஊ (/u, uː/) and diacritics with the taught consonant set
3		• எ, ஏ (/e, eː/) and diacritics ெ and ே with the taught consonant set • ஐ (/ʌj/) and diacritic ை with the taught consonant set • ஒ, ஓ (/o, oː/) and diacritics ொ and ோ with the taught consonant set • ஔ (/ʌʊ/) and diacritic ௌ with the taught consonant set • Lists with standard vowel order, including lists with ஃ The entire akshara matrix of V, C and CV combinations

Examining the variety in Tamil instruction

We sampled instruction in Grades 1 to 3 in three schools that followed the ABL methodology in Karaikudi, a Class 1 town in the southern district of Sivaganga in Tamil Nadu. Karaikudi with a population of more than a hundred thousand has literacy rates that are closely similar to the statistics of the whole state (e.g., as per Census of India, 2011, 81.5% of over seven year olds in Karaikudi were considered as in-school or literate, the number is 80% for Tamil Nadu state). The main language of instruction in schools in our survey was Tamil (Tamil medium schools) and catered to families with equivalent socio-economic profiles. All the schools were run under private management but teacher salaries, ABL materials, children's uniforms and school meals were funded by the government. Despite the centralised ABL curriculum and controlled supply of ABL materials to schools, we found differences in literacy instruction. Each school had a different focal point from which the curriculum was delivered. While school 1 focused on the activity materials, school 2 stayed close to the textbook and school 3 transferred the focus to the black board. A point to note is that anecdotal evidence suggests that this variety is also present in government schools, although a systematic survey of classroom literacy practices within this sector is currently unavailable.

School 1: Activity based learning

Teachers in this school were amongst the first in the district to be trained in the ABL methodology, and had experience with activity cards without a textbook and, approximately ten years on, were using both activity cards and the textbooks. Each day in school started with the teacher announcing the subject for the day. Children then individually picked up the card corresponding to the level they were in for that subject and settled down to work. Some children remembered their level but when unsure the teacher told them where to begin. In parallel, children identified their working group for the day from a set of six logos also shown on the activity card. Since two or more children could be at the same level, oftentimes children had to share a card. Mobility was not restricted; children could pick up their card and choose to write or work through the activity card at the low-level black board running the perimeter of the room. There was usually no group teaching unless two or more children were progressing at the same pace and were at the same level. The school also encouraged children to bring low or no-cost materials that were easily available in their neighbourhood

(e.g., *rangoli* colours, vegetables, pumpkin seeds and cotton for craft work and class projects).

Teaching in this school was typically informed by guidelines implicit in the activity cards. For example, in Grade 1, writing cards provide arrows to show the formation of the akshara(e.g., for அ, the circle is shown first, followed by the curve below, followed by the horizontal line and finally the vertical line). We found these guidelines were used although not consistently. In addition, teachers also wrote on the board to show the sequence of strokes for each akshara.

School 2: Textbook led learning

The textbook was the main teaching resource in this school. In Grade 1, morning sessions typically had children naming and labelling pictures, and then copying and reading these picture names. These words were linked to target akshara (see Table 1). When a new word was introduced there was a lot of writing. First, the word was written with a pencil into the notebook and then practiced five times by writing with chalk onto a handheld slate. The same words were again assigned as homework to be written on a slate. Initially the teacher wrote new words in each child's notebook but six months on children were expected to copy independently from the blackboard. Afternoon sessions were dedicated to spoken language activities (picture description, singing, storytelling).

Ideas implicit in the ABL ladder were often renegotiated. Group activities and activities that needed locally sourced material were skipped. Opportunities for exploring the language through expressive language activities were limited. In second and third grades, any given day typically focused on the teaching of one subject with a new lesson from the Tamil textbook coming up about once a week. The typical routine differed across teachers. One teacher taught each lesson by following a set sequence of activities: She put down on a handheld slate the textbook questions at the end of the lesson along with a standard answer and asked children to copy this into their notebooks. She occasionally included topics beyond the textbook focused on general knowledge. Another teacher had a very different sequence. She introduced the lesson one day, asked questions given at the end of the chapter the next day and, as children answered, she wrote on the blackboard. Children then copied this teacher-written answer into their notebook, carried it home, and copied both the question and the answers again three times as homework. This particular class also had daily spelling tests,

and poor spellers retook the test. The second teacher appeared to have a closer engagement with what children wrote and learned from each lesson.

School 3: Blackboard led learning

In this school, each class had big-sized blackboards on three walls. When a new chapter was introduced, teachers wrote key words using a fraction of one blackboard. The next day the word list was moved to a board on an adjacent wall where it was left for a long time. We found that some information, depending on space and teachers' judgement of what still needed revision, could be left on one of the adjacent walls for up to a term. Since new words were constantly added to the adjacent boards, almost all the words for the term were available for a child to see. Every morning children read all the words on the board in a chorus. While choral lessons were the norm, the strategy for word-level reading changed by grade. In Grade 1, children read akshara-by-akshara (<a-ni-l> for /ʌɳil/ ('squirrel'), <si-*ng*-ga-m> for /siŋgʌm/ ('lion'). This sequence of sounding out and then blending occurred several times in a day. There were subtle exchanges during the sounding out to focus children's attention on vowel length and phonological realisation of confusable consonants. In Grade 2 the read aloud sessions did not have akshara-by-akshara sounding out. The writing on the wall and the chorus captured the teaching philosophy of this school*: A lot of reading aloud practice improves reading, and keeping word lists in the classroom allows for a long period of practice.* Approximately 90% of textbooks appeared to be covered with this method of instruction; the group activities and group projects were done to a lesser extent.

Even though there was a lot of teacher-written print in the classroom, children did not see teachers write very much. Most of children's writing was to copy from the large boards, and sometimes to write from memory. For Grade 1 the copying was of akshara and word lists; for Grade 2 it was questions and answers which were again recited every morning (e.g., *Who stands at the traffic signals? Traffic police.*). While children worked on a slate or the low-level black board, teachers moved around. Corrective feedback or further explanation occurred when a teacher stopped behind a child upon spotting an error. There was a lot of drilling, reciting, and copy writing practice in this school.

Taken together, we found distinct interpretations of the state mandated ABL programme in each school. For example, at the level of akshara instruction, teachers in school 1 brought focus on vowel length with intonation and an accompanying gesture. Children repeated

words with this dramatisation at the appropriate point in a word. The routines were repeated when children read words to us, accompanying long vowels in words with lengthy articulation and an additional tilt of the head (e.g., /k**a**ːŋbi/, /mʌ**sa**ː**la**ː/). Success rates on vowels appeared to be higher in Grade 1 compared to other schools, perhaps because that is when the practice is the most prevalent in this school (School 1, 2 and 3 at 48%, 21% and 30% respectively). However by Grade 2 all instruction strategies appeared to show similar vowel learning (Schools 1, 2 and 3: 58%, 40% and 56% respectively). The use of dramatised cues for long vowels appears to help jump start orthography-phonology mapping. But, from the limited data we have, there is no reason to believe that the sing song recitation followed in the other schools slows down symbol learning.

The similarities across schools are also worth highlighting. All schools kept a record of children's performance on state- supplied templates for formative assessment and summative assessment. All children, irrespective of their performance, compulsorily moved to the next grade at the end of each academic year. A struggling first grader had to start on activities for second grade when he moved to the next grade even if he needed more of the earlier placed activities. Equally, even if a first grader performed well, he moved to activities assigned to Grade 2 only in the next academic year. At the level of infrastructure, all classrooms had a low level black board running along the perimeter of the room with individual spaces for children to write. Slates were used often because the government supplied notebooks were too slim to last a full year of writing practice. In all schools teachers wrote on the board to show sequences of strokes to write the akshara and we saw feedback to individual children, but this occurred in a serendipitous fashion. Interviews with teachers and a review of classroom resources such as the charts, material written on blackboards, and children's note books showed that children were not taught the vowel diacritic in isolation but always in a symbol block with the consonant. Instead of this, teachers in all schools liberally used the rich Tamil vocabulary to describe the orthography, particularly the interior of the akshara. Finally, none of the schools expressly devalued dialect in language usage, but the tacit understanding between teacher and child was that standard Tamil pronunciation is expected when reading or spelling, and this was easily reinforced in joint read aloud sessions. This invariance in pronunciation for written Tamil by primary school teachers has been recorded by other authors as well (Geetha, 2012; Sugathapala De Silva, 1986).

Orthographic knowledge in Tamil

Performance on single word reading and spelling requires multiple layers of orthographic knowledge (see Nag, 2017 for developmental trajectories in akshara-based decoding). Greater mastery of the rules of orthography may allow for higher accuracy and automaticity in reading and spelling. The very small body of literature that is available in the field of Tamil orthographic knowledge suggests that transparency assists in learning and that the younger reader and poor reader are particularly vulnerable to phonological errors (Aaron & Joshi, 2005; Bhuvaneshwari & Padakannayya, 2009; Nag-Arulmani, 2003). We extend this limited earlier literature with reports from two studies looking specifically at three aspects of orthographic knowledge: word reading, spelling and, for the first time in the literature, on early writing routines in Tamil.

Single word reading

Balambigai and Purushothama (2011) examined the role of orthographic knowledge in word reading with a group of 60 native Tamil-speaking children studying in Grade 3 in four government schools following the ABL curriculum described above. Most schools in the sample were close to the daily routines described under school 1 (Activity Based Learning). In each school, the class teacher was asked to rank order the students according to their academic performance, and for a contrastive analysis the highest achieving 10% and the lowest achieving 10% were taken for the study. Children were tested for speech and hearing difficulties, and if the results were positive they were dropped from the study. School records showed that all of the children in the high achieving group (30 children) were placed towards the upper end of the ABL ladder and the low achieving group (30 children) were at the bottom. Children from Grade 3 were chosen because that is probably when 'learning to read' changes to 'reading to learn' and this is the youngest possible age group who may be considered as proficient enough to assess several aspects of orthographic knowledge.

A word list of 71 words was taken from the Grade I Tamil textbook, and this allowed for ensuring all words were familiar to participating children in school. The list included at least two words with each vowel in the Tamil repertoire, and all consonants and consonants with vowel combinations. Thus a child who read the whole list correctly would have applied all possible akshara-level Tamil orthographic rules. The task was to quickly read words briefly exposed on a computer screen. The words were printed using the 'Baraha Tamil' font with font size adjusted automatically according to the length of the word. The longest word had five akshara (மிதிவண்டி, /miɖiʋʌɳɖi/) though the maximum number of disconnected units

was seven (ஔவையார், /ʌʊʌjjaːr/).The word stimuli were randomized and presented in the same order. The exposure time for the brief exposure condition was controlled to be at one second duration. The interval between words (the inter-trial interval) was held constant.

The child sat comfortably in front of a computer screen on which the words were presented. Before the test, each child was instructed individually in Tamil to read words appearing on the screen as quickly and as accurately as possible. As the child read each word the vocal reaction time was recorded. This is the time between onset of the word on the screen and onset of the child's vocalisation as they read the word aloud. If an error was made, the error was noted down by the experimenter. This procedure was followed for all the 71 words before the second part of the study began. In this phase, words that each child skipped or misread during brief exposure were presented again to be read at a long exposure. This time the child was encouraged to read aloud each akshara, or spell out the word. This instruction was introduced because it could perhaps help differentiate whether the child's problem with reading a word was at the level of automaticity (that is, the child could read when given enough time but not when exposure to word was too quick) or at the level of knowledge of the orthography (that is, even with enough time to decode the child still struggled). Once again, the child's responses were noted down verbatim.

As expected, good readers read significantly more words than poor readers. When words were shown in a brief exposure, the good readers read 94.8% of words correctly while the poor readers read 65.31% correctly. The accuracy rate of the good readers was almost uniform across the group (SD= 2.78) reflecting the properties of the simple word list; an alternate list with low frequency words may have dispersed the scores even among the good readers. There was a scattering of scores among the poor readers (SD=13.74), suggesting that, although all poor readers had lower accuracy when a word was briefly shown, some were much worse than others. Among the good readers, the majority of the words on the list (85.91%) were read correctly by all the children. Among the poor readers, all read 32.39% of words correctly and these words only had akshara with the inherent vowel, and the vowel diacritics for /i/ and /aː/.

To further determine the nature of orthographic knowledge, children's scores for reading at brief exposure and reading at long exposure were analysed. In the long exposure condition, good readers could read 71.81% of the earlier skipped or misread words suggesting that their greater difficulty was in decoding under the pressure of time, when given enough time they could apply their orthographic knowledge accurately. In the brief exposure condition, good

readers struggled to identify words with grantha akshara, akshara with an inconsistent ligaturing rule (the vowels u and u:, e.g. சூ, டூ), and akshara with non-linear arrangements. Most of these difficulties reduced with long exposure though accuracy remained low on words with diacritics for dipthongs /ʌʊ/ and /ʌj/, the vowel diacritic for /o/ (ொ), low frequency akshara /ɲʌ/ (ஞ) and the grantha akshara /ʤʌ/ (ஜ). In other words, some aspects of the symbol set and orthographic rules were still being mastered by the highest achieving readers in Grade 3.

The group of poor readers also improved their performance when given time in the long exposure condition. They could read correctly 67.11% of the words skipped or misread at brief exposure. Although poor readers improved their performance when they had time to decode a word, error analysis suggested shaky orthographic knowledge. More than 33.33% of poor readers had difficulty reading the following akshara at brief exposure: ahu (ஃ), ʌj (ஐ), ʋa: (வா), ʤʌ (ஜ), hʌ (ஹ), ʤi (ஜி), n̪i: (நீ), lu: (லூ), ɖu: (டூ), n̪o (நொ), mo: (மோ), ko: (கோ), ʂʌ (ஷ), pʌʊ (பௌ), se (செ), and s (ஸ்). Sixteen percent had particular difficulty reading the following akshara at brief exposure: o (ஓ), e (எ), ʌj (ஐ), ʌʊ (ஔ), su (சு), ku: (கூ), su: (சூ), ʤu: (ஜூ), mʌj (மை), pe (பெ), ne (னெ), and kʌʊ (கௌ). But even when read in leisure, the poor readers had difficulty with words containing the vowel diacritic ou (ொள), o (ொ), u: (நூ, லூ); and words with low frequency akshara (e.g., ஞ, நு), the grantha akshara (e.g., ஸ், ஜ்), and the symbol ஃ. Other errors were with a disconnected unit of an akshara which lead to a change or omission in vowel diacritics, confusion among visually similar akshara, and poor knowledge of akshara in context. The latter error type is particularly interesting because it demonstrates the possible role that semantic knowledge could have to avert confusion when orthographic knowledge is shaky: as described earlier, the symbol பௌ can be read as /pʌʊ/ or /pe/ + /l̪ʌ/, and must be decoded depending on the word context; thus the word may be / pʌʊr̪ɳʌmi/ (‘full moon’) as against the nonword /pel̪ʌr̪ɳʌmi/. Similar disambiguation is needed to read /gʌʊr̪i/, /mʌʊnʌm / and /ʌʊʊʌjja:r/ (‘a girl’s name’, ‘silence’ and ‘ouvaiyyar, name of a Tamil poet’) rather than the alternative orthographic possibility of /kel̪ʌr̪i/, /mel̪ʌnʌm/ and /ol̪ʌʊʌjja:r/, all of which are nonwords. Poor readers did not self-correct when their decoding strategy left them with a nonword.

The trends in this study show that poor readers in Grade 3 were yet to become adept at attending to the details of print. The vocal reaction time data mirrored this finding. Good

readers were much faster in reading out a word when compared to the poor readers (see Figure 2). The average reaction times in milliseconds for the good readers was 927.49 (SD = 77.48) and the poor readers was 1176.08 (SD = 79.07), and this was statistically a significant difference. Among the good readers, words that took the shortest time included words with primary vowels, consonants with inherent vowels or the vowel /i/ and phonemic consonants. Words that had the longest reaction time were words that included vowel diacritics with more disconnected units /o, oː, ʌʊ/, the low frequency /ʤʌ/, and the symbol ஃ. Unsurprisingly, words with more akshara also had longer vocal reaction times. Among the poor readers, words that took the shortest time were the same as for the good readers, but they still took longer when compared to the good readers. Words that took the longest included all of the instances noted for good readers and also the /uː/ diacritic marker. In summary, greater mastery in orthographic knowledge is demonstrated with greater accuracy and speed of word recognition.

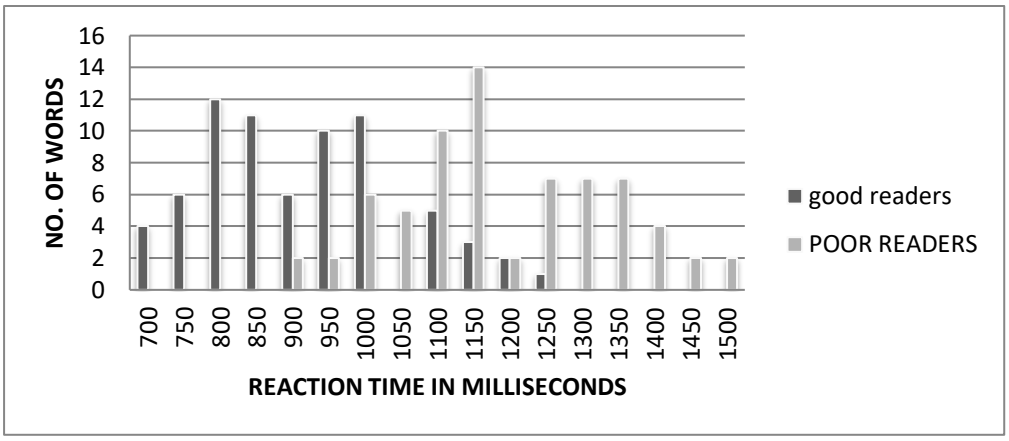

Figure 2: The vocal reaction times of words read by Grade 3 good and poor readers following an exposure of 1 second per word (the brief exposure condition).

Spelling

The Tamil akshara system is particularly interesting for the study of spelling development because mastery is needed at several levels. First, as with other akshara orthographies, the number of symbols in Tamil can be expected to take time to learn. The sheer number of symbols also implies a potential for confusion when selecting the appropriate akshara for spelling, and this confusion may be at the level of closely similar sounds (phonological confusion) or similar looking symbols (visual confusion). In addition, two akshara features may reduce accuracy in spelling, namely, the presence of disconnected elements in the akshara block and non-linear arrangements. A further feature specific to Tamil orthography

is the drop in transparency because one symbol represents both the voiced and voiceless consonants and the stop symbol represents the affricate and fricatives (ச for /sʌ/, /t͡ʃʌ/, /ʃʌ/, although the later can be written with a grantha consonant—ஶ்ரீ as well).

Finally, diglossia may disrupt spelling accuracy in Tamil. We examined the impact of all these features in the spelling of 193 monolingual Tamil speaking children from Grades 1 to 3. All children were from the three Karaikudi schools described above.

The words chosen were highly familiar and acquired before age five. They were balanced for parts of speech (nouns and verbs), mostly bisyllabic but some trisyllabic, all mono-morphemic and no loan words. To understand the nature of children's orthographic knowledge, an akshara-by-akshara analysis of spelled words was done after setting aside akshara that were illegible. The pattern of attainments in this sample broadly replicated uneven attainment patterns reported for the early grades in other akshara-based spelling studies (e.g., Kannada: Nag, Treiman & Snowling, 2010). Accuracy was higher on consonants than vowels and knowledge of visual features of the orthography was higher than knowledge that accurately maps phonology to spelling. All children, including the first graders, gave a spelling that broadly fitted the length of the dictated word; this is evident from the low occurrence of errors due to the addition of an extra akshara or omission of a phonemic marker within an akshara (1.75% and 6.30% respectively). Instead, the common error was of substitution (91.97%) with more than half of these substitutions on vowels rendered as a diacritic marker (56.85%).

A 'pure' visual confusion such as னை written as ன was rare (0.19% of all errors), and confusion with the orthotactics of the writing system was seen in slightly less than 5% of all errors. Instead, most errors could be traced to confusion with a phonological source or a source where the role of orthography and phonology could not be separated. Among the few orthotactic errors seen, the types of confusion in order of frequency of occurrence were a) ignoring position constraints either on the vowel form (e.g., the primary vowel form after a consonant) or in writing the non-linearly arranged vowels in the same sequence as the spoken stream, and b) thinking of a disconnected element in an akshara as an akshara-like unit and substituting the full akshara with only these elements. None of the children including Grade 1 spellers made errors on ligaturing position, which, in Tamil, can become quite changeable for only two vowels (recall that vowel marker /u/ and /u:/ ligatures at different places depending on the shape of the base consonant). Unusually though, such advances in accuracy was not

the case for unconnected markers suggesting that an increasing number of unconnected markers makes symbol learning more complex. The most vulnerable was the marker named tuṉaikkāl (◌ா), (close to 40% of all errors), because this is a visual element shared by multiple vowel markers. Errors in order of frequency were a) omission of the /aː/ marker, b) omission of the element in the /o/ and /oː/ marker, c) addition of the element to the /e/ marker, and d) substituting other vowel markers with the element.

The phonological errors covered the multiple sound families where distinction was blurred in speech even while retained in spelling. For example, there was confusion between the dental lateral (/l/), the retroflex lateral (/ɭ/) and retroflex approximant (/ɻ/) so that a word like /kiɻi/ was written as <kiɭi> (*கிழி* – *கிளி*). Similarly, there was confusion between the dental (/n̪/), alveolar (/n/) and retroflex (/ɳ/) nasals where, for example, the word /mʌɳi/ was written as <mʌni> (*மணி* – *மனி*). Similar trends were found on words with trills, with children mixing up the akshara for the dental tap (/r̪/) and the alveolar trill (/r/). All of these errors were seen across all school types and across grades, with no evident influence attributable to the children's dialects. Other errors suggested that mastery over some aspects of the Tamil phonology was still emergent, such as ending a word with an illegal diphthong or vowel. But by far the most confusing for children in all schools was the short and long vowels and the diphthongs, though for the latter we saw some children approximating the word in spelling (e.g., /ʊʌj/ written as <ʊʌ.j>, *வய்* for *வை*).

A surprising finding from our survey, but in-line with trends reported in Nag (2013) and Aaron and Joshi (2007), was that the effect of diglossia on spelling accuracy was rare. Several words on our list should have attracted diglossia influences because of the spoken variety in the Karaikudi region. For example, the spoken form of <uʈkaːɾ> is /ukkaːɾu/ (*ukkaaru*, 'sit'), <pejʌɾ> is /peːɾ/ (*paer*, 'name') and <kaːɳbi> is /kaːmi/ (*kaami*, 'show'). In other words, although we began with the reasonable expectation that diglossia will be a potential source of spelling errors in Tamil, we did not find evidence for this. The data instead suggests that spelling conventions may be quickly learned, with one possible mechanism being a growing awareness among children that acceptable spellings (written Tamil) must follow a 'standard' phonological register.

Lastly, an intriguing error type, although rare in this sample (0.51%), was the use of English letter substitutions for Tamil consonants (e.g., 'b' for *பி*, as seen in *காண்b* for *காண்பி*, and

தம்b for தம்பி). Children in this survey had minimal exposure to English in school, but the errors appear to suggest that there may have been English tutoring at home.

In summary, a survey of early grade spellings in Tamil shows overlapping advances in knowledge about multiple properties of the orthography. Children show emergent mastery over explicitly taught akshara sets for the vowels, consonants with inherent vowel, and consonants with vowel diacritics. Among the consonants with vowel diacritics, mastery is uneven, with mastery of a consistent ligatured marker among the first to emerge. Slower to learn are those vowel markers that are disconnected from the consonant, and among these, the slowest to stabilize in learning are those vowel diacritics that are ambiguous because they are shared by multiple vowel markers. Spelling studies are particularly rare in the Indic alphasyllabaries, but our limited survey has allowed us to outline the variety of parallel learning demands that may explain the nature of orthographic knowledge and spelling development in Tamil.

Writing routines

The sequence with which children write out the word in a dictation task can shed light on the pressure points for learning the visuo-spatial arrangement of a Tamil akshara. We noted the writing routines children followed as they wrote the dictation task described above. All akshara in the word lists had been taught although schools differed in focus on writing routines and nature of practice to learn spellings (described earlier). We found that 93% of all children followed the routine of writing left to right, completing the writing of all elements of an akshara before moving to the next akshara in a word. The remaining seven percent returned to an earlier written akshara to insert missed elements. Table 2 gives a case-by-case sample of unusual writing routines. Returning to write down a missed part of a word (returns) never occurred for consonants. Instead, all returns were on vowels and these vowels were always a diacritic marker and not in the primary form. Among the diacritic markers, the markers most vulnerable for returns were the ones written in-line and placed in a non-linear arrangement. This is a counterintuitive finding since the available literature in the akshara orthographies suggests that in-line nonlinearities are more vulnerable to errors.

Several points are worth noting from this pattern of writing routines: First, recall that fifty percent of Tamil vowel diacritics have an element of nonlinearity and are written before the base consonant. These are orthographic representations which are in mis-sequence of the speech stream (/ke/ written as /ek/). Despite the higher frequency of these arrangements,

93% of children wrote left-to-right ignoring the spoken sequence (ie. writing /ek/ even as they said /ke/). Importantly, the 7% of children who returned to insert a missed orthographic element did so more often for the missed linear elements rather than the non-linear elements (compare the last four columns, Table 2). A further point to note about returns on linear arrangements is that the tuṉaikkāl (◌ை) attracted the most returns in the writing routine. Since the tuṉaikkāl represents the vowel /aː/ and is also part of the vowels /o/ and /oː/, we propose that factors related to transparency may be more important than linearity for managing returns in writing routines. In summary, information from children's writing routines shows that orthographic knowledge related to linearities and nonlinearities are quickly learned by most children and across most Tamil akshara, while a drop in transparency slows down learning.

We also examined if there were differences in return-rate across schools. Returns to finish incomplete units were seen more often in the activity focused school (14%), than the textbook and blackboard focused schools (only 6.5% and 7%, respectively, returned). It may be the case that the focus on copywriting practice in the textbook- and blackboard-focused schools helped scaffold learning of stroke sequences. But it could also be that the children who returned to insert missing visual elements were the better readers and spellers who had spontaneously recognized errors in their writing and self-corrected, and the activity-focused schools nurtured this self-monitoring and self-correction. A comparison of errors on diacritics across the three schools suggests the later may be the case: Returns when writing are perhaps a correction strategy used by better readers and spellers. More research is needed to examine these relationships between instruction, writing routines and spelling accuracy.

Table 2: Akshara writing during a spelling task: Returns to insert vowel diacritics[1]

Child (Grade)	Initial written sequence	Final written output	In-line, linear arrangement		In-line, nonlinear arrangement	
			In-line, linear	Inserted unit	In-line nonlinear	Inserted unit
CH (1)	உத	உதை			✓	ை
KI (1)	சறு	சிறு	✓[2]	ி		
SR (1)	ச	செடி			✓	ெ
AB (2)	வ	வை			✓	ை
	உத	உதை			✓	ை
SA (2)	படு	பாடு	✓	ா		
RO (2)	பசா	பைசா			✓	ை
TH (2)	வேண்டம்	வேண்டாம்	✓	ா		
JA (2)	கண்	காண்பி	✓	ா		
ME (2)	பேடு	போடு	✓	ா		
MU (3)	செல்	சொல்	✓	ா		
	கெட்டு	கொட்டு	✓	ா		
	சதம்	சாதம்	✓	ா		
KA (3)	பெம்	பொம்மை	✓	ா		
TA (3)	கண்பி	காண்பி	✓	ா		
AR (3)	வேண்டம்	வேண்டாம்	✓	ா		
AN (3)	சத	சாதம்	✓	ா		

Note: [1]In a sample of 193 children there were no returns to enter a missed consonant. [2]With the exception of this one instance, no child returned to insert a ligatured diacritic. All other returns were for disconnected elements of the vowel diacritic.

End note

In this chapter we set out to describe the development of orthographic knowledge after providing a description of Tamil orthography, phonology and morphology, and the variety in early literacy instruction. Of particular interest was the relative impact of orthography-specific transparent and opaque features and language-specific phonological and diglossia-related features. Our focus was on single word reading, spelling and writing routines of children in the first three years of literacy instruction when children are actively learning the surface layout of individual symbols and rules of the orthography. Vowel errors outnumbered consonant errors and phonological errors outnumbered errors because of visual confusion. We found orthography-specific features and ambiguity in orthography-phonology linkages intruding more than diglossia-related features. This trend is interesting since our classroom observations had not recorded specific instruction to raise children's awareness about intrusions because of the diglossia.

We acknowledge that our studies were set up to capture word level processing and it is possible that diglossia effects in Tamil may be more prominent at the level of syntax and sentence level processing. We also acknowledge that a dictation task may not be sensitive to capture diglossia effects (which may also explain the pattern of results in Aaron and Joshi, 2007). While trends from Nag (2013) suggest that increased opportunity for expressive writing reduces those spelling errors that are due to intrusions from the spoken variety, such spontaneous opportunities were rare in the schools in our survey. Our impression is that diglossia effects on spelling is perhaps more evident when children write spontaneously. Anecdotal reports from teachers seem to confirm our impression: They report that when children write memorized answers to questions set out in their book, diglossia effects are again low, but when children write stories or explain a taught concept in their own words, many words may be written as they are spoken (diglossia effect). The absence of strong diglossia effects in children's spelling is perhaps linked to two factors. First, our classroom observations had shown a strong focus on learning through writing, and this practice may have crystallized orthographic representation for literary Tamil. Second, there seemed to be no confusion in teachers' minds about the spelling of words, or when spellings reflect diglossia effects then they should be dismissed as incorrect. This assurance within the Tamil teacher community was recorded by Sugathapala De Silva (1986) in Tamil schools several decades earlier in a sociolinguistic analysis of teachers' attitudes to what the author termed

the 'linguistic cleavages' due to diglossia. Tamil primary school teachers interviewed between 1974 and 1976 were unlike Sinhala, Kannada and Telugu teachers interviewed around the same time. While the latter teachers showed indecision, ambivalence or a diminishing allegiance, respectively, to teaching the 'high' literary form and accommodating diglossia influences, teachers of Tamil showed 'an unquestioned allegiance to the classic norm for formal use' (p. 305). The focus of literacy instruction was defined by all Tamil teachers as 'the ability to read and write *correctly*' (our highlight), where correctness was defined as 'the observance of the correct rules of spelling and grammar" (p. 309).The hold of Literary Tamil in literacy instruction in the early grades was more recently recorded in the areas of teacher education and children's literature in India (Geetha, 2012). The situation does not seem too different in Singapore, prompting the evocative description of the local Tamil population being 'tongue-tied' because of the continued focus on Literary Tamil in children's materials and literacy programmes (Schiffman, 2003). More research is clearly needed to understand when and how diglossia impacts Tamil literacy development, even when orthographic knowledge has been firmly guided to represent the standard language.

One orthography-specific process that demands further research attention is the processing of the visual element called tuṉaikkāl (ு). Errors related to the tuṉaikkāl are particularly interesting because in some instances the errors appear to be because of a confusion with long and short vowel sounds (a phonological error) but in other instances because of shaky knowledge about visual elements in an akshara (orthographic error). Disentangling the processing of this visual element may bring further insights into the phases of disambiguation in intra-symbol processing. Stepping back from the cognitive processes involved in the use of this visual element, the tuṉaikkāl is a symbol of macro-ecological factors influencing the learning task such that they increase the learning demand for the novice learner. As shown in Figure 1, it was through orthographic reform that this element proliferated in the orthography. A second orthography-specific process we noted was the ambiguity that stems from multiple disconnected units in an akshara. Added to this are the demands placed by the voicing-devoicing phenomenon and the morpho-phonology in the language. All these issues would gain from systematic future research.

References

Aaron, P. G., & Joshi, R. M. (2005). Learning to spell from print and learning to spell from speech: A study of spelling of children who speak Tamil: A Dravidian language. In R.M. Joshi & P.G. Aaron (Eds.), *Handbook of orthography and literacy*. Mahwah, NJ: Lawrence Erlbaum.

Amritavalli, R. (2017). Morphosyntax of Dravidian Languages. *Oxford Research Encyclopaedias of Linguistics*.

Anandalakshmy, S. (2007). *Activity based learning: A report on an innovative method in Tamil Nadu*. Tamil Nadu Sarva Shiksha Abhiyan, National Informatics Centre. Chennai, India. Retrieved from http://www.ssa.tn.nic.in/docu/abl-report-by-dr.anandhalakshmi.pdf

Annamalai, E., & Steever, S.B. (1998). Modern Tamil. In S. B. Steever (Ed.), *The Dravidian Languages* (pp.100-127). London: Routledge.

Aslam, M. (2016). Evaluation: more than an afterthought? India's Activity-based Learning programme. Paper presented at *Learning from learning assessments to leave no one behind,* REAL Centre, University of Cambridge. Retrieved from https://www.educ.cam.ac.uk/centres/real/downloads/REALCentre_Conferencebooklet_2.pdf

Balambigai, N., & Purushothama, G. (2011, Dec). *Role of Knowledge of Orthography and Automaticity in differentiating Good and Poor Readers in Tamil*. Poster presented at the Symposium on Language, Literacy and Cognitive Development, The Promise Foundation, Bangalore, India.

Bhuvaneshwari, B., & Padakannaya, P. (2014). Reading in Tamil: a more alphabetic and less syllabic akshara-based orthography. In H. Winskel & P. Padakannaya (Eds.), *South and Southeast Asian Psycholinguistics* (pp. 192-201). Cambridge: Cambridge University Press.

Christdas, P. (1988). *The phonology and morphology of Tamil*. (Unpublished Doctoral Dissertation). Cornell University, Ithaca, NY.Drèze, J., & Sen, A. (2013). *An Uncertain Glory: India and Its Contradictions*. Princeton, NJ: Princeton University Press.

Geetha, V. (2012). Literacy and Reading: A Tamil Experiment. *Contemporary Education Dialogue*, 9, 63-84.

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge, United Kingdom: Cambridge University Press.

Keane, E. L. (2004). Tamil. *Journal of the International Phonetic Association*, 34, 111 – 116.

Krishnamurti, B. (2003). *The Dravidian Languages.* United Kingdom: Cambridge University press.

Krishnamurthy, R. (2011, August 13). *செயல்வழிகற்றல் உருவானது எப்படி? புரட்சி அல்ல இது: பரிமாண வளர்ச்சி* (How did ABL develop? It is not a revolution but an evolution.), *Dina Malar*. http://www.dinamalar.com/News_Detail.asp?Id=61541

Malini, R. H. (1993). *A sociolinguistic study of phonological variables in Tamil learning context at the primary school level* (Unpublished Doctoral Dissertation). Department of Linguistics, Bharathiar University, India. Retrieved from http://shodhganga.inflibnet.ac.in/handle/10603/101289

Nag-Arulmani, S. (2003). Reading difficulties in Indian languages. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 235–254). London: Whurr.

Nag, S. (2007). Early reading in Kannada: the pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading 30*(1), 7-22.

Nag, S. (2013). Low literacy attainments in school and approaches to diagnosis: an exploratory study. *Contemporary Education Dialogue*, 10(2), 197-221.

Nag, S. (2017). Learning to read alphasyllabaries. In K. Cain, D. Compton, & R. Parrila (eds.), *Theories of reading development* (pp.75-98) Amsterdam: Benjamins.

Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research*, 2(1), 41–52.

Neisz, T., Krishnamurthy, R., & Mahalingam, V. (2012). *A history of the activity based learning movement in Tamil Nadu. An independent report submitted to Sarva Shiksha Abhiyan, Tamil Nadu*. Retrieved from http://docs.ehhs.kent.edu/ABLinTN.pdf

Office of the Registrar General & Census Commissioner, India (2017). Annual report on Vital statistics of India based on the CRS-2015. Retrieved from http://www.censusindia.gov.in/2011-Common/Annual_Report.html

Perumal, B. (1989). *Descriptive Study of Standard Spoken Tamil in Malaysia* (Unpublished Doctoral Dissertation). Department of Indian Studies, University of Malaya, Kuala Lumpur, Malaysia.

Rajan, K., & Yateeskumar V. P. (2013). New evidences on scientific dates for Brahmi script as revealed from Porunthal & Kodumanal excavations. *Pragdhara*, 21-22, 279-295.

Ramasamy, M. D. (2010). *Topics in the Morpho-Phonology of Standard Spoken Tamil (SST): An Optimality Theoretic Study* (Unpublished Doctoral Dissertation). University of Newcastle upon Tyne, Newcastle upon Tyne, UK.

Schiffman, H. F. (2003). Tongue-Tied in Singapore: A language policy for Tamil? *Journal of Language, Identity & Education*, 2(2), 105-125.

Sugathapala De Silva, M. W. (1986), Typology of diglossia and its implications for literacy. p. 304 – 311. Bh. Krishnamurti (Ed.) In South Asian Languages. Structure, convergence and diglossia. Delhi: Motilal Banarasidass.

Tamil Lexicon Notation (1939). University of Madras, India. https://archive.org/details/in.ernet.dli.2015.85194

Yang, N. H. (2011). *Loan word adaptation in Tamil*. (Unpublished Master's Thesis). Department of English Language and Literature, National University, Singapore.

Appendix 1 : Tamil linguistic terms with the Tamil Lexicon Notation (Madras Notation) and the International Phonetic Alphabet (IPA).

Srl No.	Madras Notation	IPA	Meaning
1	tamiḻ	t̪ʌmiɻ	Tamil
2	centamiḻ	seṉt̪ʌmiɻ	Pure Tamil
3	koṭuntamiḻ	koɖuṉt̪ʌmiɻ	Spoken Tamil
4	ṭāmili	da:mili	Tamil-Brahmi
5	tamiḻariccuvaṭi	t̪ʌmiɻ ʌrit͡ʃt͡ʃuʊʌɖi	Tamil orthography
6	uyireḻuttu	ujiɾeɻuṯṯu	The vowel and diphthong set
7	meyyeḻuttu	mejjeɻuṯṯu	The consonant set
8	neṭil - kuṟil	ṉeɖil and kuril	The short and long vowel
9	valliṉam	ʊʌllinʌm	The hard plosives
10	melliṉam	mellinʌm	The soft nasals
11	iṭaiyiṉam	iɖʌjjinʌm	The medium consonants
12	āyutaeḻuttu	a:juɖ̪ʌeɻuṯṯu	Symbol ஃ
13	grantha		The borrowed symbol set from Sanskrit
14	puḷḷi	puɭ̪i	The schwa suppression marker
15	tuṉaikkāl	t̪unʌjkka:l	Name for the unattached visual element ◌ா
16	ottaikkompu	oṯt̪ʌjkkombu	Name for the unattached visual element ◌ொ
17	reṭṭaikkompu	ɾeʈʈʌjkkombu	Name for the unattached visual element ◌ோ
18	reṭṭaicuḻi ‘n’	ɾeʈʈʌjsuɻi *'n'*	Name for the type of visual elements in ன
19	mūṉṟucuḻi ‘n’	mu:ndrusuɻi *'n'*	Name for the type of visual elements in ண
20	canti	sʌɲɖi	Morpho-phonology rule
21	peyarccol	pejʌrt͡ʃt͡ʃol	Noun
22	uriccol	uɾit͡ʃt͡ʃol	Adjective
23	viṉaiccol	ʊinʌjt͡ʃt͡ʃol	Verb