Developing Tetun for an independent East Timor

Seminar at Melbourne University, 31/5/2002.

Abstract: East Timor faces an enormous linguistic challenge as it moves to independence. Indonesian, the language of state for the last 25 years, has been replaced by Tetun and Portuguese as the official languages. However Portuguese is spoken well only by a small elite, and Tetun has never been used for government and academic functions. This seminar reports on a project (sponsored by the Independent Electoral Commission and UNDP) to develop a glossary of electoral terms in Tetun.

We concluded that Tetun <u>can</u> be developed to serve in the public sphere, but found that there are a wide range of issues to face. These include limitations in several constructions that are beloved of dictionary writers elsewhere (such as passives, nominalisations, and relative clauses), the choice of language for borrowings, and East Timorese' experience of the world which made some electoral concepts seem not just incomprehensible but also unbelievable.

1. Introduction

Report on a project in Dili to develop election-related terms for Tetun, the lingua franca and newly-official language of East Timor.
Will give background on the language situation in East Timor, before talking about devising terms and dictionary definitions in Tetun.
This is not a technical talk – will touch on East Timor, lexicography, communication, sociolinguistics, grammar, and semantics. Having gotten those 'big words' out of the way, here we go...

2. Background: language situation in East Timor in 2002

Tetun and Portuguese are the official languages, Indonesian and English are working languages.

- Portuguese – pre 1975, official language, Dili residents used to be bilingual in Tetun and Portuguese; 5 / 10 / 15% speak it; people educated before 1975 (hence over 40 years) are fluent.
- Indonesian – language of government and education from Indonesian invasion in 1975 (actually ~1978 before people came down from the hills) – 1999; 63 / 75 / 80% speak it; virtually all educated people are fluent.
- English – language of UN administration from 1999 to independence in 2002; < 1% speak it.
- Tetun – lingua franca, oral, liturgical since 1981, government since UNAMET in 1999; 91% speak it; hardly used in writing.
  - o Lots of Port loans – 25% even in ordinary speech.
  - o Tetun translations – poorly understood (loans, translationese, poor understanding of source).

So: Tetun understood orally for everyday subjects, Indonesian typically used for academic discussions but not approved of anymore as a source of loans or as a future official language, and Portuguese selected as an official language and approved of as a source of loans, but spoken well by very few people. English was used as the language of UN rule.

3. The IEC/UNDP project

Goal: establish a glossary of electoral terms, complete with definitions. e.g. candidate, ballot box, preferential voting.
Method:
- initial draft from Melbourne in 4 languages
- re-work this in Dili, by myself and 3 Timorese.
Show the results.

4. Choice of terms

4.1. Source of terms

- Options: use a native Tetun word with this meaning, borrow, use a term with a wider/narrower meaning, use a compound/phrase, use morphology (but Tetun Dili has hardly any). We used most of these strategies, except for the morphology one.
Loans:
- Loan source was consistently Portuguese.
  - Tetun has long history of borrowing from Port, having Port as the standard source for loans seems to be universally accepted.
  - Indonesian loans are easier to understand for most people, but we avoided it as there is (at least in some senior circles) a strong bias against it. My co-workers were happy to accept some Indonesian loans, especially when the alternatives were unknown Portuguese ones. In the end, we used Indonesian loans for little-used words in the in-house training manuals (e.g. permanent marker) but avoided it in the glossary, which was to be widely distributed.
  - Tetun Terik loans/affixes – attractive from nationalistic point of view, and the Tetun-Terik speaker on our team was keen to use some native Tetun words (e.g. *loka laran* instead of *kabine* for voting booth), but they are not understood, except for a few well-known symbolic ones (e.g. *no* 'and', *hahuu* 'start').
- Often formal/academic Port words sound similar to their Indonesian counterparts (due to latter being from Dutch, from French?): *informasaun / informasi, demonstrasaun / demonstrasi*.
  - Can be helpful: Portuguese loans with same meaning as similar-sounding Indonesian term were readily accepted: *paralelu - paralel*
  - There are 'false friends': Portuguese loan with different meaning or connotation to similar-sounding Indonesian term is risky:
    - *propaganda* Port: 'advertising'; Indon. 'lies'!
    Sometimes the existing Tetun use of a Port loan word also conflicts with the proposed use:

- *maioria simples* Port 'simple majority (i.e. more than the others, not necessarily over 50%); Tetun *simples* means 'ordinary, without fuss, not complicated'
- *reinu* Port 'kingdom'; Tetun 'people of the kingdom, citizens'
- We tried to avoid unknown loans, but didn't succeed. Testing showed that some words were not known – e.g. *justu / adil* 'just (fair)'.
  - Sometimes we put an Indonesian loan in brackets after the Portuguese.
  - Sometimes we just accepted that the term wouldn't be known, and didn't need to be known except by those who would work in the field of elections.

Wider/narrower meaning:
- opinion poll: *husu opiniaun* 'ask opinion'
- sister: *alin feto / biin*

Compound/phrase:
- disenfranchise: *lakon direitu atu vota* 'lose the right to vote'
- constitution: *lei inan* 'mother law' / *konstituisaun* 'constitution'

4.2. Little nominalisation

English glossaries include many abstract or deverbal nouns: demonstration, coalition, opinion poll, registration of parties, nomination of candidates.

Tetun has few abstract or deverbal nouns:
- Loans from Portuguese. These tend to be higher-level and less-understood than the corresponding verbs, regardless of whether the verbs are Tetun (*nomeasaun / foti kandidatu*) or Portuguese (*votu / vota*).
- Some Tetun verbs can be used as nouns: *hanoin* 'think; opinion'; but many cannot.
- So we often chose a verb to correspond to an English/Portuguese/Indonesian noun.
  - This makes the text more intelligible to Timorese (by avoiding high-level loans), but creates problems for translators, who then need to restructure a text rather than translate phrase by phrase.
    Registration of parties will close on 15th March. ->
    Partidu bele tau naran too 15 Marsu deit. 'Parties can register only until 15 March'

  - How to structure bilingual glossaries? For Tetun-English glossary we used a verb gloss since this more accurately reflects the Tetun, but for the English-Tetun glossary we listed the English noun as well since people may search for the word under the noun.
    English-Tetun: registration of parties: *rejista (hanesan partidu); tau naran (hanesan partidu)*
    Tetun-English: *tau naran (hanesan partidu)*: register as party

4.3. Number and gender in Portuguese phrasal loans:

In Portuguese, both noun and adjective show number and gender, and the order is often adjective-noun. In Tetun, there is no number or gender inflection, and the order is strictly noun-adjective.

- Standard compounds: *primeiru ministru* – this is now a standard loan. Should the plural be *primeirus ministrus*? Should a woman be *primeira ministra*? Or should we change the order to *ministru primeiru*? We opted for keeping it with the Portuguese order since it is a fixed term, but would recommend that it be used with invariant masculine singular form, since this is the number/gender for borrowings.
- Analytic phrases: *primeiru prezidenti* – this is not a fixed phrase, so we opted to reverse the order to the standard Tetun one: *prezidenti primeiru*.
- Terms frequently used in the feminine form: Most adjectives are borrowed in masculine form; however some typically occur in feminine phrases, so that the feminine form is better known. Should these be borrowed as feminine, as masculine, or display gender agreement?: *igreja Católika / uma kreda Católika / ema Católiku / institutu katolik* (Indonesian)

In practice, the choices made on gender and number agreement, and on noun-adjective order, depend partly on the role of Portuguese in society, and partly on the extent to which people want to keep loans in Tetun distinct from their Port source. At present, I suspect (but can't prove) that people who can speak Portuguese, and people in the press, are more likely to use these Port feature than are others. If, as expected, the percentage of people who can speak Port well rises markedly, then perhaps the use of number and gender agreement, even in ad hoc phrases, will rise.

4.4. Other issues (ignore for seminar):

- plurals – Portuguese, Tetun definite *sira*, both
- how important is *nian* (*surat vota nian*)?
- Portuguese loans – form for modifier is N/V/Adj? (*surat vota/votu*)
- nominalised VO -> N O nian (*distribuisaun kadeira nian*) / N *ba* O (*distribuisaun ba kadeira*) / N *ba* O *nian* (*distribuisaun ba kadeira nian*)

5. Style of definitions

5.1. Typical definitions in English

Three typical features which are problematic in Tetun:

- **Passives**: writ (Collins Concise English Dictionary): a document under seal, issued in the name of the Crown or a court, commanding the person to whom it is addressed to do or refrain from doing some specified act.
- **Relative clauses**: voter (Collins Concise English Dictionary): a person who can or does vote.
- **Abstract and deverbal nouns**: propaganda (Collins Concise English Dictionary): the organized dissemination of information, allegations, etc., to assist or damage the cause of a government, movement, etc.

5.2. No true passive

Alternatives:
- OV – but sometimes O can be interpreted as S
- Use a generic subject – *ema* 'someone, people'
- Specify actors – but then need to know who the actors are; this can be problematic for translators unless they have access to the source writer, or know about the topic.

Ballot paper:
- English AEC: **Ballot paper**: official form on which a voter's choice(s) is recorded.
- Tetun AEC: *Formulariu ofisiál nebee votantes sira uza hodi marka sá mak sira hili.*

  Official form which the voters use to mark what they choose.
- IEC: ***Surat vota*** *nee formuláriu ofisiál nebee votante sira uza hodi marka partidu, kandidatu ka alternativu nebee mak sira hili iha eleisaun.*

  A **vote paper** is an official form which voters use to mark which party, candidate or alternative they choose in an election. [Detail added after an English-speaker said: sounds like a menu!]

5.3. Not many abstract / deverbal abstract nouns.

Situation already discussed above.

5.4. Restricted relativisation

Relativisation is only on subject (e.g. 'a person who votes') and object and time (e.g. 'the day people vote'), and marginally on location; so it can't be used for all definitions. Hence we used full sentences.

Ballot box
- Collins Cobuild: A **ballot box** is the box into which ballot papers are put after people have voted.
- IEC:

  ??kaixa nebee votante sira hatama surat vota nebee sira marka tiha ona
  ??box (into) which voters insert ballot papers which they have marked.

  *Votante sira hatama surat vota nebee sira marka tiha ona, iha **kaixa** nee.*
  Voters insert ballot papers which they have marked, into this box.

5.5. Insufficient information / experience of life contradicts what is read

Testing results
- Too short a definition. Some definitions can only be understood if you already know about the topic. (See definition of 'registration of parties' below). To help people understand a definition, it seems to help a lot to give examples, and to give reasons why something is done. This breaks down the dictionary – encyclopaedia boundary).

- Unbelievable results based on their current assumptions (apparently blocking understanding totally). This clearly challenges the 'pour water into a cup' model of understanding.
  - Two-round system – in this system, if no one candidate wins over 50%, there is a run-off election for the top two candidates; to be used for Timorese presidential elections. Comment: but if someone won a (simple) majority (i.e. more than other candidates but not over 50%), it isn't just to not declare him winner! Democracy means the person with the most votes wins.
  - Preferential voting: (when my co-developers finally understood it) we'd kill each other if we did that! (since the one who wins on first preferences can lose the election)
  - Independent Electoral Commission – but how can you be independent if the state sets you up?
  - Neutrality – how is that possible?
  - "Invisible ink" (visible only under ultraviolet lamps, and used by UNAMET to ink voters' fingers so we could know whether people had already voted) – "We heard that this was a lie by UNAMET to frighten us into not voting twice!"

5.6. Style of definition used: examples

Collins Cobuild style definition is much clearer than the traditional: gives the word in a grammatical context, and tends to give useful extra semantic information (eg typical actor, associated 'feeling').

We followed this style: full sentences (though sometimes not useful) in which the word occurs in context.

vote:
- Wordsmith: to express a choice in an election or decision by casting a vote
- Collins Cobuild: When <u>you</u> vote, you indicate your choice <u>officially</u> at a meeting or in an election, <u>for example</u> by raising your hand or writing on a piece of paper.
- IEC:
  > *Kuandu votante ida **vota** (ka foo nia **votu**), nia hili partidu ka kandidatu, ou hili dalan ida atu hakotu lia. Iha eleisaun, votante sira vota ho segredu, liu hosi surat vota. Iha reuniaun ruma, vota bele segredu ka nakloke, no bele liu hosi foti liman, ka liu meius selu-seluk tan.*
  >
  > When a voter **votes** (or gives his/her **vote**), s/he chooses (a/some) party or candidate, or chooses a way to resolve a dispute. In an election, voters vote in secrecy, via a ballot paper. In meetings, voting can be secret or open, and can be via raising hands or via other means.

Registration of parties and candidates:
- AEC: the <u>act</u> of enrolling political parties and candidates to participate in the elections, on the basis of eligibility <u>criterion</u> and *submitted* <u>signatures</u> and <u>deposits</u>.
- IEC:
  > *Kuandu partidu ida **tau naran** (ka **rejista**), nia baa hatama nia naran iha komisaun eleisaun atu bele hadau kadeira iha eleisaun. Atu tau naran, partidu nee tenki tuir kritériu nebee iha. Por ezemplu, iha rai balu nia tenki aprezenta*

lista ho asinatura hosi ema nebee apoiu nia. Iha rai balu moos, nia tenki selu osan depózitu ba komisaun eleisaun nian, hanesan pinoor. Se partidu hetan votu liu hosi mínimu nebee hametin ona, mak foin bele foti fali osan nee.

When a party **puts name** (or **registers**), it goes and enters its name <u>at the electoral commission</u> in order to be able to compete for seats in an election. In order to register, the party must fulfil the criteria that exist. <u>For example</u>, in some countries it must present a list with signatures <u>from the people who support it</u>. In some countries, too, it must pay deposit money <u>to the electoral commission</u>, <u>as a pledge</u> [something left in a pawnshop]. <u>Only if the party gets more votes than the minimum specified, can it get the money back</u>.

Valid vote:

- English AEC: a vote cast in an election, marked according to the rules.
- Tetun AEC: ***Votu nebee valido***: *Votu nebee tau iha eleisaun ida, marka tuir instrusaun nebee mak simu.*

    Vote which.is valid: Vote which (is) put in an/one election, mark according to instructions which (are) received).

- IEC:

    Surat vota *vale: Kuandu votante marka nia surat vota tuir fatin, para ema nebee sura votu bele hatene saida mak votante nee hakarak atu hili, depois votante nee hatama surat nee ba kaixa vota,* ***surat vota*** *nee* ***vale***. *Surat vota nee mak sei sura.*

    <u>valid</u> **ballot paper**: If/when a <u>voter</u> marks his/her <u>ballot paper</u> according to the place [ie in the appropriate place], <u>so that the people who count votes can know what the voter wants to choose, then the voter puts the paper in the ballot box</u>, this **ballot paper** is **valid**. <u>It is this/such ballot papers that will be counted</u>.

    Extras: voter marks ballot, local knowledge (valid = can be counted, even if it doesn't follow the rules; importance of the valid/invalid distinction = is counted / not counted)

6. Orthography

- There is no standard in use – Hull's system is provisionally used by the linguistics institute, and apparently approved by the education department, but not used by anyone else that I know of.
- Discussion – the lecturer in the group was pleased to be able to discuss it logically, and urged consistency.
- Testing – participants read it easily.
- Publication – the eduction ministry wouldn't publish it without a covering note, since it didn't follow the official spelling.

7. Conclusions

- It is possible to develop a technical vocabulary in Tetun.

- Political and short-term practical considerations can conflict – choice of Portuguese / Indonesian for loan source.
- Preconceptions can get in the way of intelligibility.
- Definitions which require knowledge of the field to be understood are inadequate. To be useful, definitions require 'extra information'.

Developing Tetun for an independent East Timor

Catharina Williams-van Klinken
cvk@webfront.net.au

Melbourne University, 31/5/2002

Independent Electoral Commission and UNDP

Goal: establish a glossary of Tetun electoral terms, complete with definitions.
Process:
- initial draft from Australian Electoral Commission in Melbourne in 4 languages
- the Tetun was re-worked in Dili, by myself and 3 Timorese colleagues
- Tetun tested for intelligibility with interested groups
- Feedback from IEC Tetun translators
- English back-translation tested for content by international electoral experts

Result:
- Monolingual dictionary of 200 terms (first-ever in Tetun?)
- Bilingual word-lists from Tetun into Portuguese, English and Indonesian, and vice-versa

Choice of terms

Loans – always Portuguese
- many similar to Indonesian (via Dutch):
    - *informasaun / informasi* 'information'
    - *paralelu / paralel* 'parallel'
- some conflict with known meanings:
    - *propaganda* Portuguese 'advertising' / Indonesian 'propaganda, lies'
    - *maioria simples* Portuguese 'simple majority' / Tetun *simples* 'not complicated, not fussy'
- some little-known
    - central: Portuguese *justu* / Indonesian *adil* 'just, fair'
    - technical: Portuguese *kuota* 'quota'
- number, gender, and word order in Portuguese loans:
    - fixed: *primeiru ministru* 'prime minister'
    - ad hoc: *prezidenti primeiru* 'first president'
    - feminine: *igreja katolika* 'Catholic church', *ema katolika/u?* 'Catholic person'

Compound/phrase:
    - 'disenfranchised' *lakon direitu atu vota* = 'lose (the) right to vote'
    - 'constitution' *konstituisaun / lei inan* 'mother law'

Wider / narrower meaning:
- o 'opinion poll' *husu opiniaun* 'ask opinion'
- o 'sister' *alin feto* 'younger sister' / *biin* 'older sister'

Abstract noun -> verb
- o 'registration' *tau naran* / *rejista* 'register'

Style of definitions

Three typical features of English definitions which are problematic in Tetun:
- **Passives**: writ = a document under seal, <u>issued</u> in the name of the Crown or a court ...
- **Relative clauses**: voter = a person who can or does vote
- **Abstract and deverbal nouns**: propaganda = the organized <u>dissemination</u> of information ...

No true passive

Alternatives:
- OV – but sometimes O can be interpreted as S
- Use a generic subject – *ema* 'someone, people'
- Specify actors

Ballot paper:
- AEC: **Ballot paper**: official form on which a voter's choice(s) is recorded.
- IEC:

      Surat vota nee formuláriu ofisiál nebee votante sira uza hodi marka partidu, kandidatu ka alternativu nebee mak sira hili iha eleisaun.

      A **vote paper** is an official form which <u>voters</u> use to mark which party, candidate or alternative <u>they</u> choose in an election.

Restricted relativisation

Relativisation on:
- subject ('a person who votes')
- object ('a ballot paper which people discard')
- time ('the day people vote'
Relativisation on location disputed, on other constituents unavailable.

Ballot box:
- Collins Cobuild: A **ballot box** is the box into which ballot papers are put after people have voted.
- IEC: *Votante sira hatama surat vota nebee sira marka tiha ona, iha **kaixa** nee.*
          Voters insert ballot papers which they have marked, into this **box**.

Testing

Insufficient information
- Too short a definition. Some definitions can only be understood if you already know about the topic. Add examples, reasons, local detail.

Not believable
- Two-round/preferential systems: "Democracy means the person with the most votes wins"
- Independent Electoral Commission: "But how can you be independent if set up by the state?"
- Neutrality: "How is that possible?"
- Block vote: "But you can only vote for one person"

Communication is clearly <u>not</u> writing on a blank sheet, or pouring water into a cup!

Examples

vote:
- Wordsmith: to express a choice in an election or decision by casting a vote
- Collins Cobuild: When <u>you</u> vote, you indicate your choice <u>officially</u> at a meeting or in an election, <u>for example</u> by raising your hand or writing on a piece of paper.
- IEC:

  *Kuandu votante ida **vota** (ka foo nia **votu**), nia hili partidu ka kandidatu, ou hili dalan ida atu hakotu lia. Iha eleisaun, votante sira vota ho segredu, liu hosi surat vota. Iha reuniaun ruma, vota bele segredu ka nakloke, no bele liu hosi foti liman, ka liu meius selu-seluk tan.*

  When a voter **votes** (or gives his/her **vote**), s/he chooses (a/some) party or candidate, or chooses a way to resolve a dispute. In an election, voters vote in secrecy, via a ballot paper. In meetings, voting can be secret or open, and can be via raising hands or via other means.

Registration of parties and candidates:
- AEC: the <u>act</u> of enrolling political parties and candidates to participate in the elections, on the basis of eligibility <u>criterion</u> and <u>submitted</u> <u>signatures</u> and <u>deposits</u>.
- IEC:

  *Kuandu partidu ida **tau naran** (ka **rejista**), nia baa hatama nia naran iha komisaun eleisaun atu bele hadau kadeira iha eleisaun. Atu tau naran, partidu nee tenki tuir kritériu nebee iha. Por ezemplu, iha rai balu nia tenki aprezenta lista ho asinatura hosi ema nebee apoiu nia. Iha rai balu moos, nia tenki selu osan depózitu ba komisaun eleisaun nian, hanesan pinoor. Se partidu hetan votu liu hosi mínimu nebee hametin ona, mak foin bele foti fali osan nee.*

  When a party **puts name** (or **registers**), it goes and enters its name <u>at the electoral commission</u> in order to be able to compete for seats in an election. In

order to register, the party must fulfil the criteria that exist. <u>For example</u>, in some countries it must present a list with signatures <u>from the people who support it</u>. In some countries, too, it must pay deposit money <u>to the electoral commission, as a pledge</u> [something left in a pawnshop]. <u>Only if the party gets more votes than the minimum specified, can it get the money back</u>.

Valid vote:
- AEC English:
  a vote cast in an election, marked according to the rules.
- AEC Tetun:
  Votu nebee valido*: Votu nebee tau iha eleisaun ida, marka tuir instrusaun nebee mak simu.*

  Vote which.is valid: Vote which (is) put in an/one election, mark according to instructions which (are) received).
- IEC:
  Surat vota vale*: Kuandu votante marka nia surat vota tuir fatin, para ema nebee sura votu bele hatene saida mak votante nee hakarak atu hili, depois votante nee hatama surat nee ba kaixa vota, **surat vota** nee **vale**. Surat vota nee mak sei sura.*

  valid ballot paper: If/when a voter marks his/her ballot paper according to the place [ie in the appropriate place], so that the people who count votes can know what the voter wants to choose, then the voter puts the paper in the ballot box, this **ballot paper** is **valid**. It is this/such ballot papers that will be counted.

Exit poll:
- AEC:
  assessment of the election results obtained by canvassing voters as they leave the polling stations
- IEC:
  husu opiniaun bainhira sai*: Kuandu ema '**husu opiniaun bainhira sai**', bainhira votante sira foin sai hosi sentru votasaun, ema husu sira hodi hatene partidu ka kandidatu nebee mak sira hili ona. Informasaun nee sira uza hodi buka hatene lailais partidu ka kandidatu nebee mak atu manaan kadeira iha eleisaun. Ema nebee husu nee laoos staf eleisaun nian ka observadór sira, no la bele obriga votante sira para hataan sira. Iha Timor, atividade ida nee la iha.*

  ask opinion when departing: If someone '**asks opinion when departing**', just as the voters leave the polling centre, people ask them which parties or candidates they have chosen. This information is used to seek to quickly know which parties or candidates will win seats in the election. The person/people who ask are not electoral staff or observers, and can't force voters to answer them. In Timor, this activity doesn't occur.