Developing discourse-informed Early Grade Readers

Charles E. Grimes, PhD (AuSIL, ANU, UBB)

Abstract: How a story is put together, even for early readers, can significantly hinder or facilitate reading for meaning, which is fundamental for continuing on to higher levels of education. In addition to original stories in a local language, it is also common to adapt Early Grade Readers available in national and international languages. From experience over several years working on around 150 titles in a dozen languages in the Indonesian province of NTT and the country of Timor-Leste, we are initially getting junk that needs major reworking to be useful for children. Without such reworking, the children learn to read some of the words, but can have difficulty following the story. This initially poor quality material is produced by semi-trained local language teams, as well as by well-intended professional linguists and educators. What is missing at a very high level is an awareness of common patterns in narrative discourse in local languages, such as handling background and foreground information, introducing participants, tracking participants, principles for paragraph breaks, handling direct quotes, etc.

Many languages in a region share high level patterns, even across different families. When a source collection is developed in a language of wider communication that attends to these issues, it is relatively easy to develop good quality by-products in many other languages. And we can produce these daughter versions better and faster.

This paper provides examples of these and related issues in the interests of helping people in other regions provide high quality resources for the children (SDG-4).

Key Words: Early Grade Readers, Language in Education, Applied Discourse Analysis, Education in Timor, Presentational Clauses, Tail-Head Linkage

Outline:

1. Background ..................................................................................................................................................................................2
   1.1 High-level assumptions backed up by research ...........................................................................................................................2
   1.2 Language research behind these suggestions ..................................................................................................................................3
   1.3 Clusters of languages for MLE materials: EMBLI & UBB ...............................................................................................................3
   1.4 Discovering local discourse patterns: oral vs. written sources .....................................................................................................4
   1.5 The benefits of a discourse-informed Front Translation in an LWC ..............................................................................................5
   1.6 Key resources for applied discourse analysis and translation .......................................................................................................6
2. Discourse-informed Early Grade Readers .......................................................................................................................................7
   2.1 A framework for cohesion and discontinuity in narrative discourse.................................................................................................7
   2.2 Introducing major participants .........................................................................................................................................................8
   2.3 Tracking participants .......................................................................................................................................................................10
   2.4 Tracking location ..............................................................................................................................................................................11
   2.5 Tracking time ....................................................................................................................................................................................12
   2.6 Tracking events, themes .....................................................................................................................................................................13
   2.7 Repetition, Tail-head linkage, Point-of-departure ............................................................................................................................13
   2.8 Quote formula and complementizers .............................................................................................................................................15
   2.9 Direct vs. Indirect speech ...............................................................................................................................................................17
   2.10 Discourse openings .........................................................................................................................................................................18
   2.11 Discourse closings ............................................................................................................................................................................18
   2.12 Framing questions ..........................................................................................................................................................................22
   2.13 Connectors: use them or not? .........................................................................................................................................................23
   2.14 Natural Rl. order (logical relationships) .....................................................................................................................................23
3. Related translation and cross-cultural issues ......................................................................................................................................25
   3.1 The problem of foreign things ..........................................................................................................................................................25
   3.2 The problem of foreign words ..........................................................................................................................................................25
   3.3 The problem of semantic range of meaning and category mismatches .............................................................................................26
   3.4 The problem of structural or grammatical mismatches ..................................................................................................................27
   3.5 The problem of conventionalized usage and cultural categories ..................................................................................................28
4. Summary and implications ..................................................................................................................................................................29
5. References ...................................................................................................................................................................................................29

1 Paper presented at the 5th International Conference on Language & Education: Sustainable Development through Multilingual Education. 19-21 October 2016, Bangkok. AuSIL = Australian Society for Indigenous Languages, Darwin, Australia. ANU = Department of Linguistics; School of Culture, History & Language; College of Asia & the Pacific; Australian National University. Canberra, Australia. UBB = Unit Bahasa & Budaya (Language & Culture Unit), Kupang, west Timor, Indonesia.
1. BACKGROUND

After many years of heavy duty analysis of the grammars, discourse and lexicons of a number of languages, along with extensive translation experience, I was surprised to see the poor quality of many Early Grade Readers (learn-to-read level, from kindergarten through Year 2-3) intended for schools in many languages.

In some cases, language-aware educators have been involved in producing many of these materials. But it turns out that many educators know little about linguistics, the principles of meaning-based translation, local patterns of narrative discourse, or cross-cultural mismatches in the semantics.

In other cases, well-intended professional linguists (several of them friends of mine) also produced poor quality materials. It turns out that many people with MAs and PhDs in linguistics, have had little or no training in translation principles, orthography design, discourse analysis, language-in-education issues, or cross-cultural communication.

In every case, the usefulness of their materials can be improved by attending to a few very common issues. Many of these issues are addressed and illustrated in this paper.

1.1 High-level assumptions backed up by research

Many countries, societies and local communities are multilingual. Yet many theories of education and teaching degrees are developed in effectively monolingual countries, in which the critically important questions of the role of language in education have never really been asked or addressed. While the growing body of research by those who do ask about language-in-education considers issues of policy, language-of-instruction, curriculum, language of literacy, and L1 as a subject to be studied, my focus in this paper is on producing quality early grade reading materials in the L1—the primary language spoken daily in the home and community—for enabling the most number of children possible to learn to read for meaning. Quality early grade reading materials are an essential foundation to sustainable development.²

- **Reading for meaning** is the foundational skill to learning in all formal education. (Gove & Cvelich, 2011)
- Knowing how to read with understanding is **essential for continuing on to higher levels** of education. (Gove & Cvelich, 2011)
- **Learning to read** for the first time is easiest and most effective through a language one knows well, specifically the L1. (Dutcher, 1995; Grimes, 2009; Pinnock, 2009a,b; SIL, 2008, 2009; UNESCO, 2008; World Bank, 2005, 2006.)
- The skill of learning to read in one language transfers automatically to reading in other languages. Learning to read only has to happen once. (Baker, 2006; Smits et al., 2008; UNESCO, 2008)
- Well designed and well implemented programs of bilingual education result in **more children learning the official languages better**, and continuing on to higher levels of education. (Gove & Cvelich, 2011; SIL, 2008, 2009; Thomas & Collier, 1997; UNESCO, 2008; World Bank, 2005, 2006.)

² Other abbreviations used in this paper: 3s = third person singular; 3G = third person Genitive (singular or plural); 3sp = third singular Possessive; AIS = Aboriginal Interpreter Service; APP = applicative; DIST = distal (not near in time, space, or reference); EGRA = Early Grade Reading Assessment; EMBLI = Edukasaun Multilinge Bazeia ba Lian-Inan (mother tongue based multilingual education pilot project in Timor-Leste); EXIST = existential ‘there is’; G = genitive; IRR = irrealis (purpose, intent, future, hypothetical, about to); ISO = International Organization for Standardisation, which has adopted the 639-3 standard 3-letter unique identifier codes from the Ethnologue, first developed by the author’s father, Joseph E. Grimes; L1 = first language, the primary language spoken daily in the home and community; Lit: = literally; LOC = locative; LWC = language of wider communication; MuLok = Muatan Lokal, locally developed curriculum materials in Indonesia; NP = noun phrase; NOM = nominalizer; NTT = the Indonesian province of Nusa Tenggara Timur, which includes the western half of the island of Timor and surrounding islands; PL= plural; POSS = possessive; PRF = perfective aspect; PROX = proximal (near in time, space, or reference); REL = relativizer; SVO = Subject-Verb-Object (and other typologies); ¶ = paragraph break; § = see section.
1.2 Language research behind these suggestions

What I present below is not a bunch of abstract ideas that have not been tested. It is based on over 35 years of first-hand research on the grammars, lexicons, and discourse patterns of many Austronesian languages (both VSO and SVO), a few Papuan languages (SOV) and languages from other families. It is based on the experience of community checked and professionally peer-reviewed translation materials published in several languages numbering over 6,000 pages. It is based on assisting in the development of around 150 titles of Early Grade Readers in a dozen languages or so.

First-hand research has resulted in linguistic publications by the author on the phonologies, grammars, discourse, orthographies and lexicons in the following languages of eastern Indonesia and Timor-Leste (ISO codes are in [ ]): Amarasi [aaz], Ambon [abs], Baikeno [bkk], Buru [mhs], Dhao [nfa], Galolen [gal], Hawu [hvn], Kemak [kem], Kupang [mkn], Indonesian [ind], Mambae [mgm], Tetun [tet], Tetun Dili [tdt], and several others. I also speak several of the above languages with reasonable fluency.3

First-hand comparative and sociolinguistic research by the author has also been published on the languages of South Sulawesi, the languages of North Maluku, the languages spoken past and present on the island of Buru, and the languages of Nusa Tenggara and Timor-Leste.

Second-hand involvement with language data has also occurred in the capacity of supervising PhD and MA students, examining PhD theses, helping others prepare papers for linguistic conferences and academic journals, and running several grammar and discourse analysis workshops spanning nearly 30 years for participants working on data from languages throughout Maluku, Papua, Nusa Tenggara, Timor-Leste, and Australia.

1.3 Clusters of languages for MLE materials: EMBLI & UBB

In the Indonesian province of Nusa Tenggara Timur (which includes the western half of the island of Timor), we are working at the Language & Culture Unit (UBB—Unit Bahasa & Budaya), in well over 20 Austronesian and Papuan languages doing language documentation, grammars, dictionaries, translation, and local language literacy materials. To date we have published over 150 titles in around 18 languages. The aim is to develop a body of local language reading materials that can be used in government and private schools as part of the MuLok curriculum. MuLok is the acronym for Muatan Lokal. The government allows for up to 20% of the school curriculum to be developed locally. The current educational policies also make space for the use of local languages as the language of instruction in the classroom. Unfortunately, many teachers in rural areas are unaware of this, and express feelings of guilt when using the local language to facilitate the teaching-learning process (Kedoh, 2010; Grimes, 2013).

In the nation of Timor-Leste on the eastern half of the island of Timor, I have been providing input spanning several years to mother-tongue based multi-lingual education policy, design of practical orthographies, translation and discourse input in the development of Early Grade Readers in several languages in the EMBLI pilot project (Edukasaun Multilinge Bazeia ba Lian-Inan). I have been directly involved in various aspects of the production of around 70 titles in several languages, and there are many additional materials in other languages in which I have not been directly involved. The World Bank EGRA report for Timor-Leste (2010) showed that “More than 70% of students at the end of grade 1 could not read a single word of the simple text passage they were asked to read. 40% of children were not able to read a single word at the end of grade 2.” The bilingual EMBLI pilot project in government schools in three languages only began in earnest development in 2014, and is already showing impressive statistically significant improvement in children’s reading abilities, in contrast with the results from two other approaches in other government schools (Walter, 2016).

3 Those who care can see a list of academic publications at: https://scholar.google.com/citations?user=5mWNySIAAAAJ&hl=en
1.4 Discovering local discourse patterns: oral vs. written sources

Getting the higher level discourse right, particularly in translated material and also improving some native-authored stories, means that the readers and hearers can absorb and process the information in ways that they are used to, rather than struggling painfully to guess at what might be going on in a story. They are more likely to connect the dots right on how bits of the story or argument link together. They are more likely to identify what is intended to be important, or the focal lesson. They are less likely to miss things, and less likely to zero in on what is not important while losing sight of everything else. So once you have the basic building blocks in place for meaning-based translation, getting the discourse right may be THE most important thing you can do to make your material clear, natural, and useful. Descriptive (structural) linguistics tells you what the constructions are. Discourse analysis helps you understand what their functions are in context, and when to use them (or not).

1. **Type**: when trying to discover local discourse patterns, go after recording oral texts, not written, and not translated texts. Video with a good mic is even better than audio alone. Try and get high quality lighting and sound. Use a tripod if possible. In general, be suspicious of any material that starts out in written form or translated material. Too often they are not reliable for discourse analysis. Too often the structure of written material has noise from what the writer was taught in school about how stories or texts in the official language should be structured. Too often material that starts out in written form does not even illustrate common discourse patterns found in the language (such as presentational clauses, presentation of background information, tail-head linkage, repetition, use of direct and indirect quotes, parallelisms, chiasmus, etc.). This is particularly a problem for translated material.

While I have seen this several times, the most recent experience was in 2016, when I asked several speakers of one language to be ready to tell me stories later in the afternoon that I could record and study. When the time came, and the camera was on the tripod, one speaker became too shy, and declined. An older man had written out a story (even though I had specifically requested they not do that), and read it for me to record. Then another woman, who initially had not planned on saying anything, decided she would tell a story spontaneously. I also recorded that. Then the first (shy) woman decided she also had a story to tell, and we recorded that. Upon analyzing these 3 stories, we found that the written story read by the older man, showed none of the discourse features commonly found in that language and others in the region, but was consistent with what is taught about the official language in schools. By contrast, both of the oral stories told by the two women beautifully illustrated many of the common local patterns of natural discourse.

2. **Length**: to start with, go after texts that are around 30-50 sentences long. Less than 20 is often too short to show many of the patterns we are looking for. Over 100 is exhausting for doing multiple charts on the same text while learning how to analyze them.

3. **Who**: go after the best story-tellers or people who are known to be eloquent speakers. If possible, also have the narrator or another eloquent speaker do an editorial pass with you to confirm that this really is ‘good’, ‘beautiful’, or ‘sweet’ language and structure. You want the story structured and told in the Tetun/Hawu/Uab Meto/Yolŋu way, not the English/Indonesian/Portuguese way. Save the raw transcription separately, including marking all the

---

4 Published translations from children’s stories, novels, government publications, all the way up to some translations of the Bible demonstrate that native speakers can produce texts that are highly unnatural in their discourse structure and grammar, and ones that are often hard to comprehend—even though all the words may be known. Don’t fall into the trap of assuming that “Because this was done by a native speaker, therefore it must be clear and natural.” There are way too many counter examples.

5 Keep in mind that one text from one speaker gives you a glimpse of possibilities, but is not sufficient to establish discourse patterns for a whole language across all genres. The sample is simply too small! To have a higher degree of confidence for discourse analysis, translation, Early Grade Readers, grammar writing, and lexicography, one should aim to eventually build a text corpus of at least 200 pages of running text (one-and-a-half spaced, not yet interlinearised), from a variety of speakers, ages, both genders, different levels of education, and multiple genres.
pauses and false starts. For the purposes of discourse analysis, you can (but don't have to) edit out true false starts, and perhaps some other minor things, but do not edit out repetition, tail-head linkage, discourse particles that you can't gloss, or things of that sort. Do a final read-through to make sure the story flows and hangs together.

4. **Meta data:** make sure you have all the information on who told it, where, when, the context for the story, whether you have permission to use the text for this (and other) purposes, etc. Remember, some of these recorded stories could develop into excellent Early Grade Readers.

5. **Filename:** If you name it **TEXT1.DOC**, 10 years (10 days?) from now this will be useless to you and to anyone else who has to archive your valuable language material after you are not around. I recommend something like:

   **Language/ISO-YearMonthDay-Short Title-Narrator-STAGE-version.doc**. For example:
   - mhs-901022-Ena Biloro-Ety Lesnussa-RAW.doc; (RAW transcription)
   - aaz-20130928-Nahor Bani Nmaet-Roni Bani-INTER.doc (INTERLINEAR)
   - nfa-110515-Dheu Goa-C Aplugi-PARA.doc (PARAgraph breaks)
   - abs-010606-Kerusuhan-Anon-CHART-B.doc (Basic charting for discourse analysis)
   - tet-19921121-Baria Ulu-GTT-PARTICIPANT.doc (Referent Tracking issues)
   - hvn-20060424-Bee Lodo-Petrus Baky-PLOT.doc (PLOT structure; pattern of the story)

   [NOTE: be sure to copy to a new filename (different STAGE) before each new pass for a different kind of analysis. In case of a mistake, I recommend making a PDF to milestone your work at the end of each major pass.]

6. **Transcribe the story:** Accurately. It is up to you whether or not you want to put in tentative paragraph breaks at this stage. You might use RAW in the filename for this stage.

7. **Translate the story:** Make a chart with a 2-column table. Put each paragraph of the transcription of the story in its own box on the left. Put a good free translation on the right.

8. **Interlinearise the story:** Particularly if you are in the early phases of language learning. This is less critical if your proficiency is advanced. Interlinearising may also help other linguists and educators be able to help you more effectively. You might use INTER in the filename for this stage.

9. **Study the story repeatedly:** Work carefully through each text with multiple passes. Look at different things during each pass. Listen to the audio each time you work through the text. Aim for a minimum of 10 careful passes for each text. You will see new patterns emerge. You will start to get a sense of the high-level structure, flow and style of the text. You will gain fluency in the language. As you compare and contrast with other texts, you will gain additional insights.

1.5 **The benefits of a discourse-informed Front Translation in an LWC**

   The experience of UBB spanning 15 years or so shows that high quality discourse-informed materials in a language of wider communication [LWC] that attend to local and regional translation issues not only speeds up the process of drafting in Daughter Translations, but also results in significantly higher quality and more natural drafts, even from teams with minimal training. This is in contrast to each team drafting independently and directly from an Indo-European source. High quality Front Translations in Kupang [mkn] yield significantly better results than teams translating materials from either Indonesian or English. So speed, quality, and effectiveness of the output are all improved significantly.

   The experience of EMBLI is similar, albeit effectively spanning only two years so far. When extra work goes in to getting the Front Translation into good Tetun Dili [tdt] attending to many regional translation and discourse issues, the resulting Daughter Translations into other

---

* In the case of Early Grade Readers, this contrast is seen by how many changes are made in the final editing and checking stages. In the case of translation of the Christian scriptures, this has been independently verified by several outside Translation Consultants professionally accredited by Wycliffe Bible Translators.
languages is relatively quick and painless. For materials that have not gone through that process, a lot more time must be taken to make major revisions, and sometimes even restructuring in each separate language. Again, speed, quality, and effectiveness of the output are all improved significantly if translation and discourse issues are attended to first in the Front Translation.

Ben Grimes, former Program Manager for the Aboriginal Interpreter Service [AIS] in the Northern Territory of Australia reports similar results (p.c. 2015). When there is important material to be translated into several languages, they get significantly better and more effective results faster (and thus ultimately cheaper for the client), if they pay specialists to first rewrite the material into interpreter-friendly English. This interpreter-friendly English includes repackaging the discourse to be more like the discourse structure of Aboriginal languages, and avoids passives and abstract nouns, among other things (see summary in B.R. Grimes, 2015).

The use of a good Front Translation can be visualised as follows:

<table>
<thead>
<tr>
<th>Each language done independently:</th>
<th>A discourse-informed Front Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Text → Translation, Language A</td>
<td>Language J, Language A, Language B</td>
</tr>
<tr>
<td>Source Text → Translation, Language B</td>
<td>Language I</td>
</tr>
<tr>
<td>Source Text → Translation, Language C</td>
<td>Language C</td>
</tr>
<tr>
<td>Source Text → Translation, Language D</td>
<td>Language D</td>
</tr>
<tr>
<td>Source Text → Translation, Language E</td>
<td>Language G</td>
</tr>
<tr>
<td>Source Text → Translation, Language F</td>
<td>Language F</td>
</tr>
<tr>
<td>Source Text → Translation, Language G</td>
<td>Language D</td>
</tr>
<tr>
<td>Source Text → Translation, Language H</td>
<td>Language H</td>
</tr>
<tr>
<td>Source Text → Translation, Language I</td>
<td>Language G</td>
</tr>
<tr>
<td>Source Text → Translation, Language J</td>
<td>Language I</td>
</tr>
</tbody>
</table>

Results uneven; often poor quality, literal, unnatural, and takes a long time

Results uniformly better quality, more natural to the way local languages work, much faster

*Figure 1: Visualizing a good Front Translation for a cluster of languages in a region*

At UBB and EMBLI we have also found there are great benefits to improving the Front Translation with a feedback loop. Basically, as Daughter Translations are made from the Front Translation, we occasionally find better ways of addressing certain translation or discourse issues, or notice something that was not addressed at an earlier round. If it will improve the Front Translation and/or make it easier for the Daughter Translations to draft from, we revise the Front Translation, and make those improvements available to teams working in languages yet to start, as well as revisions of materials in existing languages. When the current version is out of print and needs to be reprinted, we do the revised and improved second edition based on the improved Front Translation, instead of just reprinting the first edition.

1.6 Key resources for applied discourse analysis and translation


---

A fairly accessible and practical guide to the principles of meaning-based and audience-appropriate translation is Larson (1998). Other sources include Nida & Taber (1974), Beekman & Callow (1974), and others.

2. **Discourse-informed Early Grade Readers**

When a story is well-formed and structured the local way, the readers-hearers can easily process the information and follow the story. When a story is not packaged the local way, or is foreign, or is missing important signals, or sends wrong or confusing signals, this introduces noise that *detracts from the main task of learning to read*.

Many things can add noise that distract from the basic task of learning to read. These include:

<table>
<thead>
<tr>
<th>Things that add noise:</th>
<th>Things that add noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>• spelling that does not reflect the sounds and structures of the language;</td>
<td>• Can't tell who is important;</td>
</tr>
<tr>
<td>• inconsistent spelling;</td>
<td>• Can't track participants;</td>
</tr>
<tr>
<td>• poor word breaks;</td>
<td>• Can't track time;</td>
</tr>
<tr>
<td>• awkward sentence structure;</td>
<td>• Can't track location;</td>
</tr>
<tr>
<td>• The story does not start in familiar ways;</td>
<td>• Can't track events/themes;</td>
</tr>
<tr>
<td>• The story does not close in familiar ways;</td>
<td>• Paragraphing: does not show</td>
</tr>
<tr>
<td>• Quote formula are out of place;</td>
<td>the structure of the story;</td>
</tr>
<tr>
<td>• Passives &amp; abstract nouns are used in languages that don’t have them;</td>
<td>• Background &amp; foreground</td>
</tr>
<tr>
<td></td>
<td>information are presented</td>
</tr>
<tr>
<td></td>
<td>incorrectly;</td>
</tr>
<tr>
<td></td>
<td>• Repetition not used the local</td>
</tr>
<tr>
<td></td>
<td>way;</td>
</tr>
<tr>
<td></td>
<td>• Tail–head linkage not used the</td>
</tr>
<tr>
<td></td>
<td>local way;</td>
</tr>
<tr>
<td></td>
<td>• Vocatives are incorrect;</td>
</tr>
</tbody>
</table>

**Figure 2: Examples of things that distract from learning to read for meaning**

Notice in Figure 2 above, the items in dark red are discourse-related issues.

Children learn about how good stories are formed from a very early age. Stories are told around the house. They overhear adults telling about interesting experiences at the garden, a dispute over land, a person's experiences returning from the city. Their grandparents and aunts and uncles tell folktales with cultural values during lulls at clan gatherings, village events, weddings, and funerals. They subtly begin to discern the good story-tellers from the average ones—who holds the audience in rapt attention? Who has them rolling in laughter? Who has them talking about the story over the next several days? Who has them retelling the story to others?

And they begin to notice patterns in how good stories are structured. I have seen 6-year-olds tell stories that are beautifully well-formed according to their local discourse patterns.

2.1 **A framework for cohesion and discontinuity in narrative discourse**

Givón (1990:427) and J. Grimes (1971) identify four threads of different kinds of information that are woven together to form a coherent narrative discourse. Different languages have different mechanisms for tracking each of these through a discourse. And they also have different mechanisms for marking discontinuity in these areas. Cross-linguistically, paragraphing tends to happen where there are discontinuities in these threads, where they are introduced, shifted, or reintroduced (resumed) after a gap (brought back on stage after a time off stage).

---

8 In many communities people are reporting to me that this traditional context for the intergenerational transmission of language and culture is subtly being overshadowed by the simple act of turning on the television (which uses the national/official language), in place of the traditional oral entertainment by skilled storytellers. Alerting village elders to this trend is one very simple way of combating language shift.
a) Participants (and props)
b) Time
c) Location
d) Events / Activities / Themes

I have found these four threads to also be useful in analyzing why a story does not seem to be well-formed, and how to fix it so the children can follow what is going on. An example of paragraphs being placed at shifts in these threads is illustrated below from a section of the Sarce story in the Kupang [mkn] language written by June Jacob (a Year 2 reader).

<table>
<thead>
<tr>
<th>Discontinuities</th>
<th>Kupang text</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event 2: Shift in Time = new paragraph</td>
<td>Sonde lama ju, waktu ana-ana parampuan dong ada baramen boneka di Meri pung teras ruma, ju Sarce pung mama datang bacari sang Sarce. Ais dia tanya sang itu ana-ana dong bilang, “Sarce ada di situ, ko?”</td>
<td>A little later, while the girls were playing dolls on the porch of Meri’s house, Sarce’s mother came looking for Sarce. And she asked the children, “Is Sarce there?”</td>
</tr>
<tr>
<td>Change of speaker (Participant) = new paragraph</td>
<td>Ju itu ana-ana dong manayo bilang, “Sarce sonde ada di sin, tanta. Mangkali dia ada di Mia pung ruma. Te tadi botong lia dia ada maen sikiidoka di sana.”</td>
<td>The children answered, “Sarce isn’t here, auntie. Perhaps she is at Mia’s house. Cuz (elaboration) earlier we saw her playing hopscotch there.”</td>
</tr>
<tr>
<td>Event 3: Shift of Participant; Change of Location = new paragraph</td>
<td>Ais Sarce pung mama jalan di Mia pung ruma. Sampe di sana, dia lia ana-ana parampuan dong ada baramen sikiidoka. Ju dia tanya sang dong bilang, “Sarce ada di situ, ko?”</td>
<td>So Sarce’s mother went to Mia’s house. Arriving there she saw the girls playing hopscotch. And she asked them, “Is Sarce there?”</td>
</tr>
<tr>
<td>Change of speaker (Participant) = new paragraph</td>
<td>Ju itu ana-ana dong manayo bilang, “Sarce sonde ada di sin, tanta. Mangkali dia ada di Sarloka pung ruma. Te tadi botong lia dia ada maen bamasa di sana.”</td>
<td>The children answered, “Sarce isn’t here, auntie. Perhaps she is at Sarlot’s house. Cuz earlier we saw her playing cooking there.”</td>
</tr>
</tbody>
</table>

When these things are not done (e.g. the typesetter removes paragraph breaks to save space), the teachers and children cannot follow the unfolding structure of the story. Simple things like putting paragraph breaks at change of speaker in a dialog turn out to provide significant help to the readers.

2.2 Introducing major participants

Presentational clauses (also known as ‘existential’ or ‘presentative’ clauses) are often used to introduce important participants with on-going relevance into the discourse. In many languages they often use a secondary sense of a common verb (e.g. ‘be, have, get’) or a preposition (e.g. ‘in, at’) with an existential meaning. For languages with no copula verbs, a presentational clause can be a non-verbal SUBJECT + PREDICATE. It is also quite common cross-linguistically for introducing single individuals with on-going relevance, for there to be the use of some form of the numeral ‘one’ in association with a generic category (Andrews 1985:80; Givón 1990:741-748; Payne 1997:123-125). These kinds of presentational clauses may begin a story, or may be bundled with other background information such as establishing the time and location for the whole story. For example:

(1) English: (Long ago, in a land far far away) There was a girl named YY.
(2) Kupang [mkn]: Ada satu nona, dia pung nama YY. *There was a girl, her name was YY.*
(3) Tetun Dili [tdt]: Ihga labarik ida, naran YY. *There was a child named YY.* (ihga is normally a locative preposition ‘in, at’)
(4) Galolen [gal]: Mia atabauk nehe, naran YY. *There was a child named YY.* (mia is normally a locative preposition ‘in, at’)

Grimes: Discourse-informed Early Grade Readers 11-Oct-16
There are two very common problems found in Early Grade Readers.

a. Main participants are often not introduced with presentational clauses;
b. Translators do not know what to do with English articles ‘a, the’, and often mistranslate them with a form of the numeral ‘one’—which actually signals a presentational clause.

When the storyline (also known as foreground, event line) begins with no background setting, and no presentational clause, the abruptness takes people by surprise. It feels like “you are telling us what they are doing, before you have told us who they are.” In other words, the rate of new information is too fast at the wrong part of the story. So for example:

(10) English:  
A. John and Freddy went to the shops…  
B. There were two friends, named John & Freddy. The two of them went to the shops…  

[The second (B) follows the local/regional patterns of natural discourse.]

(11) Baikeno [bkx]:  
A. Mala and Sufa were good friends.  
B. There were two children, one named (male) Mala and one named (male) Sufa. The two of them were good friends…  

[The second (B) follows the local/regional patterns of natural discourse.]

(12) Kupang [mkn]:  
A. Johnny and Pete went to play marbles…  
B. Ada dua anak, nama Anis den Pe?u. Satu hari dong dua pi maen kaleren…  

[The second (B) follows the local/regional patterns of natural discourse.]

The second common problem is when translators do not know what to do with English articles ‘a, the’, or Portuguese ‘um, a, o’, because their languages often have no such articles. In many cases I have seen these translated with a form of the numeral ‘one’—which is not unreasonable when one is translating at the word level. The problem is that this sends confusing signals in the discourse. For example:

---

* A large number of examples in this paper are taken from actual drafts and actual publications of Early Grade Readers in several languages. The publication details are not given, since the intent of this paper is to provide concrete examples of what people do and how to learn from and improve on them. There is no intent to shame people or make them feel bad.
The problem, as we saw above in examples 1-9, is that the use of ‘one’ is associated with presentational clauses—the first introduction of important participants with on-going relevance in the discourse. I have seen this pattern in several Early Grade Readers in several languages. So at a very subtle level, the discourse pattern of the translation signals that every page is introducing a new dog, and a new cat, and collectively these are going to build to something big—that never happens. Alternatively, each page is a different story about a new dog and a new cat, because ‘one’ is used in presentational clauses in discourse to introduce new participants with on-going relevance. Both teachers and children are confused—but it is so subtle that they don’t know why they are confused. The children never understand that this is intended to be a coherent story about the same dog-cat pair that track through a sequence of events. Below is a revised translation with discourse-informed changes made to fit the structures of the languages in the Timor region (and many other parts of the world).

Some languages can dispense with the definite anaphoric deictic (pointer word) ‘that’ after the two characters are established.

With these simple adjustments, the children will hear this as a coherent story about a single dog-cat pair that track through several events. Now they can learn to read for meaning at the level of the story, rather than the page.

In the languages of the Timor region, first mention of new participants should not use definite anaphoric deictics “this dog, that dog” (e.g. asu nee in Tetun Dili; itu anjing in Kupang). And tracking them through the discourse after they are introduced should not use “one dog” (e.g. asu ida in Tetun Dili; satu anjing in Kupang), except in the presentational clause where they are introduced for the first time. Paying attention to these things in Early Grade Readers is being informed by local discourse patterns and reduces the clutter so that children can focus on the primary task of learning to read for meaning.

2.3 Tracking participants

A common pattern for stories is for major participants to be named, or identified by a category (a person, a horse, a pig, a mouse) and then be tracked through that discourse using pronouns. Once the participants are clearly established, some languages only need a prefix or suffix on the verb. So there is a tendency in many languages to move down the hierarchy through an episode of a story, and restart the hierarchy when the character has been off stage for awhile and is reintroduced into the story.

---

10 For Tetun Dili examples, I use a “simplified orthography” for reasons detailed in Grimes (2011), rather than the so-called “official” INL orthography. The latter, among other things, uses letters for sounds that are not found in the language, and inconsistently uses 4 different strategies for marking double vowels, requiring each word to be looked up separately instead of one consistent pattern for all occurrences of double vowels. The official orthography introduces unexplainable distractions for teaching children to read.
Some patterns emerge in many of the languages of eastern Indonesia and Timor-Leste.

a. 3s free pronouns are only used in some languages to track main participants, never minor participants.

b. Usually only main and major participants are named.

c. Repeated use of names to track participants is considered heavy style, or very awkward.

d. Since names should be used only at first introduction in many languages, the repeated use of names to track participants can imply that there are several people with that same name in the story. This is quite confusing to adults and children alike.

Yet I notice that many children's stories from Indo-European sources tend to repeat the name of the main character on every page, and sometimes every occurrence. Pronouns are not used, or are used minimally.

(18) School English (not natural):
Johnny went to town.
Johnny went to school.
Johnny went to eat lunch.
Johnny went to the shops.
Johnny went to see his auntie.
Johnny went home.

(19) Improved English (more natural): [better for using as a Front Translation, see §1.5]
Johnny went to town.
Arriving there, he went to school.
Then he went to eat lunch.
Then he went to the shops.
Then he went to see his auntie.
Then he went home.

Names and pronouns should be used in Early Grade Readers in the same ways and frequencies and distributions that they are used in natural oral discourse.

2.4 Tracking location

Some stories seem “odd”, because they are not tracking location properly. For example:

(20) Amarasi [aaz]: Early version
Title: Peʔu prepares to go to school
p.1-11: His father asks, “Do you have notebook, lunch, drinking water, hat, etc.?”
p.12: Rarit, Peʔu ro he nfaʃu neʔ in umi, nautin in nnikaʔ nain in taas.
Then Pete had to go back to his house, because he had forgotten his (school) bag.

A few people sensed that something was not quite right with this story. So I asked a group of 44 people, “On pages 1-11, where do you think Pete and his father were?” The unanimous response, “At home.”

The problem on p.12 was, we had to get Pete away from home first, so that he could then go back home to get his school bag. The revised version tracked the location more clearly.

---

11 Drawings from early grade readers in this paper are taken from Big Books and readers in the EMBLI program. They were drawn by Dionisio da Silva Sipa and José Seixas dos Santos. They are gratefully used with permission.
Another story also ‘felt’ a little odd, and people weren't sure why. It turns out that the grammars and discourse of many Austronesian languages in central and eastern Indonesia and Timor-Leste pivot around mechanisms for indicating irrealis (unreal, not yet achieved, intended, future, purpose, hypothetical) and realis (real, realised, factual).

The problem, ‘going to the market’ is on the way, but not yet arrived (in Fillmore's case terms (1968), they are still along the Path, and have not yet reached the Goal). So in the discourse patterns of the local languages, we have to get them to the market, before they can actually buy anything. This is easily done with the use of the tail-head linkage pattern very common in local narrative discourse (see more discussion of tail-head linkage in §2.7 below).

Paying attention to tracking the location through a discourse can be used forensically to figure out why a story doesn't quite work right. Fixing it makes it a much better resource for more children learning to read for meaning.

2.5 Tracking time

Time phrases are often fronted to establish a time frame or mark shifts in the structure of a story.

(24) English: Long ago, in a land far far away, there was a girl named YY. [establish timeframe]

(25) English: Once upon a time, there was a girl named YY. [vague time; unreal]

(26) English: There was a girl named YY. One day she.... [signals shift from Setting to Storyline]

Fronted time phrases at the paragraph level are often used to signal shifts to new events as they track the movement of time through a discourse.

(27) Baikeno [bkx]: (discourse-level time phrases underlined):

Mbi neon meseʔ, au ama? nnao .... On the first day (=Monday), my father went ...
Mbi neon nuaʔ, au ama? nnao .... On the second day (=Tuesday), my father went ...
Mbi neon tenuʔ, au ama? nnao .... On the third day (Wednesday), my father went ...
Mbi neon haaʔ, au ama? nnao .... On the fourth day (Thursday), my father went ...
Mbi neon nimaʔ, au ama? nnao .... On the fifth day (Friday), my father went ...
Mbi neon neeʔ, au ama? nnao .... On the sixth day (Saturday), my father did not go ...
Mbi neno ?naek, au ama? ka nnao fa .... On the big day (Sunday), my father did not go ...

It is quite common in many languages to use a time phrase such as ‘one day’ or ‘one time’ to signal the shift from the Background Setting (establishing who, when, where, and sometimes the general state of affairs) to the Storyline (Foreground, Event line).
(28) Buru [mhs]:


(29) Kupang [mkn]: Early version
Bet de' bet pun kwo n oot kawri buk....
I and 4 of my friends are looking for firewood out bush....

(30) Kupang [mkn]: Revised version
Situ k, bet de' bet pu' kwo n oot kawri buk....
One time, I and 4 of my friends went looking for firewood out bush....

(31) Kupang [mkn]: Revised version
Situ k, bet de' bet pu' kwo n oot kawri buk....
One time, I and 4 of my friends went looking for firewood out bush....

2.6 Tracking events, themes
Actions, events and themes (what is being talked about) can also help track the development of a story. Again, where they shift helps identify the ‘chunks’ or boundaries of the higher-level parts of a story.

(33) Tetun Dili [tdt]:

Early draft: (discourse-level events underlined):
One day, I and my mother went to the market. My mother bought rice, My mother bought fish, My mother bought chilli, My mother also bought vegetables. Finally, my mother bought ginger. What might my mother be going to cook?

Many of the features already discussed above may be brought together to further improve this story.

(34) Tetun Dili [tdt]:

Revised story: (discourse-level events underlined):
One day, I and my mother went to the market. Arriving at the market, my mother bought rice. Then she bought fish. And she bought chilli. And she also bought vegetables. Finally, she bought ginger. Think carefully! Earlier my mother bought rice, fish, chilli, vegetables, and ginger. According to your thinking, what might she be going to cook?

2.7 Repetition, Tail-head linkage, Point-of-departure
Children’s Early Grade Readers need to be structured a little closer to the oral style that the children are used to hearing, rather than the stripped down and edited written style so common in Indo-European languages. Repetition is one feature that is commonly found in oral
narrative styles in the languages of eastern Indonesia and Timor-Leste. Repetition can be used to slow down the rate of presenting new information, or it can mark the Peak of a story, or it can be used to build tension or anticipation before the Problem is resolved. And there are many kinds of repetition (J. Grimes, 1975; Dooley & Levinsohn, 2001; Levisohn 2015a,b).

**Repetition** is a common feature of narrative discourse in the languages of this region (as in many places around the world). They are not a series of false starts by the narrators.

(35) Buru [mhs]: From a Buru folktale (repetition is underlined):

... Gofot naa, da tane fiat di enhelan, da hapu unet, da tane suran pa rahe. Petu da oli. Da oli eta beto rua sepo, da saki suba dli, po wela toho fi saka, pa da beta unet di haik. Da beta unet dii, pa gofot na holik. Holik wela, petu rihe reka wela dii. Da ego kau turen dii, pa da reka wela dii, fene...

From a Buru folktale (repetition is underlined):

... This land turtle [that we are talking about] planted the banana sapling, he tied a snare trap, he planted spike traps in the ground. Then he went home. After he had gone home for two days, he returned crossing the threshold [of that place] the monitor lizard had come down, and had tripped the snare. He tripped that snare and then the turtle freed him. [He] freed the lizard, then he beat the lizard. He took a club and he beat the lizard, saying....

A particular kind of repetition is known as ‘tail-head linkage’. The end or last clause of the previous sentence (‘the tail’) is repeated as the first part of the second sentence (‘the head’). This is sometimes used at paragraph breaks to bridge episodes in a narrative, but may also be used within a paragraph.

(36) Amarasi [aaz]: (tail-head linkage underlined):

Neno io, gam NB nmaet. ¶ Oras in nmate, te, in aan moen jes, kaann ee, nai? F, anhain nain noup. In ka nhan je ruum aah fa te, nhan niiraj je te, nrame, Nraem je reko-reko...

(tail-head linkage underlined):

Today, father NB died. ¶ When he died, his son named F, had already dug his grave beforehand. He did not just dig it empty (= plain dirt walls), [when] he finished digging it, he plastered it with concrete. He plastered it nicely....

The following narrative from Ambon has 19 instances of tail-head linkage in a story with 99 sentences.

(37) Ambon [abs]: (tail-head linkage underlined):

... Lulu don pigi. ¶ Lulu don samuo barankat pi ka asrama di T. Don lewat BM bawa. ¶ Don lewat BM bawa, tarus waktu don su ada di T, ada kajadian di P, di luar kota. Itu kajadian akan carita bagini:

(tail-head linkage underlined):

... So they went. ¶ So they all left going to the dormitory at T. They went by way of lower Red Rock. ¶ They went by way of lower Red Rock, and then while they were at T, there was an incident at P, outside the city. The story of that incident goes like this: ....

Earlier I showed how tail-head linkage can be used to make the story more coherent. That example is reproduced below.

(38) Baikeno [bkx]: Revised version


One day, I and my mother went to the market.


Arriving at the market, my mother bought rice.

Another kind of repetition is known in some sources as ‘reduced adverbial clauses’, and in others as ‘point-of-departure’. It is a kind of reduced tail-head linkage that can point back to a previous sentence, a previous paragraph, or a whole episode. It often marks the boundary of a new paragraph or a new event or episode.

(39) Buru [mhs]: (point-of-departure underlined):

Kain YT nake empeit, do, poson pei da keha toka dimen. Da rasa kakun rohin gam ba fasah. Ethfewen waen lagen mihat gam di tonal ethfo waraha. Da muta tu begilufuh. ¶ Sepo fi dii, A eptea na WK odo ...

(point-of-departure underlined):

Elder sibling YT’s sickness, well, his chest hurts rising up to the xiphoid process. He feels like his backbone is being cut with a sawing motion. His urine appears reddish-brown like a cuscus urinates blood. He vomits and feels dizzy. ¶ After that, A stayed here in WK for another ...

Point-of-departure can be used to help a story link together and flow better.
Different languages have different patterns for how reduced these point-of-departure clauses can be. Some would have to expand the example above to include a subject and say Bainhira tiu haree nunee, nia haksoi... “When the old man saw that, he jumped...”

2.8 Quote formula and complementizers

One set of issues I have encountered repeatedly revolves around quote formula (also known as quote margins) and complementizers, and the related issues of direct and indirect speech (§2.9). Some of these issues relate to the position or distribution of quote formula. English, like many Indo-European languages, can be quite flexible in where quote formula occur.

(42) English: Mary said, “I’m going to the market. Would you like to come too?” [at the front]
“I’m going to the market,” Mary said. “Would you like to come too?” [in the middle]
“I’m going to the market. Would you like to come too?” asked Mary. [at the end]

Notice that the normal SVO typology for English uses an unusual VS order in the quote formula in the third variation above.

In the Austronesian languages of eastern Indonesia and Timor-Leste, the overwhelming pattern I and others have found is for the quote formula to always be at the front. Never in the middle, and never at the end. For example, in Buru [mhs] in 99 narrative and non-narrative texts, 100% of quote formula are at the beginning—no exceptions.

A variation is illustrated by the Yolŋu Matha languages of northeast Arnhemland, Australia, in which the quote formula optionally brackets the quote contents, identifying the beginning and the end of direct speech.

(43) Wangurri [dhg]:
Ga bitjana garrun, “Yo, nhâmi way yolju” bitjana. [bracketing front & back]
Ga bitjana garrun, “Yo, nhâmi way yolju” [OR front only]
And saying thus, “Yes, you are an [important] person.” (he said)

However, I have found several SVO languages in which the quote formula are always at the front in natural oral texts, in which educators try to be ‘literary’ by putting the quote formula at the end in drafts of Early Grade Readers, and using a VS order. I have also seen this in some newspapers, where local journalists are forcing the language to follow Indo-European patterns. This is completely unnatural, and even ungrammatical. It is guaranteed to cause confusion among children learning to read. Don’t try to be literary for Early Grade Readers for which the primary purpose is for the children to learn to read for meaning.

(44) Tetun Dili [tdt]: Early version (artificial, ungrammatical, forcing the language)
“Tha ona, apa,” dehan Apeu.
“Yes, I have it, father,” said Pete.

(45) Tetun Dili [tdt]: Revised version (following patterns of natural discourse)
Apeu hotaan, “Tha ona, apa,”
Pete responded. “Yes, I have it, father.”

A complementizer is a kind of subordinating conjunction that introduces a complement clause. Think of a complement clause as a whole clause that functions as the object of a
transitive verb. In many languages, many verbs of *speaking*, *cognition*, and *perception* take complement clauses. (In the following examples in English, the complement clause is **bolded**, and the complementizer is ***boxed**.)

(46) English: 

She said **that** she would like to go to the market.  
He wondered **whether** he should go to the market.  
She saw **that** four people had returned from the market empty-handed.  
She told him **to** go to the market.

Paul Schachter (1985:50) observes, “A good many languages have a complementizer that is rather transparently derived from the verb meaning ‘say’.” That is certainly true for many of the Austronesian languages of eastern Indonesia and Timor-Leste. As with other languages, many of these can be used not only with speech-act verbs, but also with verbs of cognition and perception. The range of possibilities and complexity for each of the languages is much more interesting than what is presented in the limited data in the survey below, but these examples suffice to establish a pattern for the region. (In the following examples the main verb meaning ‘say’ is **underlined**, and the same word used as a complementizer is ***boxed**.)

(47) Kupang [mkn]: Tane bilan dia mau iko bantu oraŋ susa di pulo Semau.  
Nathaniel **said** he wanted to go along and help the people in difficulty on the island of Semau.

(48) Tane kasi tau **bilan**, “Beta mau iko bantu oraŋ susa di Semau.”  
Nathaniel told [us] **saying** “I want to to go along and help the people in difficulty on Semau.”

(49) Tane su deŋar **bilan** oraŋ bam-banya su tarima ame bantuan di Semau.  
Nathaniel has heard **that** many people have received assistance on Semau.

(50) Tane sonde tau **bilan** oraŋ bam-banya su tarima ame bantuan di Semau.  
Nathaniel does not know **that** many people have received assistance on Semau.

(51) Tane sonde tau **kalo** oraŋ bam-banya su tarima ame bantuan di Semau.  
Nathaniel does not know **whether** many people have received assistance on Semau.

(52) Tetun Dili [tdt]: Entaun katuas dehan, “Hau nee, liurai nia atan.”  
So then the old [respected] man **said**, “I am the king’s servant.”

(53) Nia hataan dehan, “O la bele lori hau nia oan baa iha nebaa.”  
He answered **saying**, “You may not take my son there.”

(54) Bainhira nia rona dehan, ema kaer tiha ona nia sobrinhu.  
When he heard **that** people had already arrested his nephew…

(55) Ita hotu hatene dehan, Nai Maromak la rona ema aat...  
We all know **that**, God does not listen to evil people...

[John 9:31]

(56) Amarasi [aaz]: Rarit in nak, “Karu hi mtoup au haan totis feʔe na te, au ?toit he hi...”  
Then he **said**, “If you accept my earlier request, [then] I ask that you....”

(57) Rarit uisf ee nataan ee **mnak**, “Nansaa? am es ho mmoe? on naan?”  
Then the king asked him **saying**, “Why did you do that?”

(58) Oras in nneen **nak**, akonot naan gubernur ate, in nhunun naheer-heer am **nak**, “Hoe!  
When he heard **that** the passer-by was the governor, he shouted out loudly **saying**, “Hey!...

(59) In ka nroim fa he atoni? nahin **nak** in etare naan.  
He didn’t want people to know **that** he was at that place.

(60) Dhao [nfa]: Haia naŋu peka na, “Jaʔa keʔa boe naŋu atu mia!”  
Then she **said**, “I don’t know where he is!”

(61) Haia na manangi peka na, “Ana. Hiak ju jaʔa kinu ciki.”  
Then he requested **saying**, “Child. Give me a bit to drink.”

(62) Loɗo naŋu tadani peka na, aʃu ɲaŋaʔa le atu naí...  
When he heard **that** there was already food available there [remote]...

(63) Buru [mhs]: Tu riŋe fen, “Amaro! Eta kimi musik yoko, do, kimi sali penik fi di ya naŋ huma.”  
And he **said**, “Gentlemen! If you favour me, well, come into my house for a bit.”
An observation I have made from many Early Grade Readers translated from Indo-European languages is that the draft translations often use **no complementizers**, when in fact the use of complementizers in these languages is the norm for speech-act verbs, and also often with verbs of cognition and perception. The absence of these complementizers is not natural, even though speakers educated in and influenced by the official language may accept it.

Buru [mhs] also has an interesting twist on the use of quote formula. When a narrative has clearly established two main characters in conversation, a bare-bones *fene/fen* can be used by itself to simply signal “change of speaker” without overtly identifying the participant.

For Early Grade Readers, if this pattern is found in a language, it needs to be tested with appropriately aged children to see if they can correctly identify and track the participants. If not, the participants should be identified explicitly, and this pattern should probably not be used until Years 3-4 and above.

A third pattern is also illustrated from the same Buru folktale above. Direct speech may be used to make the story more vivid by having **no quote formula**. This pattern is called **drama**.

As with the pattern for a stripped down “change of speaker” above, for Early Grade Readers, if this pattern of drama is found in a language, it needs to be tested with appropriately aged children to see if they can correctly identify and track the participants of who is speaking. If not, the participants should be identified explicitly, and this drama pattern should probably not be used until Years 4-5 and above.

### 2.9 Direct vs. Indirect speech

Another common difference between stories in Indo-European languages and the languages of eastern Indonesia and Timor-Leste is the prevalence of the use of indirect speech—or not. While the differences between the two can be subtle, often merely a difference in prosodic features such as intonation, one common diagnostic (telltale) difference is a contrast in pronoun. See Noonan (1985:111ff) and Givón (1990:515-555) for more discussion.
The differences are often most subtle with the use of third person pronouns.

Some languages never use indirect speech. Some use them only rarely. Indo-European languages tend to use indirect speech quite frequently. So Early Grade Readers in languages that use them rarely need to convert indirect speech to direct speech to make them natural.

Stylistically direct speech can make a story more vivid. But in the many languages of the world that do not use indirect speech, or use it very little, converting it to direct speech can also make the story understandable. In the Timor region, Early Grade Readers that use indirect speech should be treated with a high degree of suspicion.

2.10 Discourse openings

Some discourse openings are simply a non-verbal oral title, identifying the main characters. Many folktales follow this pattern.

Other discourse openings introduce the story as an intentional narrative. These discourse openings may be structured as verbal or non-verbal clauses.
Kode. Agora au naha konta kaet-keen en la berlaki...
good now I want tell a little about bridewealth
Good [+discourse shift]. Now I'm going to talk briefly about bridewealth...

A third common pattern for discourse openings is to make explicit high level evidentials. What right or authority does the narrator have for telling this story? In many cultures, stories are like property that can be owned. Some are communal, in the public domain, or owned by everyone. But telling a story that one does not have the social rights to can be risky. People understand that this can result in sickness or death for the narrator or their relatives; it can result in crop failure or failure on the hunt (Grimes 1991, ch. 23). There can be varying degrees of elaboration. Each language can have its own hierarchy of authority/credibility/reliability related to the telling of a story, and there are often common patterns shared in a region. This is all prior to the actual story beginning (eventline, storyline), but can shade into the Setting.

Buru [mhs]:
Yako la dohik ro-oin fi di geba Blanda-ro
Is narrate-APP DUP-small LOC DIST person Dutch-PL
Epte na Wae Katin.
sit PROX water pandanus
Endohin naa, yako caan fi di ya naa jina,
NOM-narrate-G PROX Is hear LOC DIST Is ISP mother
tahin-tina, R-tina.
avoid-G-female teknonym R-female teknonym
I'm going to tell-the-story a little bit about the Dutch people living here in [the village of] Wae Katin. This story I heard from my mother whose female (taboo) teknonym was R's mother.

Buru [mhs]:
Yako kita iyer dli tu naan raman emsikan.
Is see thing-PL DIST with IsP eye-G REF
Ya dapakor fi di geba dikat sa moo.
Is get-APP-PL LOC DIST person different one NEG
I saw these things with my own eyes. I did not get them from anybody else.

For many Early Grade Readers, the folktale pattern of simply giving an oral title identifying the one or two main proponents, followed by a presentational clause (§2.2) can be sufficient to follow local patterns.

Baikeno [bkx]:
Title: Na? Mala nok na? Sufa
(male name) Mala and (male) Sufa
There were two children, one named (male) Mala, and the other named (male) Sufa.
The two of them were good friends.

Tetun Dili [tdt]:
Draft title A: Hamoos duut
Draft title A: Weeding the garden (lit: cleaning the grass) [Artificial, ungrammatical, no subject]
The story begins A: Apeu ho nia amaa baa toos atu hamoos duut. Apeu and his mother went to the garden to clean/weed the grass.

Revised title B: Apeu ajuda nia amaa hamoos duut [new title]
Revised title B: Pete helps his mother weed the garden [The second (B) follows local/regional patterns and is more grammatical.]

Grimes: Discourse-informed Early Grade Readers 11-Oct-16
The story begins B (revised): *Iha labarik ida, naran Apeu.*  
*Iha loron ida, nia ho nia amaa baa toos atu hamoos duut.*  
There was a child named Apeu. [*Presentational clause (§2.2)]*  
One day he and his mother went to the garden to clean/weed the grass.

Getting the story started using local patterns in the Front Translation (§1.5) can not only make it easier to translate into local languages, but also sends the right signals to the children so they can follow the development of the story.

### 2.11 Discourse closings

Storytellers often send signals that they are shifting to wrap it up, or declaring that the story is finished.

In oral discourse, sometimes the Closing is a recap or restatement of the opening.

(87) Amarasi [aaz]: **Opening:**  
*Neno ia, aam NB nmaet. …. day PROX father (name) 3-die Today, father NB died. ….  

(88) **Closing** (42 clauses later):  
*Aam NB nmaet, in raisn ii aʔtet-tetaʔ kuun*  
father NB 3die 3s matter-PL PROX DUP-different alone-3sG  
Father NB died, and his issues (relating to his death) are entirely unique.

Where the high level Opening and Closing are identical or nearly the same, this is sometimes called a ‘sandwich structure’ in discourse studies.

It is also quite common for there to be two to three formulaic Closings in a language that signal the speaker is done speaking. Some of these can be combined. Everyone recognises these signals, and children need to learn these common Closings (also known as a discourse ‘Finish’).

(89) Amarasi [aaz]: **Closing** (42 clauses later):  
*On reʔ an, tuɹa. like REL DIST +POLITE  
Like that, sir/maam. [= that’s how the story goes/ends]  

(90) Kupang [mkn]:  
*Bagitu sa. like.that just  
That’s it.  

(91)  
*Bagitu su. like.that PRF  
That’s all.  

(92)  
*Beta puɲ carita sampe di situ sa. 1s POSS story arrive LOC there just  
My story just goes to that point.  

(93)  
*Makasi. thank.you  
Thank you.  

(94) Tetun Dili [tdt]:  
*Mak nee deît. REL DEF just  
It’s just like that.  

(95)  
*Obrigadu. thank.you  
Thank you.  

(96) Dela [row]:¹⁵  
*Baʔu a naa ena. like.that PRF  
That’s all.  

(97)  
*Duduin baʔu a naa ena. story-G like.that PRF  
That’s how the story goes.

¹⁵ Dela examples in this paper are thanks to Thersia Tamelan.
Some Early Grade Readers presented as a first person narrative should consider using Closings of this sort.

I have found a surprising number of Early Grade Readers simply hang in mid air with no closure or resolution or Finish. This seems to be particularly common where there is a list or series of things done in a sequence. Compare the high level structure of the early draft with the later revision:

(101) Tetun Dili [tdt]:

**Early draft A:**

**Title:** Lafæk nia mehi

Hau lafæk ida. Hau hamhaha loos. Hau hakarak haan ikan…

… Hau bele haan tan saida karik ee?

**Revised draft B:**

**Title:** Lafæk nia mehi

Hau nee, lafæk ida. Lafæk sira hamhaha nafatin. Iha kalan ida, hau mehi. Iha hau nia mehi laran, hau hakarak haan ikan. …

… Hau nia mehi, mak nee deit. Aogara, hau bele haan tan saida, karik ee?

A book teaching children a simple prayer did not initially close out the prayer. It ended without an ending. So as a pattern for children to learn from, it was incomplete.

(103) Kupang [mkn]:

**Early draft A:**

**Title:** Kotoŋ minya tarima kasi saŋ Tuhan

Tuhan, botọŋ minta tarima kasi, tagal Tuhan su kasi matahari ko kasi tarāŋ sang botọŋ tiap hari.

…

**Revised draft B:**

**Title:** Kotoŋ minya tarima kasi saŋ Tuhan

Tuhan, botọŋ minta tarima kasi, tagal Tuhan su kasi matahari ko kasi tarāŋ sang botọŋ tiap hari.

… Botọŋ puŋ sambayang bagitu sa, Tuhan. Makasi.

Children should learn common Closings for their own L1. And some Early Grade Readers are incomplete without a Closing. They simply hang. They are not a well formed discourse.

---

16 Both Timor-Leste and the Indonesian province of Nusa Tenggara Timur are overwhelmingly Christian in their demographics, comprised of many denominations. So not only in the private Catholic and Protestant school systems, but also in many state schools there is a strong desire among many educators to include basic religious education in a form that is acceptable to all major denominations.
2.12 Framing questions

There is much that can be said about questions in stories and questions that the teachers ask the students. And there are many types of questions, such as yes-no questions, information questions, leading questions, tag questions, rhetorical questions, etc. (see Givón 1990:782-805; Payne 1997:295-303; Saddock & Zwicky 1985:178-186). Translations of children's stories and instructions to teachers from Indo-European language sources unfortunately tend to mirror the structures of questions in the original language. However many languages in other parts of the world do not frame questions the same way. Here I touch on just two common issues.

First of all, different languages may use very different question words (Q-words) for the same social function.

(105) English: What is your name?
Spanish: ¿Cómo se llama?
Indonesian: Nama-mu siapa?

(Lit: how are you called?
Lit: name-your is who]

Secondly on the position of the Q-word, there are two very common patterns for information questions. Indo-European languages tend to follow the pattern of always fronting the Q-word. In many other regions, the Q-word remains in the position of the thing being interrogated. For example:

(106) English:
What is your name?
What are you doing?
What did they eat?
Who got hit?
What did they get hit with?
Where did they go?
When did they go?
Who are you going to see?
Why are you going?
How much did it cost?

(107) Kupang [mkn]
Lu puŋ nama sapa?
Lu ada bekin apa?
Doŋ makan apa?
Sapa yaŋ kaná pukul? /
Doŋ puku saŋ sapa?
Doŋ pake apa ko puku saŋ doŋ?
Doŋ pi mana?
Kapan doŋ pi? /
Doŋ pi dari kapan?
Lu mau pi lia saŋ sapa?
Kanapa ko lu mau pi?
Dia puŋ harga barapa?

Your name is who?
You are doing what?
They are eating/ate what?
Who is it that got hit? [fronted, contrastive, passive]
They hit whom? [Object of transitive verb]
They used what to hit them?
They went where?
When did they go? /
They went from when?
You are going to see whom?
Why do you want to go?
Its price was how much?

(108) Buru [mhs]
Kae ŋaam sane?
Kae puna sapan?
Sira kaa tenik?
Sira fláli sane?
Sane fláli geba didi?
Du fláli sira tu sapan?
Sira gam doo?
Pila saq sirä iko? /
Sira iko filim beto doo?
Kae la ku lína sane?
Kae la ku iko la emkua? /
Kae la ku iko la (nei) sapan? /
Filin pila?

Your name is who?
You are doing what?
They are eating/ate what (kind)?
They hit whom? [Object]
Who hit that person? [Subject]
They beat them with what? [Instrument]
They went where?
When did they go? /
They went from when? [Lit: which night]
You are going to see whom?
You want to go why (because of what reason)? /
You want to go why (for/towards what purpose)?
Its price was how much?

Badly framed questions easily cause confusion for the children and teachers. When I see fronted Q-words in this region, that is an automatic invitation to investigate further.
Unfortunately I have seen many translated materials in government, health, justice, and education sectors that follow the Indo-European patterns of fronting the Q-words.

2.13 Connectors: use them or not?

I have noticed that many children's books in English, and also those that are original or translated into other major languages, often have no connectors (conjunctions). Perhaps some educators feel that not using connectors makes them simpler, but the reality is that this makes them highly unnatural. Nobody would tell a story that way.

There is no good reason not to use a few simple high frequency connectors in Early Grade Readers. It does not hinder the children. The sooner they begin to recognize these high frequency connectors, the better readers they will become.

The use of these connectors makes the story hang together better, both chronologically and logically. So high frequency connectors equivalent to and, then, but, after that, so (consequently), should be used in Early Grade Readers in the same ways and same frequency as they are found in simple narrative texts.

Connectors are used to link the parts of a story together, and to mark shifts at paragraph boundaries (see §2.1).

2.14 Natural RL order (logical relationships)

There are many other issues relating to the semantics of grammar and logical and chronological relationships between clauses, sentences and paragraphs, but one in particular stands out in my working with Early Grade Readers. That is the is the order of Reason-Result clauses. They are different in different languages. They are different in different language typologies. For many Austronesian languages in eastern Indonesia and Timor-Leste, the “normal” order is Result-REASON. Various conjunctions associated with foregrounded or backgrounded reasons are boxed in the examples below.

(109) English: You should come inside, because it is raining.
Since it is raining, you should come inside.
It is raining, so (result) you should come inside.

Many languages have a fixed order for Reason-Result clauses in a sentence (or within a paragraph). Some languages can have both orders, but one is usually “normal” (unmarked) and the other is highly marked, putting that clause in greater focus (often contrastive focus) in the discourse. For many Austronesian languages in eastern Indonesia and Timor-Leste, the “normal” order is Result-REASON. Various conjunctions associated with foregrounded or backgrounded reasons are boxed in the examples below.

(110) Kupang [mkn]: Moso dolo, Eada ujan.
You should] come in now, cuz (elaboration) it's raining.

(111) Moso, Eagaell ujan.
Come in, because it's raining. [asserted reason]

(112) Tetun [tet]: Keta mata7uk, tan ha?u atu kalo o kan...
You should] not be afraid, because I will make your...

(113) Foo etu oan ida ha?u lai, Car ha?u ksalan mate ona.
Give me some rice now, because I'm starving.

(114) Tetun Dili [tdt]: Hau rona ona ema barak reklama kona ba sira, kanba sira nee aat demais liu ona.
I have heard many people complaining about them, because they are so evil.

(115) Hau halo buat sira nee hotu, kanba o tuir ona hau nia hakarok.
I am doing all these things, because you have done what I wanted.

(116) Mambae [mgm]: Ura akar fliik Maria ni kase, mahee solhati fun ura ni hua mata ba seneda. Fes fliik Maria ni kase.
He liked to listen to what Maria had to say, but sometimes it made his heart unsettled, because of listening to what Maria was saying.
Sometimes when a ‘because’ conjunction is used sentence-initial, it is really functioning as the second part of a higher-level unit.

For contact creoles such as Kupang [mkn], Ambon [abs], and Tetun Dili [tdt] the patterns can get a bit muddied because of the different L1 backgrounds of second language speakers. So for example, it is quite well known around Dili that second language speakers of Tetun Dili who come from the eastern parts of Timor-Leste and speak Papuan languages as their L1 often use a different typological order (SOV rather than SVO) for sentence-level constituents, and also frame their arguments differently. So for some people, Draft A is acceptable. Dili is located within traditional Mambae-speaking areas. And while Tetun Dili [tdt] is a Tetun-based creole [tet], much of the semantics and grammar is modelled on Mambae [mgm]. For Mambae-background speakers, and those of other Austronesian languages surrounding Dili, Draft B is much better.

Below is an example of how a page in an Early Grade Reader (Year 2-3) needed to be restructured to follow the dominant local pattern of Result-Reason.

Early draft A: [order: Reason-Result]
But Sea Cucumber said, “My thinking is like this: Tomorrow Egret will stop at the seven capes [≠ land jutting into the sea]. At each of the capes he will call me. Therefore [result], tonight seven of you must first go on ahead to each of those capes…”

Revised draft B: [order: Result-Reason]
But Sea Cucumber revealed his plan saying, “Listen to this. Tonight seven of you must go on ahead to each of the seven capes. Because [reason] tomorrow Egret will stop at each of the capes. Arriving at each of the capes he will call me.”

17 For contact creoles such as Kupang [mkn], Ambon [abs], and Tetun Dili [tdt] the patterns can get a bit muddied because of the different L1 backgrounds of second language speakers. So for example, it is quite well known around Dili that second language speakers of Tetun Dili who come from the eastern parts of Timor-Leste and speak Papuan languages as their L1 often use a different typological order (SOV rather than SVO) for sentence-level constituents, and also frame their arguments differently. So for some people, Draft A is acceptable. Dili is located within traditional Mambae-speaking areas. And while Tetun Dili [tdt] is a Tetun-based creole [tet], much of the semantics and grammar is modelled on Mambae [mgm]. For Mambae-background speakers, and those of other Austronesian languages surrounding Dili, Draft B is much better.
This issue may not appear until more complexity is introduced into the stories for Year 2 readers and above.

3. RELATED TRANSLATION AND CROSS-CULTURAL ISSUES

During the course of working with Early Grade Readers in a number of languages, I have encountered a number of related issues that recur frequently enough to see them as patterns.

3.1 The problem of foreign things

Ducks, giraffes, zebras, lions, umbrellas and kites may make cute childrens’ books in the west, but they are completely foreign to children in remote villages and mountain villages of Timor and surrounding islands. Being strange and foreign, they are a distraction in Early Grade Readers, seriously detracting from the main task of learning to read for meaning. There are several options for stories that are about foreign things.

a. Don’t use those stories;

b. Substitute local things that have similar qualities and behaviour;

c. Completely rewrite the story around local things and ideas, using the foreign story as a trigger for ideas, or a jump-off to get started.

d. Save stories of this sort until after the children already know how to read reasonably well in Year 3-4.

3.2 The problem of foreign words

When children are trying to learn how to read, words that are shaped very differently to common native words in their own language are going to create noise and confusion for the children. This includes sounds and letters, CV structure (where C = consonant; V = vowel), length of words, and stress, among other things. For example, content words (nouns and verbs) in the Austronesian languages of eastern Indonesia and Timor-Leste tend to be derived from the canonical shape \((C) V (C) V (C)\).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rock</td>
<td>fatu ~ faut</td>
<td>fatu</td>
<td>huçu</td>
<td>batu</td>
<td>fatuk</td>
</tr>
<tr>
<td>moon</td>
<td>funan ~ fuunn</td>
<td>fulan</td>
<td>haru</td>
<td>bulan</td>
<td>fulan</td>
</tr>
<tr>
<td>day</td>
<td>neno ~ neon</td>
<td>lea</td>
<td>lo/do</td>
<td>hari</td>
<td>loron</td>
</tr>
<tr>
<td>thing</td>
<td>bareŋ ~ baerŋ</td>
<td>iŋi</td>
<td>maŋa</td>
<td>baraŋ</td>
<td>buat</td>
</tr>
<tr>
<td>go</td>
<td>n-nao</td>
<td>iko</td>
<td>laʔ, kako</td>
<td>pi</td>
<td>ba, laʔo</td>
</tr>
<tr>
<td>eat</td>
<td>n-bukæ</td>
<td>kaa</td>
<td>naʔa</td>
<td>makan</td>
<td>naa</td>
</tr>
<tr>
<td>die</td>
<td>n-mate ~ n-maet</td>
<td>mata</td>
<td>maŋe</td>
<td>mati</td>
<td>mate</td>
</tr>
</tbody>
</table>

By contrast, notice the very different CV structures of the following loanwords that I have found in Early Grade Readers in the region.

(130) loanwords: papagaiu kike [from Portuguese]
layang-layang kike [from Malay]
borboleta butterfly [from Portuguese]
joaninha ladybug [from Portuguese]
sombrinha umbrella [from Portuguese]
arkuiris rainbow [from Portuguese]

If stories are actually built around these key foreign loanwords that have very different shapes from the native vocabulary, as above, consider:

a. Don’t use those stories;

b. Substitute local things that have similar qualities and behaviour;

c. Completely rewrite the story around local things and ideas, using the foreign story as a trigger for ideas, or a jump-off to get started.

d. Save stories of this sort until after the children already know how to read reasonably well in Year 3-4.
3.3 The problem of semantic range of meaning and category mismatches

First, as background to the following discussion, we talk about generics and specifics. These can work for both nouns and verbs. For example:

<table>
<thead>
<tr>
<th>Generic</th>
<th>Specifics</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>tuna, mackerel, cod, barramundi, salmon, barracuda, ...</td>
</tr>
<tr>
<td>furniture</td>
<td>chair, table, bed, desk, cupboard, wardrobe, drawers, ...</td>
</tr>
<tr>
<td>vegetables</td>
<td>beans, peas, carrots, spinach, brussel sprouts, cauliflower, ...</td>
</tr>
<tr>
<td>go</td>
<td>walk, run, drive, ride, crawl, fly, ...</td>
</tr>
<tr>
<td>carry</td>
<td>shoulder carry, carry on head, carry on back, ...</td>
</tr>
</tbody>
</table>

**Figure 6: Examples of generics and specifics**

Secondly, we talk about range of meaning. Across languages, categories often do not match neatly. There is often not an exact correspondence, and not a one-to-one correspondence. For example:

(131) English: afternoon ranges from 12:01pm until about 5:30pm
Indonesian sore ranges from around 4pm until about sundown
[Dictionaries imply the two are equivalent; they do overlap, but they are not the same.]

(132) English: red can loosely range from pink to maroon, but not all speakers would accept that
Indonesian merah red-brown, (includes cows, just-baked bread, and newborn babies)
[Dictionaries imply the two are equivalent; they do overlap, but they are not the same.]

One Early Grade Reader talked about different qualities of skin on different creatures: thin, thick, smooth, rough, brown, green, etc. Then the summary statement at the end read:

(133) English: The skin of these animals is beautiful on each of their bodies.
[listed in the story: human, crab, frog, sea turtle, crocodile, earthworm]
[‘animal’ is used in a broad generic sense here, perhaps in contrast to ‘plants’]

(134) Tetun Dili [tdt]: Animal sira nia kulit furak iha sira idak-idak nia isin.
[Problem: the Portuguese loan ‘animal’ tends to imply mammals in Tetun Dili, and not include the things named in this book.]

(135) Baikeno [bkx]: Muʔt ini es-es in aʔpooʔn ii, alekot nabii in aon aa kuun.
[Problem #1: a muʔt (metathesized plural form muʔt n) tends to refer to domesticated animals around a house and farm. There are no muʔt in this story, but it is a reasonable translation of the Tetun Dili word animal.]  
[Problem #2: the more generic kaunaʔ “creature (prototypically a ‘snake’)” does not include humans, and speakers are divided as to whether it includes all of the other creatures in the book.]

Possible solution: give two categories to cover what is listed in the story: category A and category B.

Another book talked about a man feeding his animals. The pictures show fences with the animals inside.

(136) English: On the farm.
[‘farm’ is used in a broad generic sense here, not distinguishing whether it is a) a farm where things are grown and harvested, or b) a farm where animals are raised, or c) both.]

(137) Tetun Dili [tdt]: Iha toos laran.
[toos is a reasonable translation of the English word ‘farm’]  
[Problem: toos is where you grow food crops. People work very hard to keep the animals out. Luhan is an enclosure where animals are kept and fed.]

(138) Baikeno [bkx]: Mbi lele nanan.
[lele is a reasonable translation of the English word ‘farm’ and matches Tetun Dili toos.]  
[Problem: lele is where you grow food crops. People work very hard to keep the animals out. Oʔof is an enclosure where animals are kept and fed.]
One solution would be to rewrite the story using the proper terms. However the story itself focuses on the events of **which foods** the man is feeding to **which animals**, and the location is not really in focus. So we rewrote the story to by-pass the translation problem.

(139) English: An uncle feeds his animals.

   Tetun Dili [tdt] Tiu ida foo haan nia animal sira.
   Baikeno [bkx] ʔnaef es nahao in muiʔt ini.
   Kupang [mkn]   Satu om kasi makan dia puŋ binataŋ doň.

Another story was about the behaviour of different creatures. There was a **crocodile, sea turtle, dragonfly, cuscus (marsupial), fish, bird, monkey, water buffalo, and butterfly**. The English title was: **What are these animals doing?** But in translation, in many languages there is no single term that covers all of these categories of mammals, reptiles, fish, birds, and insects. One language was able to use two terms (a bit awkwardly) to cover everything. A solution for many other languages might be simply to be vague: **What are they doing?**

3.4 The problem of structural or grammatical mismatches

The mismatch between Indo-European **articles** such as ‘a, the’ with languages that only have definite anaphoric deictics (pointer words) has already been discussed and illustrated above in §2.2.

Another common area of difficulty is translating between Indo-European languages which tend to have a **copula verb** ‘to be’, and the many languages around the world that have **no copula** verb and use other structures (such as non-verbal clauses: **SUBJECT + PREDICATE**) to accomplish the same purpose.

(140) English: She is an important teacher.

   Tetun Dili [tdt] Nia nee, mestri boot. [3s DIST, teacher big]
   Baikeno [bkx] In leʔ nane, kulu ʔnaet. [3s REL DIST, teacher big]
   Kupang [mkn] Dia tu, guru bésar. [3s DIST, teacher big]
   Buru [mhs]   Riñe dii, geba eptoke haat. [3s DIST, person teach big]

There is often a many-to-one or **one-to-many** correspondence between languages. In many cases it is impossible to translate one word for one word.

(141) Buru [mhs] riha
   English come up out of the water

(142) Buru [mhs] poda
   English come over the crest of a hill or mountain,
   or where a slope transitions from quite steep to a bit flatter

(143) Kupang [mkn] mete aer
   English stay awake at night waiting for water to start coming through the pipes
   to fill up cisterns in a semi-arid climate

Words in one language may be a **grammatical phrase** in another language. Many Austronesian languages use a phrasal Head + Modifier NP for things that Indo-European languages have single words for.  

(144) English puppy

   Kupang [mkn] anjing ana [Lit: dog offspring]
   Tetun Dili [tdt] asu oan [Lit: dog offspring]
   Amarasi [aaz] aus anaʔ [Lit: dog offspring]
   Buru [mhs]   asu anan [Lit: dog offspring-G]

---

18 In Grimes (2011) I have pointed out how confusing and unfortunate it is for people to try to force Tetun Dili grammatical phrases to function as if they are single words in the orthography (e.g. asu-oan, ai-tahan), simply because they are single words in English or Portuguese. One has to already know those Indo-European languages to be able to predict how they will be spelled in Tetun Dili! This is an unhelpful and untenable approach to spelling rules, since most citizens of Timor-Leste are not proficient in an Indo-European language. And it fails to recognise the level of **phrase** in the language.
In this category of grammatical mismatches, we could also list the following:

<table>
<thead>
<tr>
<th>Indo-European languages</th>
<th>Austronesian lgs of eastern Indonesia &amp; T-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>grammatical gender in the pronouns</td>
<td>no gendered pronouns</td>
</tr>
<tr>
<td>grammatical gender in noun classes</td>
<td>simple noun classifier (or counter) systems</td>
</tr>
<tr>
<td>frequent use of several kinds of passives</td>
<td>no passives, or very limited passives; non-referential 3p as passive equivalent</td>
</tr>
<tr>
<td>tense-prominent</td>
<td>aspect-prominent (no grammatical tense)</td>
</tr>
<tr>
<td>frequent use of abstract nouns</td>
<td>no abstract nouns, or used very infrequently</td>
</tr>
</tbody>
</table>

**Figure 7: Additional grammatical mismatches**

### 3.5 The problem of conventionalized usage and cultural categories

Many translations have highly artificial and heavy phrasing, because of forcing equivalence in systems that actually function differently, such as kinship systems. Even though one could say it that way, nobody would.

Indo-European languages have built-in gender in many of their kinship terms. The Austronesian languages of eastern Indonesia and Timor-Leste handle these quite differently. Certain terms can be modified to specify gender, but normally they are not. When they are, it is a highly marked construction, often to set up a contrast.

(146) Kupang [mkn]: p.1 *Bapa Metu mau cabu ubi kayu.*  
Mr. Metu tries to pull up the cassava.

p.2 *Dia puŋ bini mau cabu ubi kayu.*  
His wife tries to pull up the cassava.

p.3 *Don puŋ ana takı-loki mau cabu ubi kayu.*  
Their son tries to pull up the cassava.

p.4 *Don puŋ ana parampuan mau cabu ubi kayu.*  
Their daughter tries to pull up the cassava.

p.5 *Don puŋ anjıng mau cabu ubi kayu....*  
Their dog tries to pull up the cassava....

(147) Tetun Dili [tdt]: p.1 *Tiu Metu atu fokit ai farina.*  
Mr. Metu tries to pull up the cassava.

p.2 *Nia ferik oan atu fokit ai farina.*  
His wife tries to pull up the cassava.

p.3 *Sira nia oan mone atu fokit ai farina.*  
Their son tries to pull up the cassava.

p.4 *Sira nia oan feto atu fokit ai farina.*  
Their daughter tries to pull up the cassava.

p.5 *Sira nia asu atu fokit ai farina....*  
Their dog tries to pull up the cassava....

In the story above, without page 3 setting up the contrast with page 4, *ana* (or *oan*) would normally be used by itself for either son or daughter. If gender must be marked at all, it is usually only at the first introduction of the participant into the story (see §2.2). But normally the name, or a conventionally gendered activity are sufficient clues (e.g. weaving [female]; cooking at a house [female]; cooking in the jungle [male]; hunting [male]; building a house [male]; turning the soil [male]; planting seeds [female], etc.)

There are different kinship systems. And when translations force all the meaning bits to be in the text, this is more often than not awkward, forced, and artificial. Some of these mismatches are illustrated below.

<table>
<thead>
<tr>
<th>Broad category</th>
<th>Indo-European</th>
<th>Timor+ region (normal equivalent)</th>
<th>Rarely marked for gender</th>
<th>Meaning in Timor+ region</th>
</tr>
</thead>
<tbody>
<tr>
<td>offspring, child</td>
<td>son [Portuguese]</td>
<td><em>ana</em> [Kupang]</td>
<td><em>ana laki-laki; ana parampuan</em></td>
<td>minus 1 generation from ego or ego’s siblings; not normally marked for gender;</td>
</tr>
<tr>
<td></td>
<td>filha</td>
<td><em>oan</em> [Tetun]</td>
<td><em>oan mane; oan feto</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>liʔana</em>? [Baikeno]</td>
<td><em>liʔaan mone; liʔaan feto</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>anat</em></td>
<td><em>ananhana; anaan</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>ana</em> [Dhao]</td>
<td><em>anafina</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>ana</em> [Tetun]</td>
<td><em>ana mone; ana fi</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>asu</em></td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I have seen quite a few children's books in Tetun Dili force the use of modifiers in ways that are completely unnatural. These little bits of forced and unnatural language add little bits of distraction for children trying to learn to read and do not model natural use of the language.

(148) Tetun Dili [tdt]:  
\[
\begin{align*}
\text{oan mane} & \quad \text{‘offspring male = son’} \\
\text{oan feto} & \quad \text{‘offspring female = daughter’} \\
\text{avoo mane} & \quad \text{‘grandparent male = grandfather’} \\
\text{avoo feto} & \quad \text{‘grandmother’}
\end{align*}
\]

\[
\begin{align*}
\text{[more natural: oan ‘offspring, son, daughter’]} \\
\text{[more natural: oan ‘offspring, son, daughter’]} \\
\text{[more natural: avoo ‘grandparent’]} \\
\text{[more natural: avoo ‘grandparent’]}
\end{align*}
\]

4. SUMMARY AND IMPLICATIONS

Can the children follow what the story is about? Is the story structured and told the local way? Does it matter? Yes, it does, because each issue adds another layer of interference into the mix that detracts from the children being able to stay focused on the task of learning to read for meaning.

In my experience, it takes 3-4 years of training and mentoring to get a good team aware of and routinely attending to all of the issues described in this paper. Regardless of how much classroom training people have at the beginning, there is a strong tendency to start out fairly literal to the words. And most Departments of Education do not have good access to linguistic and translation experts in the local languages who are also discourse-informed. And they often have a high turnover of personnel, so they can't always keep their trained and capable personnel.

Given all of that, it is of vital and strategic importance that at least the Front Translation in a language of wider communication be done to a high standard of excellence in translation and discourse awareness. The best people need to be given priority to work on these materials—and rework them, and rework them again. To do so is to at least give a hope that the Daughter Translations into other local languages will be of reasonably good quality, and allow significantly more children to learn to read for meaning. To fail to do so is to almost guarantee that many local language materials will be poorly done, and will leave more children confused and without a strong skill in reading—so essential to remaining in school and reaching higher levels of education.

5. REFERENCES


