Building tools for assessing linguistic vitality

Margaret Florey

Monash University, Australia margaret.florey@arts.monash.edu.au

Issues

- Field of language endangerment studies has grown rapidly over 20 years
- Methodology has not kept pace
 - characterizations of language endangerment scenarios and speaker ability are often ad hoc
 - little common use or understanding of terminologies
- > Lack of comparability across languages

Assessing language endangerment

- Researchers draw on a range of different factors, such as:
- ➤ domains of language use
- Ianguage transmission
- ➢ size of speaker community
- ➤ range of linguistic resources, e.g.
 - speech levels, ritual language
- > grammatical restructuring
- ➤ speaker fluency

Terminological issues

Classifying speaker fluency

- strong, fully competent, or fluent speaker
- > older fluent speaker
- > younger fluent speaker
- imperfect speaker
- ▹ semi-speaker
- rememberer
- terminal speaker
- > word-inserter
- passive (and near-passive) bilingual

- Classifying speech
- communities
- healthy, strong
- shifting
- threatened
- endangered
- ➤ moribund
- > obsolescent, extinct

Terminological issues

- > How useful are these terms?
- > What do we understand by them?
- > Where is the boundary between them?
 - e.g. shifting vs. threatened vs. endangered
- > What does it mean, for example, to be a "near passive bilingual"?
- How do we compare speaker fluency in relation to grammatical innovations and emerging varieties?

Goals - testing linguistic vitality

Develop a standardised tool to

- gain an informed overview of linguistic vitality in a site based on empirical data rather than less formal tools (survey, self-reporting or observation)
- learn how linguistic ability varies in and between communities: e.g.
 - age / generations, gender, religious affiliation, special roles
- > permit comparisons of linguistic vitality between sites
- > assess intergenerational transmission of linguistic and other indigenous knowledge

Linguistic uses

- Test data provide a comparative crosslinguistic database for
- > sociolinguistic analyses of language shift and language variation
- > typological analyses of linguistic features
- > historical-comparative analyses of linguistic relatedness and language change

Community uses

Community and researchers can draw on the findings to

- > provide feedback to the wider community
- raise community awareness of language vitality and language shift / loss
- > assess language maintenance needs
- > develop appropriate language learning materials
- > request funding and support from local agencies



Four research languages

The research team has been working with four ethnolinguistic communities in central Maluku which vary across a range of features:

- > geography
 - urban, coastal, mountainous interior
- > demography
 - populations range from 200 19,000
- religion
 - both Muslim and Christian villages
- linguistic vitality
 - ranges from strong to moribund





2. Translation sentences

productive ability in TL

Ambonese Malay to TL

designed to test creative

designed to test

translation from

sentences

ability in TL

3. Discourse

total 75 translation

- 3 sets

Language vitality test

Methodology and the 3 test components

Community language workers

Test components

1. Lexical recognition

designed to test

receptive ability in the

target language (TL)

no productive language

skills are necessary

total 53 test items

first three sets test recognition of common

• 4th and 5th sets test

comprehension of

simple sentences

nouns

- 5 photo sets

Research team includes community language workers (CLWs) in each site:

- > selected in consultation with the community
- > work in apprenticeship with the linguist and with language speakers
- > trained in documentation methodologies data collection, transcription, materials production

Assessing linguistic vitality

- > Realistic assessments of linguistic vitality take time!
- > We allow about 1 hour per respondent
- > Step 1: Lexical recognition
 - approximately 30-45 minutes per respondent
- > Step 2: Translation task
 - in Lohiasapalewa, task took from 7 minutes for older fluent speakers to 25 minutes for youngest speakers (primary school age)
- > Step 3: Discourse task
 - approximately 15 minutes per respondent

Preliminary work

Linguist

- records descriptions of the lexical recognition items in the TL with fluent speakers
- Community language workers
- > map the village and number each house
- ➤ select
 - adult test population: e.g. from every 5th house
 - child test respondents from the class roll (e.g. every 4th girl, etc.)

Mapping Rutah village (Seram Island)



Sampling: minimum test population per site

Generation	Adults	Female	Male	Total
G1	50 years +	3	3	6
G2	30-50 years	3	3	6
G3	18-30 years	3	3	6
	Sub-total	9	9	18
	Young people			
G4a	Senior High Yr. 11 (SMA2) (av. age 16)	3	3	6
	Junior High Yr. 8 (SMP2) (av. age 13)	3	3	6
G4b	Primary school Yr. 4 (SD4) (av. age 9)	3	3	6
	Sub-total	9	9	18
	TOTAL	18	18	36

1. Lexical recognition

This task allows us to

- learn if/how receptive linguistic ability varies according to factors such as age, gender, religious affiliation
- determine if and in which generation transmission failure (linguistic tip) has taken or is taking place
- identify language shift
- compare and map <u>broad</u> differences in linguistic vitality between language communities
- begin to demarcate speaker and non-speaker populations

Lexical recognition methodology

- First photo set is displayed in front of the respondent
- In Ambonese Malay, CLW explains the content of each photo
- CLW plays taped description in target language of first item
- Respondent listens to the description and selects a photo which matches it
- Researcher notes response
- Procedure continues through 1st set, and then through next 4 sets of photos













2. Translation task

A more sophisticated tool which allows us

to

- Iearn if/how productive linguistic ability varies
- identify linguistic similarities and differences between the speakers
- analyse grammatical restructuring
- map conservative and emerging linguistic varieties

Tran	slation task methodology	
CLW	explains the translation task	
CLW	reads or plays recording of 1st test sentence in Ambonese Malay	
Respondent	attempts translation into target language	
CLW	may prompt respondent with contextual clues for a lexeme	
Respondent	if reasonably confident with the first translation set moves on to translation set 2 and then on to translation set 3	

Translation task instructions

Speaker	Ambonese Malay	English
BL	Jadi beta ada bicara dengan siapa?	So who am I speaking with?
KM	Kristofer.	Kristofer.
BL	Kristofer. Kristofer kalau katorang bilang dalam bahasa Indonesia begini, terus Kristofer bilang akan dalam bahasa Alune ya?	Kristofer. Kristofer, if we say something like this in Indonesian, then Kristofer says it in Alune, yeah?
KM	lya.	Yeah.
BL	Oke.	Okay.

Example translation task sentences

Ambonese Malay	English
Beta lapar.	"I'm hungry."
Beta pigi di kebun.	"I went to the garden."
Dia sudah pi di air.	"She has gone to the river."
Beta makan papeda dingin.	"I ate cold sago porridge."
Beta seng makan kusu.	"I don't eat cuscus."
Katong pung papa ada tidur.	"Our father is sleeping."
Jang menanggis!	"Don't cry!"
Beta pi diuntuk jual pisang.	"I went toto sell bananas." [insert name along coast in Y direction]
Dia pulang dari	"She came home from" [insert name in upwards direction]
Dong cari durian di gunung.	"They're looking for durian in the mountains."
Dia naik pohon kelapa yang tinggi.	"She climbed a tall coconut tree."
Dia keku kayu satu ikat yang besar.	"She carried a large bundle of wood on her head."
Bapak ambil kayu untuk bikin tempat tidur.	"Father fetched wood to make a bed."

Translation task

- > On the basis of our testing experience, respondents who correctly identified ≥ 66% (35/53) of the items on the recognition task were considered able to move on to task 2
- > Three sets of translation sentences, each with 25 sentences which increase in complexity
- Sentences test for productive knowledge of the lexicon and of various grammatical constructions: e.g.
 - pronoun paradigm, predicate types, possession, negation, questions, conjunctions, adjectives, deixis and spatial reference

3. Disco	urse task		[Discourse task i	nstructions	
 Respondents who complete all 3 translation sets move on to the discourse task 6 photos from the levical 	 Stories are recorded for later analysis With young children for whom the concept of the 	Ę	Speaker	Ambonese Malay	English	
 > c) photos from the toxical recognition set are used to trigger discourse > CLW displays a trigger photo and prompts the respondent to talk in the TL for a short while about the scene, or to make up a story 	 discourse task may be difficult to understand, the test may be adapted Instead of the respondent telling a story, the CLW may instead have a short conversation with the child in the TL 	E	BL	Jadi beta ada minta supaya Kristofer ada bikin cerita sedikit dengan enam gambar ini, dalam bahasa Alune.	So I'm asking Kristofer to make a short story with these six pictures, in Alune.	0



Pilot testing

Earlier testing and validation of the test methodology in three Alune villages



Pilot test results for translation task

- Indicated that speakers could be grouped according to their shared use of certain linguistic features: e.g.
 - Alune vs. Malay word order
 - Malay loanwords
 - innovative grammatical strategies (e.g. possession, spatial reference)





Sam	nple pilot test resu	Ilts for discourse task imrod (m 8 yrs)
Speaker	Alune	English
CLW	Nime, memane a 'eu etia?	Nimrod, where did you go yesterday?
NM	Depa bolate.	Cut a pop-gun.
CLW	A depa bolat 'ai sidei?	Who did you cut the pop-gun with?
NM	Anteni.	Anteni.
CLW	Ho' imi depa etia?	So where did you all cut it?
NM	Kubure.	Graveyard.
CLW	Amure mise pi be mo?	Was yours good or not?
NM	Mise.	Good.
CLW	Ele'i leu lomei, imi dunu lomai etia?	So when you came home here, where did you shoot at each other?
NM	Di Bu Zekeus rumah.	At Mr Zekeus' house.
CLW	Ale 'ai sidei?	You and who else?
NM	Anteni, Petuka.	Anteni, Petuka
CLW	Sidi danire? Sai dani moyo?	Who cried? Wasn't there someone who cried?
NM	Mo.	No.

Sam	nple pilot test resu	Ilts for discourse task limrod (m 8 yrs)
Speaker	Alune	English
CLW	Imi 'eu depa bolate, imi 'eu lo' etia? 'eu lulu lalan dia?	When you all went to cut pop-guns, where did you all go? Which road did you take?
NM	Jalan raya.	The main road.
CLW	Ele'i doma mpei, ului esa ei ono ei nie pi behe Anteni ei ono mue?	Then when you got there, did each of you make his own or did Anteni make yours?
NM	Au ono aukure sendiri.	I made mine myself.
01144		

NM	Au ono aukure sendiri.	I made mine myself.
CLW	A nte'wa bolat pende?	Do you know how to make a pop-gun?
NM	Peneka.	Already.
CLW	Leu lomei mei hena, ale kai sidei luami dunu lomaije?	When you came home to the village, you and who else fired at each other?
NM	Beta deng Teni	Me and Teni.
CLW	Imi bo'ala ei naie 'eu ndi Laturakere mina.	The lot of you ran to Laturake.
NM	Katong ikut.	We followed along.
CLW	Bei mei henare imi doma mpe Papa olas ila?	From the village, what time did you all get to the Papa?
NM	Sepuluh.	Ten.
CLW	Ele'i imi lo' etia le'we?	Then where else did you go?
NM	Lolete.	Up there.









Results — religion

- Literature suggests that maintenance of indigenous languages in central Maluku is greater in Muslim than in Christian villages for complex historical reasons
- Results demonstrate that this pattern no longer holds: language shift is now a serious issue across religious boundaries
- Christian Allang and Muslim Rutah have been following the same trajectory of language shift for more than forty years
- > 25 years later, Muslim Tulehu is shows the same trajectory of tip from high school to primary school aged children
- > Language shift has also commenced among high school and primary school children in Christian Lohiasapalewa
- In several Christian and Muslim villages (e.g. Allang, Rutah, Soahuku) indigenous languages are no longer being transmitted and there are very few speakers older than 40 years

Results — indigenous knowledge

- Little traditional ecological knowledge has been retained in coastal villages or in the city. Results indicate loss of knowledge relating to:
- > traditional implements such as
 - fishing spear, fish trap, harvesting tools
- ▹ botany
 - plant species such as nutmeg (*Myristica fragrans*), betel nut palm (*Areca catechu*), taro (*Colocasia esculenta*)
- sago technology
 - sago processing techniques, tools, cooking methods, serving implements

PC	oten	tial 1	task	2 r	esp	ond	ents	\$
	Lohiasa	apalewa	Tu	ehu	Ru	tah	All	lang
	Female	Male	Female	Male	Female	Male	Female	Male
Primary grade 4	74	88	50	51	27	16	30	28
Junior High 2	93	82	89	84	31	24	21	29
Senior High 2	87	94	84	74	31	35	39	36
18-30	95	95	96	94	47	35	50	48
30-50	94	99	97	94	67	65	62	69
50+	99	95	97	98	92	89	90	89
	avera	age % sco	ores by g	ender an	d village/	language		
	Primary grade 4 Junior High 2 Senior High 2 18-30 30-50 50+	LohiasFemalePrimary grade 4Primary grade 4Junior High 293Senior High 28718-309530-509450+99avera	LohiasapalewaFemaleMalePrimary grade 47488Junior High 29382Senior High 2879418-30959530-50949950+9995average % scor	Image: Contract tease Lohiasapalewa Tul End Male Female Primary grade 4 74 88 50 Junior High 2 93 82 89 Senior High 2 87 94 84 18-30 95 95 96 30-50 94 99 97 50+ 99 95 97 average % scores by g	Image: Logit constraints Lohiasapalewa Tuletu Image: Logit constraints Female Male Female Male Primary grade 4 74 88 50 51 Junior High 2 93 82 89 84 Senior High 2 87 94 84 74 18-30 95 95 96 94 30-50 94 99 97 94 50+ 99 95 97 98 average % scores by gender and	Image: Construction of the second s	Image: Constraint of the image of the image. Image of the image of th	Image: Construction of the construction of

Translation task

In Allang and Rutah, not all participants who scored $\geq 66\%$ were willing or able to take the translation task. Of those who qualified, the following undertook the translation task.

	Allang	Rutah
30-50	1/2	1/4
50+	6/7	6/7
	78%	64%

Village	Age Group	Sentence set 1	Sentence set 2	Sentence set 3	Total completed 3 sentence sets
Allang	30-50	1	1	1	7/37 respondents
c	50+	6	6	6	18.9%
Rutah	30-50	1	1	1	7/36 respondents
	50+	6	6	6	19.4%
Tulehu	Primary grade 4	1	0	0	
	Junior High 2	6	6	6	26/37
	Senior High 2	5	5	5	respondents
	18-30	6	5	3	70.3%
	30-50	6	6	6	
	50+	6	6	6	
LohiaS	Primary grade 4	5	5	5	
	Junior High 2	6	6	6	36/37
	Senior High 2	6	6	6	respondents
	18-30	6	6	6	21.070
	30-50	7	7	7	
	50+	6	6	6	

Translation task respondents

		Kinstolei (m. i	995)
Speaker	sentence	Malay / Alune	English
	6	Ikan ini sedap.	This fish is tasty.
Λ		lane meje, ntelete.	This fish is tasty.
	7	Beta seng makan kusu.	I don't eat cuscus.
1		Au 'ane marele mo.	I don't eat cuscus.
3L	57	Beta sakit karena hujan.	I'm sick because of the rain.
M		Ale <u>sa'i</u> le ulane tetu ale.	You're sick because rain fell on you
BL	59	Beta bikin bakal par bawa di kebun.	I made a picnic to take to the field.
M		Au ono masate le <u>bawa</u> mei mlinu.	I made a picnic to take to the field.
L	68	Beta pi di gunung dan beta lihat kusu.	I went to the mountains and I saw a cuscus.
M		Au 'eu lete	I went up to
-		dan beta lihat kusu.	and I saw a cuscus.
N		ulate au selu marele.	the mountains I saw a cuscus.







Alune — Maria (f 1977)			
Alune	English		
Au <u>sementara</u> meje au piala manue,	At the moment, I'm raising chickens,		
bata, tulale.	hens, roosters.		
Masie bo'a bo'a mosa po <u>usaha-usaha</u> mina,	Although there aren't many yet but (you) have to make an effort,		
le, be manane.	because, for food.		
Ma'eri'e bo'a 'wat ho'o sae 'eu ono mlinu mo.	There's a lot of work, so there's no going to the garden.		
'epene, <u>cu'up</u> mo lo sabe ala,	Money, (there's) not enough to buy rice,		
po, au <u>sementara</u> meje au <u>usaha-usaha</u> mina.	but, at the moment I'm making an effort.		

Identifying paths of language shift

Paths of language shift

- 1. Gradual shift
- ~10-15% decrease in linguistic ability from one generation to the next (measured on standardised test)
- 2. Abrupt transmission failure or "tip"
- "In terms of possible routes toward language death, it would seem that a language which has been demographically highly stable for several centuries may experience a sudden "tip", after which the demographic tide flows strongly in favor of some other language." (Dorian 1981: 51)
- > ≥ 20% decrease in linguistic ability from one generation to the next (measured on standardised test)







Age group	LohiaS	Tulehu	Allang	Rutah	
50+	97%	97%	89%	90%	Group 1
30-50	96%	96%	65%	66%	Group 2
18-30	95%	95%	49%	41%	Group 3
High school	89.5%	83%	31.5%	30%	Group 4
Primary school	81%	50.5%	29%	21.5%	

2. Combining receptive and productive ability

- Combining results from receptive (lexical recognition) and productive (translation and discourse) tasks allows the researcher to refine speaker/non-speaker groups based on:
 - productive knowledge of the lexicon
 - shared innovations (word order, grammatical features, use of loans)
 - creative use of the target language over a range of genres
 - etc.

2. Combining results from receptive and productive tasks

- > distinguishes primarily between higher scoring groups (≥ 80%) where strong receptive ability need not correlate with strong productive ability
- ➤ indicates that respondents scoring ≤ 40% have neither receptive nor productive ability



utah	ng R	Tulehu	LohiaS	Age group
90% Gro	9%	97%	97%	50+
66% Gro	5%	96%	96%	30-50
41% Gro	9%	95%	95%	18-30
30% Gro	.5%	83%	89.5%	High school
21.5%	9%	50.5%	81%	Primary school

Receptive and productive ability and speaker/non-speaker groups

- > Lohia: 3 groups of speakers rather than just 1
 - Identifies restricted productive ability and grammatical innovation amongst LohiaS G4a and clusters G4a in Group 2 with Tulehu G3 and G4a
 - Identifies the limited productive ability and greater extent of grammatical innovation amongst LohiaS G4b and clusters G4b in Group 3 with Rutah and Allang G2
- > Tulehu: 3 groups of speakers rather than 2
 - Identifies a distinction in speaker ability in G3 between Tulehu and LohiaS, despite identical lexical recognition scores (95%) and clusters Tulehu G3 in Group 2 with LohiaS and Tulehu G4a
- > Rutah: 3 groups of speakers rather than 4
 - Identifies lack of productive ability amongst Rutah G3 and clusters G3 in Group 5 with Rutah and Allang G4a and G4b

Standardising identification of speaker groups

Fluent speakers (group 1)

- fluent productive ability
- ability to speak over a range of topics, genres
- little use of loanwords

Fluent innovative speakers (group 2)

- fluent productive ability
- restricted range of topics, genres
 some grammatical changes in
- comparison to fluent speaker normsome code-switching / use of
- loanwords

Semi-speakers (group 3)

- limited productive ability
- very restricted range of topics, genres
- frequent code-switching / extensive use of loopworde
- of loanwords greater grammatical change in
- comparison to fluent innovative speakers
- word order changes

Passive bilinguals (group 4)

- no productive ability in TL
- receptive ability only

Non-speakers (group 5)

neither receptive nor productive ability

in TL

Speaker groups in Maluku

Fluent speakers

- Lohia ≥ 18 years (G1, 2, 3)
- Tulehu \geq 30 years (G1, 2)
- Allang ≥ 50 years (G1)
- Rutah ≥ 50 years (G1)

Fluent innovative speakers

- Lohia 12-18 (G4a)
- Tulehu 12-30 (G3, 4a)

Semi-speakers

- Lohia ≤ 12 years (G4b)
- Allang 30-50 (G2)
- Rutah 30-50 (G2)

Passive bilinguals

- Tulehu ≤ 12 years (G4b)
- Allang 18-30 (G3)

Non-speakers

- Allang ≤ 18 years (G4a, 4b)
- Rutah ≤ 30 years (G3, 4a, 4b)

Gauging and standardising definitions of linguistic vitality

Threatened language

- (rapidly) increasing innovation
- measurable language shift
- decrease in language use among younger people (following pattern identified in other areas)

Endangered language

- restricted contexts for use among fluent speakers
- linguistic tip > 25% reduction in speaker ability (using standardised test)
- receptive ability only among youngest generation

Moribund language

- very restricted contexts for use among fluent speakers
- fluent speakers limited to oldest generation
- severe linguistic tip
- parent generation have receptive ability only

Linguistic vitality in Maluku

Threatened (Lohiasapalewa)

- G1-G3: fluent speakers
- G4a: fluent speakers of innovative varieties
- G4b: semi-speakers

Endangered (Tulehu)

- G1-G2: fluent speakers
- G3-G4a: fluent speakers of innovative varieties
- G4b: passive bilinguals

Moribund (Allang and Rutah)

- G1: fluent speakers, very restricted contexts for use
- G2: (very few) fluent speakers of innovative varieties
- G1-G3: severe tip of 49% in Rutah, 40% in Allang
 G3: passive bilinguals in Allang; non-
- G3. passive billinguals in Allang, hor speakers in Rutah
- G4a, 4b: non-speakers in Allang or Rutah

Where to from here?

- Combine measures of linguistic vitality (LVT) and traditional ecological knowledge (TEKVI)
- 2. Develop Vitality Index for Indigenous Languages and Traditional Ecological Knowledge (VIILTEK)
- 3. Pilot VIILTEK in two research sites in different language groups (Venezuela, Maluku)
- 4. Develop a template which could be applied in different ethnolinguistic groups

Endangered Maluku Languages Project

The concepts and materials in this presentation have been produced by members of the research team "Endangered Moluccan languages: eastern Indonesia and the Dutch diaspora". This project is funded by the *Hans Rausing Endangered Languages Project* and the *Australian Research Council*.

These concepts and materials may not be reproduced without permission.

For further details please contact Dr Margaret Florey Monash University, Australia Phone: +61 (0)3 9905-2237 Fax: +61 (0)3 9905-8492 Margaret.Florey@arts.monash.edu.au

http://www.arts.monash.edu.au/ling/maluku/

