



Pacific Linguistics

Electronic Publication E-6

SEALS XVI

Papers from the 16th
annual meeting of the
Southeast Asian
Linguistics Society
2006

edited by
Paul Sidwell
and
Uri Tadmor

Pacific Linguistics
Research School of Pacific and Asian Studies



The Australian National University

Pacific Linguistics E-6

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: John Bowden, Malcolm Ross and Darrell Tryon (Managing Editors),
I Wayan Arka, David Nash, Andrew Pawley, Paul Sidwell, Jane Simpson

EDITORIAL ADVISORY BOARD:

Karen Adams, *Arizona State University*
Alexander Adelaar, *University of Melbourne*
Peter Austin, *School of Oriental and African Studies*

Byron Bender, *University of Hawai'i*
Walter Bisang, *Johannes Gutenberg-Universität Mainz*

Robert Blust, *University of Hawai'i*
David Bradley, *La Trobe University*
Lyle Campbell, *University of Utah*
James Collins, *Universiti Kebangsaan Malaysia*

Bernard Comrie, *Max Planck Institute for Evolutionary Anthropology*
Soenjono Dardjowidjojo, *Universitas Atma Jaya*

Matthew Dryer, *State University of New York at Buffalo*

Jerold A. Edmondson, *University of Texas at Arlington*

Nicholas Evans, *University of Melbourne*

Margaret Florey, *Monash University*

William Foley, *University of Sydney*

Karl Franklin, *Summer Institute of Linguistics*

Charles Grimes, *Universitas Kristen Artha Wacana Kupang*

Nikolaus Himmelmann, *Ruhr-Universität Bochum*

Lillian Huang, *National Taiwan Normal University*

Bambang Kaswanti Purwo, *Universitas Atma Jaya*

Marian Klamer, *Universiteit Leiden*

Harold Koch, *The Australian National University*

Frantisek Lichtenberk, *University of Auckland*

John Lynch, *University of the South Pacific*

Patrick McConvell, *Australian Institute of Aboriginal and Torres Strait Islander Studies*

William McGregor, *Aarhus Universitet*

Ulrike Mosel, *Christian-Albrechts-Universität zu Kiel*

Claire Moyse-Faurie, *Centre National de la Recherche Scientifique*

Bernd Nothofer, *Johann Wolfgang Goethe-Universität Frankfurt am Main*

Ger Reesink, *Universiteit Leiden*

Lawrence Reid, *University of Hawai'i*

Jean-Claude Rivierre, *Centre National de la Recherche Scientifique*

Melenaite Taumoefolau, *University of Auckland*

Tasaku Tsunoda, *University of Tokyo*

John Wolff, *Cornell University*

Elizabeth Zeitoun, *Academica Sinica*

SEALS XVI

**Papers from the 16th annual meeting of
the Southeast Asian Linguistics Society
2006**

edited by
Paul Sidwell
And
Uri Tadmor



Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Pacific Linguistics E-6

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

Copyright in this edition is vested with Pacific Linguistics
First published 2008

National Library of Australia Cataloguing-in-Publication entry:

- Author: Southeast Asian Linguistics Society Meeting (16th : 2006 : Jakarta, Indonesia)
- Title: SEALS XVI [electronic resource] : papers from the 16th annual meeting of the Southeast Asian Linguistics Society, 2006 / editor, Paul Sidwell.
- ISBN: 9780858835863 (pdf)
- Notes: Bibliography.
- Subjects: Southeast Asia--Languages--Congresses.
- Other Authors/ Contributors: Sidwell, Paul J.
Australian National University. Research School of Pacific and Asian Studies. Pacific Linguistics.
- Dewey Number: 495

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
AUSTRALIA

<http://pacling.com>

Table of contents

Preface	vii
Pragmatic intricacy of Javanese figures of speech: too hard to understand? <i>Djatmika</i>	1
Negotiated grammar and distributed cognition in Allang <i>Michael Ewing</i>	5
Relative clauses in K'cho <i>Kee Shein Mang and George Bedell</i>	21
West Coast Bajau: syntactically ergative or symmetrical voice? <i>Mark T. Miller</i>	35
Cases of linguistic assimilation and etymological doublets in Javanese words of Arabic origin <i>Rochayah Machali</i>	47
Malay numeral classifiers: sketching conceptual representation from a native speaker's perspective <i>Khazriyati Salehuddin and Heather Winskel</i>	65
Fataluku as a tone language <i>Ruben Stoel</i>	75
The language choice behaviors of Paiwan people in the Taipei metropolitan area <i>Kai-lin Sung</i>	85
A descriptive study of comparative constructions in Kavalan and Amis <i>Li-May Sung and Cheng-chuen Kuo</i>	109
Grammaticalization paths of the Thai verb <i>dây</i> : a corpus-based study <i>Kiyoko Takahashi</i>	121
A comparison of the expression of simultaneity in Thai and English adults and children using short animations <i>Heather Winskel and Sudaporn Luksaneeyanawin</i>	135
Reduced relative clauses in Thai and Vietnamese <i>Natchanan Yaowapat and Amara Prasithrathsint</i>	143

Preface

The Southeast Asian Linguistics Society

History and Goals

The Southeast Asian Linguistics Society (SEALS) was conceived by Martha Ratliff and Eric Schiller in 1990 as a needed forum for the linguists who have the languages of mainland and insular Southeast Asia as their primary research focus. It is our hope that the activities of the Society will lead to:

- greater communication within this group of scholars, especially across the gap which has heretofore divided researchers of mainland Southeast Asian languages and the Austronesian languages of the Pacific;
- needed publication of descriptive, theoretical, and historical accounts of these languages, in the first instance in the form of these proceedings volumes; and
- greater awareness of these languages by non-specialist linguists, many of whom attempt to make universal and typological generalizations about the human language faculty without the important corrective which knowledge of Southeast Asian languages provides.

To these ends the Society hosts an annual international meeting as a means to support these goals. Specific projects, publications, and services beyond those of an annual meeting and the publication of the meeting proceedings will be at the discretion of the members of the Society.

Scope

The Southeast Asian Linguistics Society was founded with the idea of giving language researchers with a ‘non-northern’ Asian focus a place to share their findings and ideas. In terms of genetic affiliation, investigation into any aspect of Austroasiatic, Austronesian, Hmong-Mien, Tai-Kadai, or Tibeto-Burman languages may be relevant to our members. Although the common thread we recognize in the first instance is geographical, the boundaries of the Southeast Asian area are not clear, and we would not like to draw them rigidly. For example, scholars of languages which have a historical connection to the languages of the area but which are geographically outside and/or typologically unlike those in the Southeast Asian group would be welcome to participate in our meetings and publications, as would scholars of the typologically similar Chinese languages of southern China.

The Sixteenth Annual Meeting

The sixteenth annual meeting of the Southeast Asian Linguistics Society (SEALS XVI) was held on 20-21 September 2006 in Jakarta, Indonesia. It was planned rather late in the day, after no sponsoring institution emerged from the preceding SEALS meeting. Despite the last minute planning the meeting proved a success and many excellent papers were presented. The meeting was jointly sponsored by the Institute of Language and Culture Studies (better known by its Indonesian initials PKBB) at Atma Jaya University, and the Jakarta Field Station of the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany). The program included 36 papers by scholars from Australia, Britain, Burma, East Timor, Germany, India, Indonesia, Israel, Japan, Malaysia, the Netherlands, the Philippines, Taiwan, Thailand, and the USA. The opening plenary talk was given by Paul Jen-kuei Li of the Academia Sinica in Taiwan, who, by a fortunate yet total coincidence, celebrated his 70th birthday on that very day. The closing plenary talk was delivered by Yusrita Yanti of Bung Hatta University (Padang) and Bambang Kaswanti Purwo of Atma Jaya University.

End of the Proceedings and the birth of JSEALS

A decision was taken at the seventeenth meeting of the Southeast Asian Linguistics Society (held at the University of Maryland, USA, 31 August 2 September 2007) to discontinue the publication of proceedings in their present form. Instead, the Society has decided to launch an annual peer-reviewed journal as the principal vehicle of publication, with the title *JSEALS: Journal of the Southeast Asian Linguistics Society*. JSEALS welcomes articles that are topical, focused on linguistic (as opposed to cultural or anthropological) issues, and which further the lively debate that characterizes the annual SEALS meetings. The journal will consider for publication papers presented to the annual conference as well as any other contributions. Submission deadline is at the end of each calendar year (31 December). For further information, please consult the Society's website at: <http://www.jseals.org/>.

Paul Sidwell
Uri Tadmor
April 2008

Preface

The Southeast Asian Linguistics Society

History and Goals

The Southeast Asian Linguistics Society (SEALS) was conceived by Martha Ratliff and Eric Schiller in 1990 as a needed forum for the linguists who have the languages of mainland and insular Southeast Asia as their primary research focus. It is our hope that the activities of the Society will lead to:

- greater communication within this group of scholars, especially across the gap which has heretofore divided researchers of mainland Southeast Asian languages and the Austronesian languages of the Pacific;
- needed publication of descriptive, theoretical, and historical accounts of these languages, in the first instance in the form of these proceedings volumes; and
- greater awareness of these languages by non-specialist linguists, many of whom attempt to make universal and typological generalizations about the human language faculty without the important corrective which knowledge of Southeast Asian languages provides.

To these ends the Society hosts an annual international meeting as a means to support these goals. Specific projects, publications, and services beyond those of an annual meeting and the publication of the meeting proceedings will be at the discretion of the members of the Society.

Scope

The Southeast Asian Linguistics Society was founded with the idea of giving language researchers with a ‘non-northern’ Asian focus a place to share their findings and ideas. In terms of genetic affiliation, investigation into any aspect of Austroasiatic, Austronesian, Hmong-Mien, Tai-Kadai, or Tibeto-Burman languages may be relevant to our members. Although the common thread we recognize in the first instance is geographical, the boundaries of the Southeast Asian area are not clear, and we would not like to draw them rigidly. For example, scholars of languages which have a historical connection to the languages of the area but which are geographically outside and/or typologically unlike those in the Southeast Asian group would be welcome to participate in our meetings and publications, as would scholars of the typologically similar Chinese languages of southern China.

The Sixteenth Annual Meeting

The sixteenth annual meeting of the Southeast Asian Linguistics Society (SEALS XVI) was held on 20-21 September 2006 in Jakarta, Indonesia. It was planned rather late in the day, after no sponsoring institution emerged from the preceding SEALS meeting. Despite the last minute planning the meeting proved a success and many excellent papers were presented. The meeting was jointly sponsored by the Institute of Language and Culture Studies (better known by its Indonesian initials PKBB) at Atma Jaya University, and the Jakarta Field Station of the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany). The program included 36 papers by scholars from Australia, Britain, Burma, East Timor, Germany, India, Indonesia, Israel, Japan, Malaysia, the Netherlands, the Philippines, Taiwan, Thailand, and the USA. The opening plenary talk was given by Paul Jen-kuei Li of the Academia Sinica in Taiwan, who, by a fortunate yet total coincidence, celebrated his 70th birthday on that very day. The closing plenary talk was delivered by Yusrita Yanti of Bung Hatta University (Padang) and Bambang Kaswanti Purwo of Atma Jaya University.

End of the Proceedings and the birth of JSEALS

A decision was taken at the seventeenth meeting of the Southeast Asian Linguistics Society (held at the University of Maryland, USA, 31 August 2 September 2007) to discontinue the publication of proceedings in their present form. Instead, the Society has decided to launch an annual peer-reviewed journal as the principal vehicle of publication, with the title *JSEALS: Journal of the Southeast Asian Linguistics Society*. JSEALS welcomes articles that are topical, focused on linguistic (as opposed to cultural or anthropological) issues, and which further the lively debate that characterizes the annual SEALS meetings. The journal will consider for publication papers presented to the annual conference as well as any other contributions. Submission deadline is at the end of each calendar year (31 December). For further information, please consult the Society's website at: <http://www.jseals.org/>.

Paul Sidwell
Uri Tadmor
April 2008

PRAGMATIC INTRICACY OF JAVANESE FIGURES OF SPEECH: TOO HARD TO UNDERSTAND?

Djarmika
Sebelas Maret University
<djarmika@uns.ac.id>

Abstract

The paper discusses how to get the real intention of Javanese speakers who use figure of speech in their daily language exploitation. To be harmonious with and to honor others in their interaction, Javanese people use their language in several strategies in which one of them is that of being indirect. If a speaker thinks that to state something clearly and ‘to the point’ is thought to be potentially face-threatening to his interlocutor, then, he in such a culture will go through ‘an (pragmatic) intricate way’ to achieve the goal of his speech act. This strategy is performed by using several types of figure of speech. However, to understand the real intention of an utterance delivered by someone who uses a Javanese figure of speech, an interlocutor should share the same cultural background knowledge with him; otherwise, the meaning of such an utterance will not be obtained, and a potential misunderstanding between the communicants will take place. Javanese language facilitates as well as accommodates this pragmatic intricacy with the several types of figure of speech it has such as *wangsalan* and *sasmita*. Examples collected from any usage of the figure of speech are discussed to show how cultural understanding helps much in designing strategies and procedures in understanding the intricacy.

Introduction

Figure of speech is a verbal expression in which words are arranged in a particular way to achieve a particular effect or for a particular intention. Most of languages in the world have this kind verbal exploitation; and there are many types of figure of speech that can be found in the languages. In common there are figures of speech which are arranged in relation to sound, grammar, or meaning exploitation. Alliteration and assonance are examples of figures of speech that are built up from the exploitation of sound, and these figures of speech can be found easily in any languages. Exploitation of grammar to make up a figure results in ellipsis and other figures of speech such as asyndeton and ellipsis, whereas the examples of figures of speech arranged in relation to the meaning are irony, metaphor, and some others.

Most of languages in the world have these “common” types of figures of speech including Javanese—a language used by people who reside in Central and East Java and other areas in Indonesia who construct a Javanese society and use Javanese as their language of interaction. However, except for the common figures of speech, Javanese speakers also have their own figures of speech. These types of figure of speech are unique and only found in Javanese. Of several Javanese figures of speech, *wangsalan* and *sasmita* are two familiar figures exploited much by the Javanese speakers in their language use. These figures are not only arranged in relation to the sounds, but also in relation to the meaning.

Wangsalan and Sasmita—Unique Figures of Speech

Figures of speech are exploited much in literary works and also in daily spoken interactions for certain rhetorical intention. Similarly, *wangsalan* and *sasmita* are two Javanese figures of speech which are used in art performances (in this case Wayang performances), Javanese traditional songs (lyrics for Javanese *Gendhings*—songs going with gamelan instruments), and also in daily usage for particular rhetoric purposes such as in delivering a speech in a wedding party. The interesting fact about the exploitation of the two figures is the “circular and intricate route” to reach the real intention of a speaker in forwarding such figures of speech. In other words, to understand *wangsalan* and *sasmita* the hearer must understand first the cultural as well as the situational context nutshelling the figures. If he/she shares every cultural and situational element with the speaker, then the problem is over. He/she will catch the real intention of the usage of the figures. On the other hand, missing one element of the background knowledge leads to a misunderstanding. This complicated route is in line with Lakoff’s (1993) suggestion on how to understand metaphors. Such a complicated route for understanding *wangsalan* and *sasmita* is explained as follows.

Wangsalan is a Javanese figure of speech a little bit similar to alliteration and assonance. The difference between them is that if alliteration and assonance are figures of speech created by repeating the consonants or the vowels in multiple words, *wangsalan* is constructed by using a sounding similar syllable of a ‘hidden’ word as a hint for expressing the real intention of a message. For example, *jenang gula dik, kowe aja lali* is a *wangsalan* having two parts—*jenang gula dik* as the first part and *kowe aja lali* as the second one. What is repeated in the second part is the (hidden) meaning of the phrase *jenang gula* (‘sugar porridge’—Javanese people have a traditional candy made from melted sugar, it is called as **gulali**). The part of the word *gulali* is then used to reveal the real message of the utterance, that is *kowe aja lali* (you don’t **forget** me). From such an example, it can be suggested that a *wangsalan* is not only constructed based on the sound (*gulali* and *lali*) but also on the meaning (*jenang gula* for *gulali*). To understand what a speaker really intends to say by delivering a *wangsalan* of *jenang gula dik, kowe aja lali* an interlocutor must go through the following route.

- He should know that the name of a traditional candy made from melted sugar in Javanese is *gulali*.
- He should know that Javanese has an expression of *lali* meaning to forget.
- He then should pragmatically relate the rhyming syllable of the word *gulali* and the word *lali* to conclude that the speaker request him not to forget her/him (the speaker).

This example is classified as ‘archaic’ *wangsalan*. Such a kind of *wangsalan* is usually exploited in art performances, especially in lyrics of Javanese songs going with *gamelan* music, or exploited for rhetoric purposes such as in delivering speeches in Javanese traditional wedding. The characteristics of this type of *wangsalan* is that the speaker states in his utterance the real intention of the message (what he really wants to say). The other examples of the archaic *wangsalan* are:

- a. *klapa muda*, ('young coconut', in Javanese named 'degan')
yen *kalegan paringa apura* (if you are in a good mood, then may you forgive me)
- b. *mbok balung janur* (coconut leave's bone, in Javanese named 'sada')
paring usada nggonku nandhang wuyung (give me a cure for my falling in love to you—the hearer).

The second type of *wangsalan* is categorized in daily *wangsalan*—this type is exploited in daily conversation. As by using this type of *wangsalan* a Javanese speaker usually does not state the second part or the message he wants to deliver, a hearer must know well Javanese cultural knowledge to get understand what the speaker really wants to say. If a hearer does not acquire the cultural items contained in a *wangsalan*, then he will be blocked to get the real message or he will fail to understand the meaning of such a figure forwarded by a speaker; put in other words he will probably infringe a maxim of the conversation in responding the figure of speech (Thomas,1996). For example, to understand an expression of '*Ojo sok seneng ngrokok cendhak*' (literally means Don't often smoke a short cigarette) used by a Javanese speaker, the interlocutor should have cultural resources to *translate* the real intention of such a speaker in using the expression through the following procedures:

- He should know that the name of a short cigarette in Javanese is *tegesan*.
- He should know that Javanese has an expression of *neges-neges* meaning trying to reveal secret matter possessed by someone (this act in certain situation is considered impolite in Javanese society).
- He then should pragmatically relate the rhyming syllable of the word *tegesan* and the word *neges-neges* to conclude that if a speaker forwards him with an imperative '*Ojo sok seneng ngrokok cendhak*', it means that the speaker (usually older than him) prohibits him indirectly for his trial in revealing a hidden matter from such a speaker.

The other Javanese figure of speech is *sasmita*. This figure is exploited by a speaker to deliver a request by giving a hint (or hints) to the hearer(s). This is usually exploited in art performances such as in a puppet show. A shadow player (or '*dhalang*'—someone who plays the puppets and he is the leader of the performance), to achieve certain mood situations for the acts in the play, often requests particular *gamelan* arrangements going along the acts or the scenes. He usually executes the request indirectly by delivering a *sasmita*. In doing this, he give hints to the hearers, in this case the music crews, that he needs a particular arrangement. For example, in requesting a kind of music arrangement entitled *Gendhing Onang-Onang* for a scene in which one of the characters is named Arjuna, the puppet player can say it in '*dhasar Raden Arjuna satriya kang kondhang kaonang-onang*'. For the audience, the expression sounds describing only the main character, but for the *gamelan* (music) crew the expression gives a hint for what arrangement they should play for the scene, that is *Gendhing Onang-Onang*—a name sounding similar to a part of the expression. The obtaining of what the expression of this type suggests bases only on the similar sound between the part of the expression and the suggested name of the arrangement. Two other examples of this type is *ingkang nedheng mengeng penggalih*e for *Gendhing Laler Mengeng*; and *mugi tansah gangsar rejekine* for *Gangsaran*, and so on.

The other type of *sasmita* is revealed from not only the similar sounds but also the meaning of the expression, for example when a shadow player forwards an expression of *kadya manyar sasra bareng neba*, the real intended name of music arrangement is *Lancaran Manyar Sewu*. The *sasmita* expression states the words of *manyar* (a name of bird) and *sasra* (archaic Javanese word for ‘a thousand’). These words sound similar to the name of the arrangement; moreover, the meaning of the two words are similar to the meaning of the name such an arrangement.

The third type of *sasmita* is the most intricate. The suggested meaning is revealed by the same strategy to the second type above; however, the revealing passes a ‘farther’ way. For example, when a shadow player suggests a *sasmita* of *penggalihe sang nata kadya surya kataweng hima* meaning ‘the mood of the king is like a sun blocked by the cloud’, the *gamelan* crews should conclude first that the (long) expression can be stated in one Javanese word *remeng-remeng*, and they know that there is a Javanese music arrangement entitled *Gendhing Remeng-Remeng*. On that account, they will expectedly catch that the music arrangement requested indirectly by the leader is that of *Gendhing*. Again this expression only works for the crews (and for the audience) who share the cultural and situational contexts with the leader. For those who don’t, this expression sounds only describing the mood the character (the king)

Concluding Remarks

Sharing the cultural as well as the situational knowledge with a speaker who delivers a *wangsalan* or *sasmita* is the key to catch the real intention of the message. Missing any aspect of the cultural and/or the situational context leads to a misunderstanding or in that case the hearer will get difficulty in catching it or he or she may fails catching it, or he or she may infringe a maxim in responding to the expression within an interaction.

References

- Lakoff, G. 1993. The contemporary theory of metaphor. In A. Ortony (ed.) *Metaphor and Thought*. Cambridge: Cambridge University Press.
- Rahardjo, S. 2003. *Pepak Basa*. Surakarta: Penerbit ITA.
- Thomas, Jenny.1996. *Meaning in Interaction: An Introduction to Pragmatics*. London: Longman.

NEGOTIATED GRAMMAR AND DISTRIBUTED COGNITION IN ALLANG¹

Michael C. Ewing
University of Melbourne
<mce@unimelb.edu.au>

Grammatical knowledge is an example of distributed cognition. Distributed cognition or socially shared cognition (Hutchins 1995, Schegloff 1991, Salomon 1993) refers to the situation in which a knowledge system and the procedures for implementing such a system do not exist completely within any one individual. Rather, such knowledge is distributed socially among multiple individuals. It is also distributed differentially (different people have different amounts and kinds of knowledge), but non-exclusively (there is overlap with some aspects of knowledge being shared among different people). The implementation of distributed cognitive phenomena is thus necessarily a social, interactive activity. In the case of language, grammatical knowledge does not exist in its entirety with any single member of the speech community, nor does it exist independently of the social and communicative interaction through and for which it is formed and utilised. While the overlap in grammatical knowledge between members of a speech community tends to be quite great in the case of robust languages, the amount of grammatical overlap between speakers decreases in language loss contexts. In this paper I present evidence for the distributed nature of grammatical knowledge in a community experiencing language loss - Allang village in eastern Indonesia - by looking at a particular speech event - the production of a history of the village in the indigenous language.

Allang is a Christian village located on the south-western tip of the Hitu peninsula of Ambon Island, Maluku in Indonesia. The resident population of Allang is approximately 4,000, with many more who identify themselves as Allang living in the provincial capital, Ambon City, as well as other parts of Indonesia. Allang shares an indigenous language with the neighbouring Muslim villages of Larike and Wakasih and the Christian village of Liliboi (Collings 1983, Laidig and Laidig 1995). As in much of Maluku, there has been a long and increasing process of shift from indigenous language to the locally used variety of Malay, Ambonese Malay. This shift has been accelerated in recent years by the widespread use in education and media of Indonesian, the variety of Malay developed as the national language of Indonesia. In the case of Allang, shift to Ambonese Malay began in the early 20th century and today only some sixty to seventy people, all over the age of 65, are Allang speakers. They themselves use Ambonese Malay as their preferred language of daily communication, even with other speakers of Allang. In Liliboi the process of shift to Malay is complete and the indigenous language is now silent there. The language is still

¹ Data collection on which this paper is based was part of the Endangered Maluku Language Project supported by Australian Research Council grant DP0343379 and The Hans Rausing Endangered Languages Project grant MDP0009. My thanks to Betty Litamahuputty for help with transcription and to Barb Kelly and Lewis Mayo for discussions on some of the ideas used here. I am also indebted to the speakers in the event analysed here, Bapa Oyang de Fretes and Bapa Nik Sabandar and to the Allang community as a whole for their support.

Paul Sidwell & Uri Tadmor, eds. *SEALSXVI: papers from the 16th meeting of the Southeast Asian Linguistics Society* (2006). Canberra, Pacific Linguistics, 2008, pp.5-20.

© Michael Ewing

widely spoken in Larike and Wakasihu, although there are reports that shift to Malay is underway there as well, among school-age children. (For a discussion of the relationship between village religious affiliation and language shift, see Musgrave and Ewing 2006). There is no generally accepted name for the indigenous language shared by these villages and, following local practice, in this paper I refer to the variety spoken in Allang village as the Allang language.

The speech event that I am investigating involved two village elders, Bapa Nik and Bapa Oyang, who together with myself, met to produce a version of the village history in the Allang language. Bapa Nik is the Tuan Adat of Allang, the head of the clan who are the traditional keepers of village custom. As such he has very high cultural status and, among other roles, is often called upon to officiate at village ceremonies. Bapa Nik is in his mid-70s and has low proficiency in Allang language. Bapa Oyang is also in his mid-70s and is recognised as one of the best remaining speakers of the Allang language. He has high social status due to his economic position and involvement in civil and church affairs but he has no ritual status and no role within the customary village structure. Indeed, Bapa Oyang's family emigrated to Allang from a neighbouring island some 150 years ago and as such are still considered newcomers. As with most village residents these two men preferred to speak Ambonese Malay to each other. At the time the recording was produced in 2004, I was at the beginning of my second fieldtrip to Allang (the first having been for two months, two months earlier). Although familiar with Allang at this time, I was not conversationally fluent and mainly interacted with these two men in a mixed (Ambonese and Indonesian) variety of Malay. The meeting to produce a history of Allang in the Allang language was initiated by Bapa Nik. I had previously interviewed Bapa Nik in his capacity as Tuan Adat. Among other information, he provided his version of the history of Allang village, told in Ambonese Malay. This history of Allang as told by Bapa Nik involves the migration several hundred years ago of a group of families from the north, who travel over a number of years, stopping at various islands, until finally reaching a location where they work with an indigenous community already occupying the area to found the present-day Allang village. At the end of that session Bapa Nik said it would be nice to have a version of the history in the Allang language. Conceding that he himself could not produce such a text, he asked me to organise a meeting with Bapa Oyang, with whom Bapa Nik knew I had been working on language documentation. The meeting occurred the next day at Bapa Oyang's home. Initially my role was to prompt Bapa Oyang, using a transcript of the Malay version of the history Bapa Nik had previously produced. Very soon however, Bapa Nik began to tell his version of the history spontaneously in a combination of Malay and Allang for Bapa Oyang to reproduce in Allang sentence-by-sentence. My role then, in addition to operating the audio recording equipment, was to write down the 'received' version as they produced it. The audio recording and transcript of this speech event form the data analysed in the present discussion. The process by which this history was produced is explicitly outlined by Bapa Nik in the following excerpt. Note that in this example Bapa Nik, in contrast to his own instructions, actually produces both a Malay and an Allang version of his sentence, which is then confirmed by Bapa Oyang (whom Nik refers to as Wate 'uncle'). In other examples, Bapa Oyang takes a much more assertive role in the production of Allang language. In this and following examples, bold type represents material produced in the Allang language while non-bold type represents material produced in Malay (whether Ambonese Malay or Indonesian).

(1)

N: Ini beta bicara deng Bahasa Indonesia baru,
this 1s say with language Indonesian then

.. Wate terjemahkan.
uncle translate

M: Ya.
yes

N: Dong tinggal di situ seng lama,
3p stay at there NEG Long

karena terlalu dekat deng tanpa perang.
because too close with place war

O: Oh ya ya.
Oh yes yes

.. Jadi e=m,
so Um

N: ... **Mati rue hama ta lou.**
3p stay there NEG long

O: **Mati rue hama ta lou.**
3p stay there NEG long

N: [speaking to O] I'll speak in Indonesian then you can translate.

M: Yes.

N: They didn't stay there long because it was too close to the battle field.

O: Oh, yeah yeah. So, uh,

N: **They didn't stay there long.**

O: **They didn't stay there long.**

This example also gives us a first impression of these two men working interactively as they produce a history of Allang village. In the discussion that follows we will see more complex examples of these men negotiating both the content and the linguistic form of the history as it is developed. The interactional work that these men do reveals the distributed nature of the knowledge they are working with. One aspect of this knowledge that contributes to it not being entirely shared between them is that it is specialist knowledge. In the case of historical knowledge, especially the received narrative of local history, is it is not surprising that this is a kind of specialist knowledge, not shared evenly by different members of the community. Indeed Bapa Nik's role as *Tuan Adat* means he is institutionally recognised in the village as someone with specialist knowledge, including, among other things, knowledge about the history of the village. Linguistic knowledge, especially regarding ritual language or literary texts, is also often specialist knowledge. General linguistic ability, however, is commonly shared among members of a robust speech community and simply being able to use the language in everyday contexts is not usually considered specialist knowledge. In a language

shift context like that in Allang, it is precisely this kind of erstwhile mundane linguistic competence that is quickly becoming highly specialised knowledge. As such, its distributed nature comes into high relief. As linguistic competence recedes, it recedes differentially in different individuals and these differences in the individual competences give rise to conflict and negotiation over some of the most fundamental aspects of grammatical and semantic knowledge.

In the event being analysed here, Bapa Nik has greater access to and authority over knowledge concerning the history of the village and Bapa Oyang has greater access to and authority over linguistic expression in Allang. At the same time, both men have some knowledge of and hold opinions about aspects of the other's area of specialisation. Throughout the interaction, knowledge of history and knowledge of language are both negotiated. Through these negotiations, not only language and content, but also the relative social status of the two interactants is also (re)established. Finally, from a linguistic perspective, these are negotiations about grammar and through these negotiations we can see aspects of Allang grammar that are changing in this process of language shift.

While Bapa Oyang clearly has greater control over the language, he is dependent on Bapa Nik to provide both the overall narrative as well as details, such as place names. In example (2) Oyang stops to ask Nik explicitly for the relevant information that will allow him to finish the sentence he is constructing.

(2)

N: **Mati rue he=**,
3p stay at

O: **Mati kene**,
3p go

mati lai hapa.
3p arrive where

N: **Lai eh=**,
arrive HES

Ulatete.
Ulatete

O: .. **Ulatete?**
Ulatete

N: **They stayed at,**

O: **The went and they arrived where?**

N: **The arrived, uh Ulatete.**

O: **Ulatete?**

The remainder of the discussion will focus mainly on linguistic aspects of this cooperative story construction. First, we see that Bapa Nik consistently makes smaller contributions than Bapa Oyang, while Bapa Oyang produces elaborations, using both

more extended lexicon and a wider range of discourse and grammatical structures than Bapa Nik. In (3) Oyang introduces a sequence of clauses with *Mati lehun lumanu* ‘they met together’ to which Nik attempts to add some material. Oyang rejects Nik’s additions and reformulates Nik’s simpler *mati husa* ‘they left’ as a more complex characterisation of their actions. He searches for a verb – *atur* ‘organise’ and *rewa* ‘search for’ – accepts Nik’s one-word contributions of *wala* ‘place’ and finishes with the more elaborate *mati rewa wala ria matir na upu anai na masa depan* ‘they searched for a place for the future of their descendents.’ Even this clause is also built up slowly, with Oyang working through two possessive construction and Nik contributing the final (Malay) word meaning ‘future’, before the sentence is acceptable to them both.

(3)

O: **Mati lehun lumanu,**
3p meet RCP

M: Hmh.
uhuh

N: **Mati lepa,**
3p say

O: **Mati=,**
3p

N: ... **Mati lehun lumanu ria mati husa.**
3p meet RCP for 3p exit

O: Nee.
no.NED

Mati lehun lumanu ria mati=,
3p meet RCP for 3p

.. **atur=,**
arrange.MAL

... **rewa,**
look.for

N: **Wala.**
place

O: **Ya.**
yes

... **Wala ... ria mas- mati matir upu ana na,**
place for 3p 3p.POSS grandchild child POSS

matir na upu anai na masa=,
3p POSS grandchild child-PL POSS time.MAL

N: **depan.**
future.MAL

O: **depan.**
future.MAL

O: **They meet together,**

M: Uhuh.

N: **They said,**

O: **They,**

N: **They meet together to leave.**

O: No. **They meet together to arrange to look for,**

N: **a place.**

O: **Yes, a place for, their descendents, for their descendents',**

N: **future.**

O: **future.**

In (4) Oyang expands the language by making explicit background material that Nik had only implied. Note that as speaker M (myself) repeats the other speakers' utterances, he is in the process of writing them down. In this case, Nik simply uses the third person plural pronoun *mati* to refer to each of two parties involved in an altercation, thus creating a sense of ambiguity. Oyang however, chooses to disambiguate by explicitly stating the identity of the two parties. This is not simply an example of more elaborate language in the sense of using larger expressions, it also demonstrates a greater ability on the part of Oyang to manipulate the language's reference tracking mechanisms with recipient design in mind. But again, the final word, crucial to the intent of the sentence, is supplied by Nik.

(4) Bapa Oyang expands by adding explicit background

O: **Lai he Hanunu,**
arrive at Hanunu

M: .. **Lai he .. Hanunu,**
arrive at Hanunu

N: **mati ta tarima mati.**
3p NEG receive 3p

O: **mati= e= supu luman la mansia=,**
3p HES meet together with person

N: ... **e Maliai.**
HES indigenous-PL

O: **Arriving at Hanunu,**

M: **Arriving at Hanunu,**

N: **they didn't receive them.**

O: **they meet with the people,**
 N: uh, indigenous people.

Similarly in (5), Oyang takes Nik's reduced phrase *Mati heten musti perang* 'They said (they) must fight' and again elaborates, not only with a more explicit characterisation of the agent of *heten* 'say' but also by explicitly mentioning the recipient. Additionally, in the reported speech segment, Oyang uses overt reference to the actors involved, rather than the implicit reference employed by Nik.

(5)

N: **Mati heten musti perang.**
 3p say must.MAL fighty.MAL

O: **Eh=m.**
 uhuh

Keuta tarima mati.
 not.want receive.MAL 3p

Jadi Patasiwa Hoamoal heten [ka] Maheri la mana rombongan,
 so.MAL Patasiwa Hoamoal say to Maheri and 3p.POSS group.MAL

N: **[heten].**
 say

M: **Hmhm.**
 uhuh

O: **eh,**
 HES

Ite musti perang.
 1pi must.MAL fight.MAL

N: **They said they must fight.**

O: **Uhuh. They didn't want to accept them. So the Patasiwa Hoamoal said to Maheri and his group,**

N: **said,**

M: **Uhuh,**

O: **Uh, We must fight.**

Another way that Oyang expands Nik's linguistic contribution is through use of transitional material lacking in the version originally produced by Nik. In (5) Nik produces a simple declarative statement meaning 'they looked for food again' followed by a request for confirmation of the appropriateness of his contribution. Then, in Malay, Nik provides background to the event. Oyang takes this background and weaves it into the more elaborate Allang construction, adding a temporal adverbial clause and also refining the main clause with a realis marker. On the one hand, the borrowed linker *kalau* 'if, when' to

mark the time adverbial shows the clear influence that Malay is having on the use of Allang by its remaining speakers; nonetheless, the use of the realis marker on the main clause is a more idiomatic Allang construction and marks that clause as something more than simply a calque on Malay structure, a characterisation that could be applied to Nik's original contribution.

(6)

N: Artinya,
meaning-DEF

mati sewa anuta peluke.
3p look.for food again

M: Hmhm.
uhuh

N: Eh Wate?
Q uncle

O: Ya.
yes

N: Barang sampai di situ dong berhenti.
that until at there 3p stop

O: **Kalau lai ma,**
if.MAL arrive there

mati rewa anuta paluke mae.
3p look.for food again RLS

N: That means, **they looked for food again.**

M: Uhuh.

N: right, uncle?

O: Yes.

N: That was when they were there, they stopped.

O: **On their arrival, they looked for food again.**

The relative level of linguistic proficiency of these two men is also demonstrated in a reverse situation in which Nik simplifies contributions from Oyang. In (7) Oyang emphasises the repetition of *paluke* 'again'. This repetition both provides an elaborating adverbial element which links these clauses to the larger discourse and also provides an aesthetic appeal by emphasising the repetitiveness of the activities the travellers experience during this portion of their journey. Nik however, when repeating Oyang's contribution, simplifies it to the bare clause.

(7)

O: **Mati kene peluke.**

3p go again

.. **Mati snapau peluke.**

3p travel again

N: **Mati snapau.**

3p travel

O: **Mati snapau peluke.**

3p travel again

O: **They went again. They travelled again.**N: **They travelled.**O: **They travelled again.**

The two men also negotiate appropriate lexical items to use. In example (8) Oyang characterises the ancestors' original religious system as the belief in stones. Nik suggests the alternative characterisation that they believed in nature, using the Malay term *alam* 'nature'. Oyang then provides an Allang language reference to a traditional power in nature *Upa Lanita*, literally the Lord of the Sky. It is interesting to note that in other contexts this same Allang term is used to refer to the Christian God as well.

(8)

O: **Jadi,**

so.MAL

ya,

yes

mati percaya ria [hatori],

3p believe.MAL for stone-PL

N: **[Alam].**

nature.MAL

O: **Iya,**

yes

mati cuma percaya ria ma,

3p only.MAL believe.MAL for that

N: **Alam XX.**

nature.MAL

O: **alam,**

nature.MAL

Upu Lanita.

lord sky

Upu Lanita.

lord sky

N: [**Upu Lanita**].

lord sky

O: [**Mati percaya**] **ria Upu Lanita.**

3p believe.MAL for lord sky

O: So, yes, they believed in stones,

N: nature.

O: Yes, the only believed in that,

N: nature.

O: nature. The Lord of the Sky. The Lord of the Sky.

N: They Lord of the Sky.

O: They believed in the Lord of the Sky.

The remaining examples all deal with negotiating of the actual grammatical structures used in the language. Example (9) illustrates negotiations over possessive constructions. It is interesting that Oyang actually concedes to Nik, sanctioning his innovative construction. Conservative speakers of Allang employ a full set of possessive case pronouns, but innovative speakers who do not have command of this case distinction employ actor case pronouns to mark possessive (Ewing 2005). In the first line in Example (9) Nik produces this type of innovative construction using the actor prefix *me-* ‘3sh’ to indicate third person possession. This is followed by a conservative third person plural possessive produced by Oyang who uses a double marking strategy (cf. Florey 2005) employing both the possessive pronoun *matir* ‘3p.POSS’ and the possessive particle *na*. In an interesting case of what might be characterised as mutual accommodation, Oyang then continues the narrative utilising *me-nalana*, the innovative third person possessive originally produced by Nik, a kind of hypocorrection (Baugh 1992). This is followed by Nik who now reformulates this same possessive reference using the conservative *mana nalana*, perhaps influenced by Oyang’s earlier conservative construction.

(9)

N: **Me-nalana Talukubessy.**

3sh-name Talukubessy

M: Tu- --

O: **Matir na pimpinan,**

3p.POSS POSS leader.MAL

hama i= [Hitu].

here this Hitu

N: **[Hitu].**

M: Hmhm.
uhuh

O: **Me-nalana.**
3sh-name

N: **Mana nalana,**
3sh.POSS name

O: .. **Talukubessy.**
Talukubessy

N: **His name was Talukubessy.** [or “he name”]

M: **Tu-**

O: **Their leader, here in Hitu.**

N: **Hitu.**

M: Uhuh.

O: **His name.** [or “he name”]

N: **His name was,**

O: **Talukubessy.**

In (10) we see a reformulation that centres around the contentious use of prepositions. Allang differentiates between two prepositions *ria* ‘in order to’ and *ka* ‘for (benefactive)’, which happen to be marked by one polysomus preposition *untuk* ‘for, in order to’ in Malay. In Nik’s first line in this example, he completes Oyang’s phrase ‘they looked for food’ with a phrase intended to mean ‘for themselves’. Nik uses the preposition *ria*, which would mean ‘in order to’ for a conservative speaker of Allang, but which, possibly on analogy with Malay, Nik intends to mean ‘for (benefactive)’. Interestingly, Oyang again accommodates, repeating the form produced by Nik, but then rejects this hypocorrection and returns to the more conservative construction with *ka*.

(10)

O: **Mati rewa antua,**
3p look.for food

M: Hm.
hm

N: **Ria mati.**
for 3p

O: **Rewa,**
look.for

mati rewa anuta ria mati,
3p look.for food for 3p

M: .. Hmhm,
uhuh

.. **Ma- mati rewa** [anuta],
3p look.for food

O: [**Anuta ka mati**].
food for 3p

N: [**Anuta**] **ria mati**.
food for 3p

O: **They looked for food.**

M: **Hm.**

N: **For themselves**, [or “in order to”]

O: **Looked for, they looked for food for themselves.** [or “in order to”]

M: **Uhun, They looked for food,**

O: **food for themselves.**

N: **food in themselves.** [or “in order to”]

Spatial reference is an important area of complexity in the grammars of languages of Maluku (Florey and Kelly 2002) as they are of many Austronesian languages and this is the case in the language of conservative Allang speakers as well. In example (11) we can see Oyang explicitly explaining this aspect of Allang grammar, in Malay, to Nik after Nik begins to present a sentence in Allang that makes (inappropriate) use of a directional.

(11)

N: **mati sanapau selat Misol.**
3p travel strait Misol

Ma- mati roti.
3p descend

O: Nanti dulu Oom Kola.
later first uncle Kola

% % **Sebabe barang beta belum tahu yang keluar dari Tidore,**
because thing 1s not.yet know REL exit from Tidore

dong bajalan di k-
3p travel at

dong bajalan di darat atau dong bajalan di laut.
3p travel at land or 3p travel at sea

N: Laut ini.
sea this

Dong turun.
3p descend

N: **They travelled to Misol. They went down.**

O: Just a minute Uncle Kola. Because the thing I don't know yet is from Tidore, did they travel, did they travel by land or did they travel by sea.

N: This was by sea. They went down (out to sea).

Here Oyang makes it explicitly clear that he needs to know whether the travellers were setting off by sea or by land. Because Tidore is a small island and Nik has said that they are departing the island, it may seem self-evident that these travellers would be going by sea. Oyang's questioning their direction of movement can be read as challenging Nik's grammatical usage rather than simply a request for information. Indeed these men seem to be speaking at cross purposes. In Malay it is common to use the metaphor of descent when talking about going to sea but in Allang the metaphor of ascent is used when going to sea. This 'mismatch' or 'reversal' between Malay and Allang was one of several examples that speakers like to point to in their construction of the otherness of Allang vis-à-vis Malay. Control of such 'reversed' constructions becomes a sign of expertise in the Allang language. In the last line of (11) Nik uses the Malay turun 'to descend' when he says they went to sea. His mistake had been to use the direct equivalent *roti* 'to descend' in Allang where a conservative speaker would expect *hara* 'to ascend'. It was this Malay influenced directional metaphor by Nik which prompted Oyang's (possibly feigned) confusion as to the protagonists' direction of travel.

Example (12) combines a number of the collaborative techniques that we have seen illustrated above. At this point in the story Nik mentions that the protagonists are taking a look or observing, without saying what they are looking at. Oyang elaborates Nik's contribution with an expanded linguistic construction that includes reference to what they are looking at. In the first instance this entails use of a directional of the type that less proficient speakers often do not control: rather than the generalised preposition he 'to, at' favoured by many speakers now, Oyang uses *lo* 'seaward'. At this point, Nik's expertise as the holder of the history comes in to play. Oyang goes as far as saying that they looked seaward toward an island, but he relies on Nik to supply the name of the island that is relevant at this point in the narrative.

(12)

N: **Mati kari'i.**
3p look

O: **Ma=ti=,**
3p

N: **Mane kari'i.**
3sh look

O: **Mae=,**
3sh

mane kari'i lo= Nusa,
3sh look seaward island

N: .. **Nusa Boano.**
island Boano

O: **Nusa .. Boano.**
island Boanao

N: **They looked.**

O: **They,**

N: **He looked.**

O: **He, he looked seaward to the island of,**

N: **Boano Island.**

O: **Boano Island.**

Historically, the languages of Central Maluku have had systems of verbal conjugation in which the initial sound of the verb changes according to person and number of the actor (Collins 1983, Collins and Kaartinen 1998). Such a system of verb conjugations is still used by speakers of the Larike variety of the language (Laidig and Laidig 1995) and data I recently collected from Wakasihu suggest that this system is still viable among Wakasihu speakers. In Allang however, the system has broken down and these days many verbs occur in pairs, which Allang speakers no longer conceive of in terms of grammatical conjugation. For some speakers, members of certain pairs are considered equally viable alternatives which can each be used in any situation. For others, members of pairs that were historically conjugations are now described as having subtle semantic or pragmatic differences. Thus which of two possible realisations of a given verb is appropriate in a given context is often contentious in Allang, but no longer follows a grammatical paradigm of conjugations. Such contention appears in example (13)

(13) Verb conjugation

O: **Mati lai he Wahai ma ria mati rewa,**
3p arrive at Wahai that for 3p look.for

N: **Sewa anuta.**
look.for food

O: **rewa,**
look.for

rewa,
look.for

rewa anuta.
look.for food

- O: **They arrived at Wahai in order to (r-) look for,**
 N: **(s-) Look for food.**
 O: **(r-) Look for, (r-) look for, (r-) look for food.**

Here Oyang uses the form *rewa* ‘to look for’ which Nik changes to *sewa* when he repeats Oyang’s first statement. Oyang then asserts his linguistic authority by insisting on the form *rewa*. Interestingly, in the context of the third person plural actor used here, Nik’s *sewa* is the form that would have been the historically appropriate conjugation. This point is not entirely moot in the current Allang situation, because text counts of natural usage by fluent speakers suggest a strong tendency still to use the alternating forms according to their historical paradigms, although when asked to make explicit conscious judgements these same speakers may allow for either form to be used in any context. Such judgements contrast with Wakasihu speakers who still insist on maintaining to the grammatical paradigm. In the Allang environment, choice between alternative forms is based on locally contingent needs rather than grammatical paradigms. In Oyang’s case, the motivation for correcting Nik could have been largely a matter of asserting his linguistic authority.

The examples discussed above have shown ways that two elders of the Allang community have cooperated in the process of developing the text of a history of Allang in the Allang language. Cooperation is a well know value in Indonesian (as it is of course in most parts of the world), often expressed in national discourse with the term *Gotong Royong*. Also common in Maluku is the Malay *kerja sama* or in Allang, the indigenous term *eseu palalaoi*, both of which can be glossed ‘working together’. The importance of the cooperative work that these two men did together was expressed by Bapa Nik, who at the end of the working session said in Malay:

Kerja sama, Artinya saling kasih pengertian. Eseu palaloi. Jadi, kalo bicara sejarah dengan bahasa Malayu bole. Mar, trima kasih keadaan trus terjemahkan deng bahasa daerah. Itu yang beta minta trima kasih. Kalo bukan ontua, juga susah.

‘Working together. That means sharing understanding with each other. Cooperating. So, I can tell the history in Malay. But, I’m thankful for this situation where it can be translated into the indigenous language. That’s what I’m thankful for. Without Bapa Oyang it would have been difficult.’

Yet the process described here is more than a division of labour or a matter of more hands make for lighter work. In the areas of historical knowledge and linguistic knowledge, including knowledge of lexicon, grammar and discourse structure, the resources needed to achieve their goal were differential distributed. It was only through a process of social interaction that these resources could be harnessed to produce the desired outcome. Another Malay phrase that I have often heard during field work in Maluku is *saling mengisi* ‘filling each other(’s knowledge)’, a recognition that no one is able to achieve their desired goals alone and it is only by utilising socially based distributed cognitive processes like those illustrated here that these goals can be reached.

Glossing conventions

-	truncated word
%	glottal sound
1pi	1st person plural inclusive
1s	1st person singular
3p	3rd person plural
3sh	3rd person singular human
DEF	definite
HES	hesitation
MAL	Malay
NEG	negative
PL	plural
POSS	possessive
Q	question
RCP	reciprocal
REL	relative clause marker
RLS	realis

References

- Baugh, John. 1992. Hypocorrection: Mistakes in Production of Vernacular African American English as a Second Dialect. *Language and Communication* 12.3-4:317-26.
- Collins, James T. 1983. *The Historical Relationships of the Languages of Central Maluku, Indonesia*. Canberra: Pacific Linguistics.
- Collins, James T. and Timo Kaartinen. 1998. Preliminary notes on Bandanese. *Bijdragen tot de Taal-, Land-, en Volkenkunde* 154.4:523-570.
- Ewing, Michael C. 2005. Diverging and converging patterns of possession: Allang in its Central Maluku Context. *Monash University Linguistic Papers* 4.1:45-64.
- Florey, Margaret & Barbara Kelly. 2002. Spatial reference in Alune' in G Bennardo (ed.) *Representing Space in Oceania: culture in language and mind*, 11-46. Canberra: Pacific Linguistics.
- Florey, Margaret. 2005. A cross-linguistic perspective on emergent possessive constructions in Central Moluccan languages. *Australian Journal of Linguistics* 25.1:59-84.
- Hutchins, Edwin. 1995. *Cognition in the wild*. Cambridge, Mass.: MIT Press.
- Laidig, Wyn D. and Laidig Carol J. 1995. A Synopsis of Larike Phonology and Syntax. In Wyn D. Laidig (ed.) *Descriptive Studies in Languages of Maluku*, 19-42. NUSA 38. Jakarta: Badan Penyelenggara Seri NUSA.
- Musgrave, Simon and Michael C. Ewing. 2006. Language and religion: a case study of two Ambonese communities. *International Journal of the Sociology of Language*. 179:179-194.
- Salomon, Gavriel (ed). 1993. *Distributed cognition: psychological and educational considerations*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. 1991. Conversation analysis and socially shared cognition. In Lauren B. Resnick, John M. Levine and Stephanie D. Teasley (eds.) *Perspectives on socially shared cognition*. Washington DC: American Psychological Association.

RELATIVE CLAUSES IN K'CHO

Kee Shein Mang
Payap University
<keesmang@gmail.com>

George Bedell
Payap University
<gdbedell@yahoo.com>

The most basic type of relative clause in K'cho is illustrated in (2), (3) and (4), which presuppose a transitive sentence like (1).¹

- (1) Om noh Yóng am pàapai pe(k) ci.
Om P Yong to flower give NF
Om gave Yong a flower.
- (2) [Yóng am pàapai pe(k) ci ah] k'chàng
Yong to flower give NF C man
the man [that gave Yong a flower]

¹ The K'cho language is spoken in southern Chin State, Myanmar, principally in Mindat Township. The population of K'cho speakers is somewhere between 10,000 and 20,000. K'cho has no generally accepted standard orthography; we have followed the most widely used conventions, as seen in Jordan (1969) or in *Ng'thu-Thai* (2001), but supplemented by distinguishing long from short vowels (doubling the former) and tones (marked only for content words; a grave accent indicates a low tone and an acute accent a rising tone; unmarked syllables have a high tone).

The symbol (*k*) indicates a velar stop inserted after any word which ends in a short vowel and is followed by the tense particles *ci* or *khai*. The symbols *k'*, *m'* and *ng'* indicate a prefixed glottal stop or nasal. A syllable final glottal stop is indicated by *h*. See Nolan (2002) for more details concerning K'cho phonology and orthography.

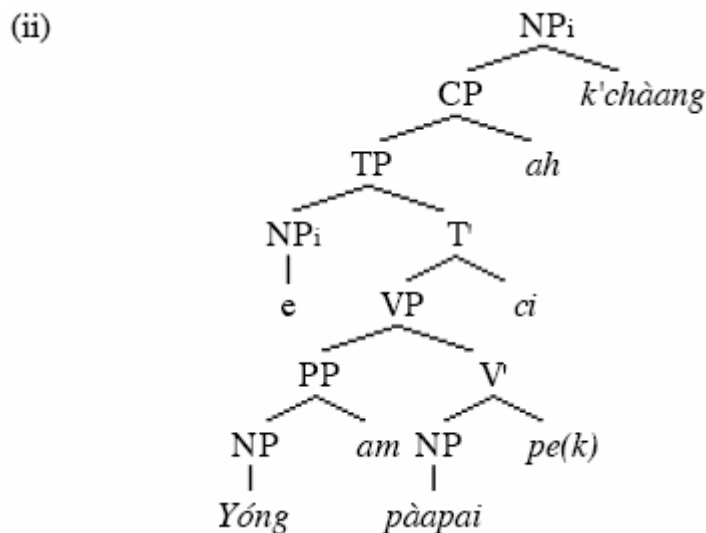
Word-by-word glosses contain the following abbreviations:

C complementizer (or conjunction)	D deictic (or demonstrative)
DIR directional	DL dual
EMP emphatic	F future
NF non-future	OPT optative
P postposition or particle	PL plural
1S first person singular	2 second person
2S second person singular	3S third person singular
3D third person dual	3P third person plural

This paper was prepared for presentation to the 16th Annual Meeting of the Southeast Asian Linguistics Society, Atma Jaya University, Jakarta, September 20, 2006. Mang is responsible for the principal examples, and most of arguments and conclusions of the paper. Bedell is responsible for the structure diagrams and for the English text.

- (3) [Om noh Yóng am a péit ah] pàapai
 Om P Yong to 3S give C flower
 the flower [that Om gave Yong]
- (4) [Om noh pàapai a péit ah] k'hngumí
 Om P flower 3S give C woman
 the woman [that Om gave a flower to]

K'cho relative clauses are marked by *ah*, the general subordinating particle. We take it to belong to the category C (conjunction or complementizer), a syntactic head which takes a clause complement. If the argument which serves as the locus of relativization is the subject, as in (2), the clause verb (*pe*) appears in the same form as in the independent sentence (1). But if the relativized argument is not the subject, as in (3) or (4), the clause verb appears in a different form (*péit*). This phenomenon is generally called 'verb stem alternation', and is a characteristic of Kuki-Chin languages.² Here *pe* is the stem I form, and *péit* the stem II form. In K'cho, in addition to the morphological change in the stem itself, a stem I verb form requires a tense particle (such as *ci*) and allows no subject third person agreement, while a stem II verb allows no tense particle, but requires a subject third person agreement particle (such as *a*).³ We take the structure of (2) to be as in (ii).

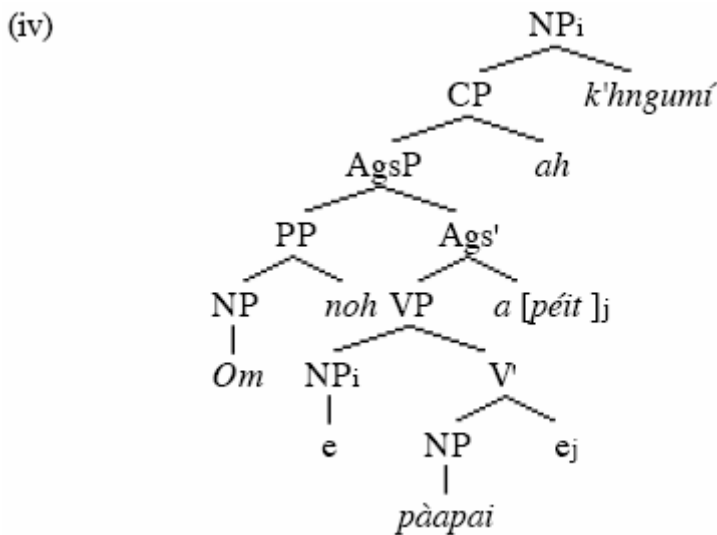
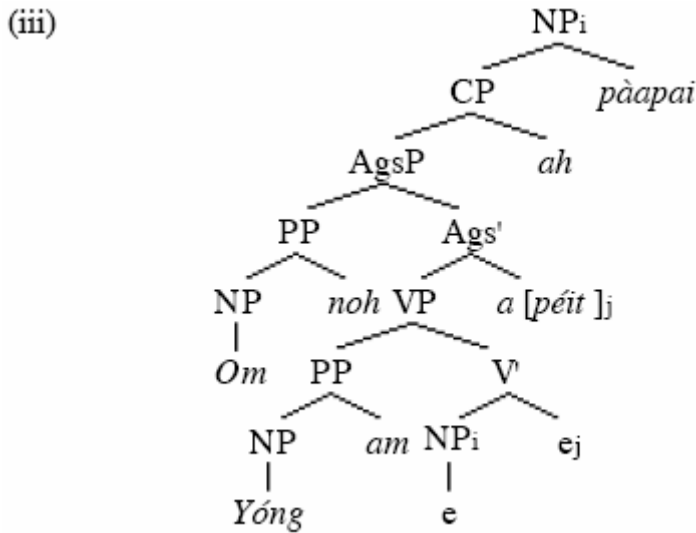


(ii) illustrates what we take to be the universal relative clause structure: the relative clause contains a noun phrase (NP) which is understood to refer to the same entity referred to by the NP which contains the clause. This is indicated by co-indexing the two NPs (NP_i). In (ii) the NP within the relative clause is empty. A transitive subject is normally accompanied by the postposition *noh* as in (1). In K'cho postpositions cannot be 'stranded', that is, appear with no overt NP complement. We take the absence of a postposition accompanying an empty NP to imply the absence of a postpositional phrase (PP) as well.

² For details of verb stem alternation in K'cho, see Nolan (2003) and Mang (2006).

³ For details of agreement in K'cho, see Bedell 2000.

The structures of (3) and (4) will then be as in (iii) and (iv).



No tense phrases (T' or TP) appear in (iii) or (iv) since the tense particle (*ci*) which would head these phrases is absent. On the other hand, subject agreement phrases (Ags' and AgsP) appear headed by the subject agreement particle *a*. The particle *a* is prefixed to the stem II verb (*péit*) which is co-indexed with any empty head V position (*e_j*). (iv) is another example of an empty NP without its usual postposition (it would be *am* in this case).

Relativization in K'cho is not limited to verbal arguments as in (2) to (4). Genitive noun phrases as illustrated in (5), (7) and (9) may also be relativized, as in (6), (8) and (10).

- (5) Ka teihpüi noh Yóng am pàapai pe(k) ci.
 1S friend P Yong to flower give NF
 My friend gave Yong a flower.

- (6) a teihpüi noh Yóng am pàapai pe(k) ci ah k'chàng
 3S friend P Yong to flower give NF C man
 the man [whose friend gave Yong a flower]

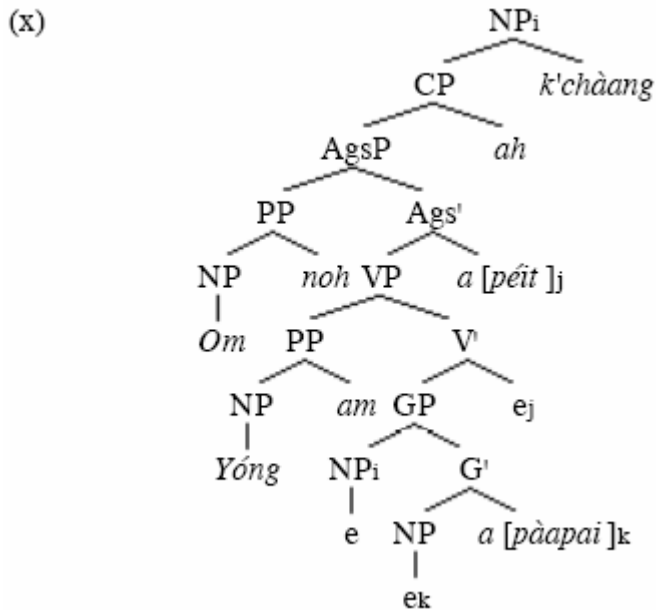
In (6), a genitive on the subject is relativized. Notice that the genitive agreement particle (here third person singular *a*) indicates where the relativized noun phrase is located. But just as in the previous examples, the relativized noun phrase itself is empty, and *a* is not to be taken as a 'resumptive' pronoun. Notice also that the clause verb appears in its stem I form, just as in (2) where it is the subject which is relativized.

- (7) Om noh ka teihpui am pàapai pe(k) ci.
 Om P 1S friend to flower give NF
 Om gave my friend a flower.
- (8) Om noh a teihpüi am pàapai a péit ah k'chàng
 Om P 3S friend to flower 3S give C man
 the man [whose friend Om gave a flower to]

In (8), a genitive on the indirect object is relativized. As in (6), genitive agreement indicates the locus of relativization. But here the clause verb appears in its stem II form, just as in (iii) where it is the indirect object which is relativized.

- (9) Om noh Yóng am ka pàapai pe(k) ci.
 Om P Yong to 1S flower give NF
 Om gave Yong my flower.
- (10) Om noh Yóng am a pàapai a péit ah k'hngumí
 Om P Yong to 3S flower 3S give C woman
 the woman [whose flower Om gave Yong]

In (10), it is a genitive on the direct object which is relativized. As in (8) we see a genitive *a* and stem II of the clause verb. The structure of (10) will look like (x).



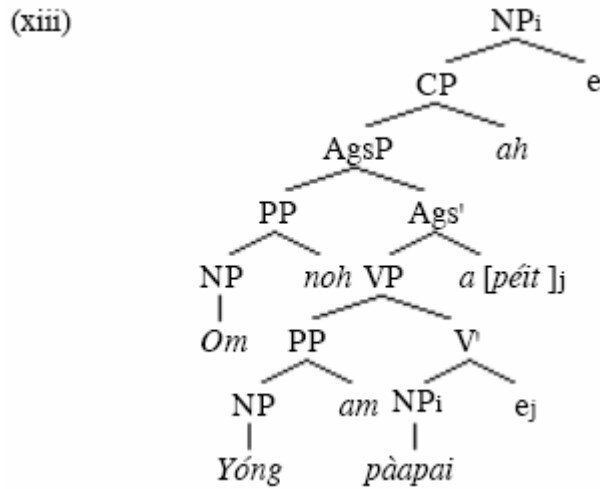
This *a*, together with others indicating different persons and numbers, belongs to the category of genitive agreement particle (G). It is a syntactic head which combines with a noun phrase to form a genitive phrase (G') and with another noun phrase to form a larger genitive phrase (GP). The head noun of the lower NP has the G prefixed to it in the same way as the subject and object agreement particles are prefixed to the verb.

K'cho relative clauses like those in (2) to (4), (6), (8) and (10) precede the head noun of the NP which contains them as shown in (ii), (iii), (iv) and (x). This type of clause is termed 'head-external'. Many Kuki-Chin languages have also relative clauses with no following head noun, but with a noun internal to the clause understood to be the head. K'cho lacks such 'head-internal' relative clauses. Such relative clauses would be as in (11), (12) and (13) corresponding to (2), (3) and (4). These are grammatical K'cho sentences, but the clauses in them cannot be understood as relative clauses. Rather, they are complement clauses with the sensory verb *hngu* 'see'. The meaning is that the event referred to by the clause is what was seen, and not any of the persons or things involved. Notice that the verb form used in these clauses is uniformly stem II.

- (11) [K'pámi tumat noh Yóng am pàapai a péit] ka hngu(k) ci.
 man P Yong to flower 3S give 1S see NF
 *I saw the man [that gave Yong a flower].
 I saw a man give Yong a flower.
- (12) [Om noh k'hngumí am pàapai a péit] ka hngu(k) ci.
 Om P woman to flower 3S give 1S see NF
 *I saw the woman [that Om gave a flower to].
 I saw Om give a woman a flower.
- (13) [Om noh Yóng am pàapai a péit] ka hngu(k) ci.
 Om P Yong to flower 3S give 1S see NF
 *I saw the flower [that Om gave Yong].

I saw Om give Yong a flower.

The structure of a relative clause interpretation of (13) would be as in (xiii).

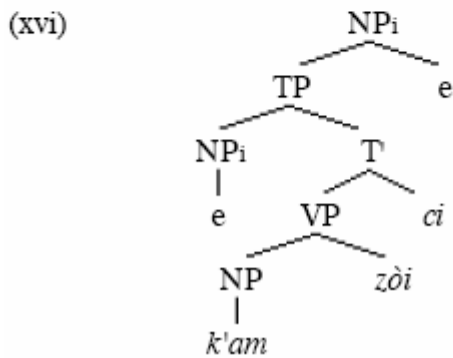
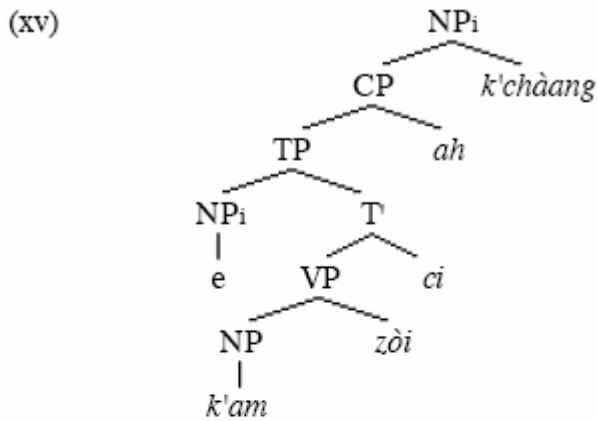


But we repeat that this is not a possible K'cho structure. The structure of (13) in its grammatical interpretation will be the same, but without the indices.

A second type of Kcho relative clause is illustrated by (16) and (18).

- (14) Om noh k'am zòi ci.
 Om P pot sell NF
 Om sells pots.
- (15) [K'am zòi ci ah] k'chàng lo(k) ci.
 pot sell NF C man come NF
 The man [that sells pots] came.
- (16) [K'am zòi ci] lo(k) ci.
 pot sell NF come NF
 The one [that sells pots] came.
- (17) [Om noh a zoih ah] k'am ka hngu(k) ci.
 Om P 3S sell C pot 1S see NF
 I saw the pots [that Om sells].
- (18) [Om noh a zoih] ka hngu(k) ci.
 Om P 3S sell 1S see NF
 I saw the ones [that Om sells].

As compared with (15) or (17), which are normal relative constructions like (2) and (3), (16) and (18) appear to have no head noun, and also to be missing the complementizer *ah*. The structure of the relative clauses in (15) and (16) will be as in (xv) and (xvi).



We take the absence of *ah* in (16) to imply no CP; but the NP at the top of (xvi) cannot be dispensed with, in spite of the absence of an overt head noun. A sentence like (16) should not be confused with one like (19) even though their meaning is very similar.

- (19) K'am-k'zòì lo(k) ci.
 pot sell came NF
 The pot seller came.

(19) contains no relative clause; *k'am-k'zòì* 'pot seller' is a kind of compound noun.

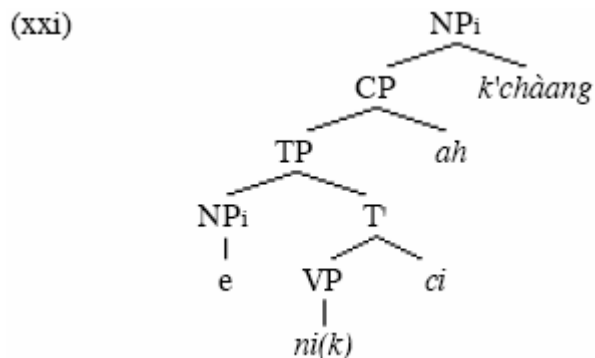
Parallel to (14) to (18) are examples like (20) to (22).

- (20) Om ni(k) ci.
 Om good NF
 Om is good.

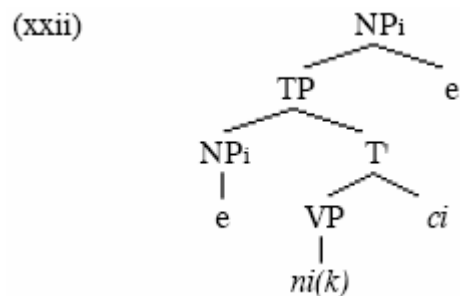
- (21) [Ni(k) ci ah] k'chàng ka hngu(k) ci.
 good NF COMP man 1S see NF
 I saw a man [that was good].

- (22) [Ni(k) ci] ka hngu(k) ci.
 good NF 1S see NF
 I saw one [that was good].

Like the relative clauses in (16) and (18), the one in (22) has no overt head and no complementizer. Unlike the transitive verb *zòì* ‘sell’ in (14), the verb *ni* ‘good’ in (20) is intransitive. If (21) and (22) seem more different in meaning than (15) and (16) or (17) and (18), this is because the linguistic context bearing on the interpretation of the empty head noun is richer in the latter cases. ‘Someone that sells pots’ or ‘something that Om sells’ are narrower in meaning than ‘something which is good’. To the extent that the linguistic context does not suffice to identify the empty noun, the extralinguistic context must do so. The structure of the relative clauses in (21) and (22) will be as in (xxi) and (xxii).



Comparing (xvi) or (xxi) with (xiii), it should be clear that the reason the latter is ungrammatical is not simply that it lacks an overt head noun. Like (19), (23) contains not a relative clause, but a compound noun *k'chàng-k'ní* ‘good man’.



- (23) K'chàng-k'ní ka hngu(k) ci.
 man good 1S see NF
 I saw a good man.

The compound nouns *k'am-k'zòì* in (19) and *k'chàng-k'ní* in (23) consist of a noun followed by a verb stem I in the former case, but by a form unique to this kind of compound in the latter. The stem I form of *ni* has a short vowel with a high tone, and the stem II form (*nii*) has a long vowel with a high tone. In (23) we see a rising tone (*k'ní*), which Nolan (2003) has called stem III.

A further set of examples is (25) to (28), based on the transitive sentence (24).

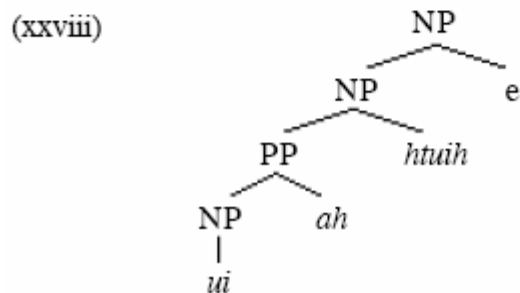
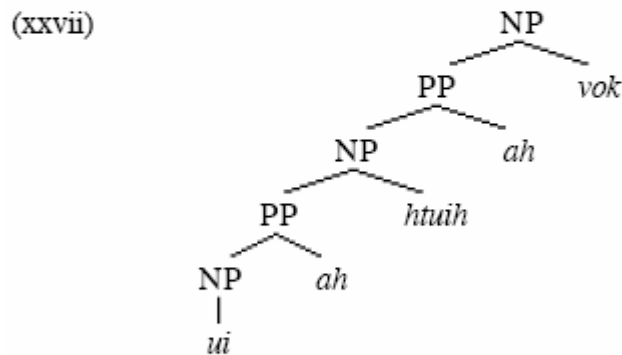
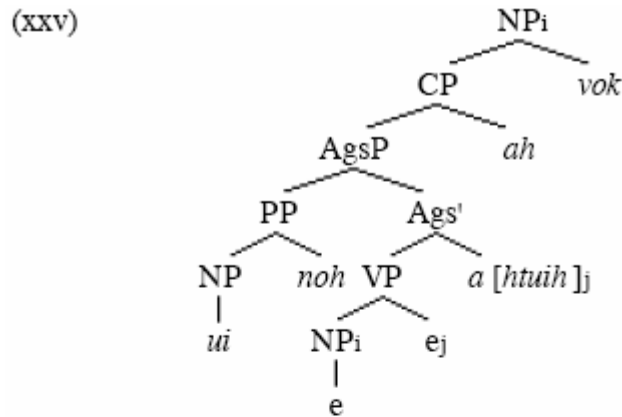
- (24) Ui noh vok htui ci.
dog P pig bite NF
A dog bit a pig.
- (25) [Ui noh a htuih ah] vok sì(k) ci.
dog P NF bite C pig die NF
The pig [that the dog bit] died.
- (26) [Ui noh a htuih] sì(k) ci.
dog P NF bite die NF
The one [that the dog bit] died.
- (27) Ui ah htuih ah vok sì(k) ci.
dog P bite P pig die NF
The pig bitten by the dog died.
- (28) Ui ah htuih sì(k) ci.
dog P bite die NF
The one bitten by the dog died.

(25) and (26) contain relative clauses, respectively headed and headless. (27) appears to have the same head noun as (25), and (28) to have the same relation to it as (25) has to (26). *Htuih* 'bite' also appears to be the same stem II form as in (25) and (26). But there is no subject agreement particle (normally required with a stem II verb form), and the noun in the clause (*ui* 'dog') is followed by *ah* rather than *noh* (normally required on the subject of a transitive verb). Structures like (27) and (28) are not possible if the subject is relativized. (29) and (30) contain relative clauses parallel to those in (25) and (26), but (31) and (32) are ungrammatical.

- (29) [Vok htui ci ah] ui sì(k) ci.
pig bite NF C dog die NF
The dog [that bit the pig] died.
- (30) [Vok htui ci] sì(k) ci.
pig bite NF die NF
The one [that bit the pig] died.
- (31) *Vok ah htui (ci) ah ui sì(k) ci.
pig P bite NF P dog die NF
- (32) *Vok ah htui (ci) sì(k) ci.
pig P bite NF die NF

This suggests that the word *htuih* in (27) and (28) may be a noun rather than a verb. It is of some interest that in English non-clausal relative modifiers are also possible, as illustrated in the glosses given for (27) and (28) 'bitten by the dog'. These are also unavailable with

relatives on a transitive subject, because only passive clauses can be so reduced. The structures of (25), (27) and (28) will be as in (xxv), (xxvii) and (xxviii).



If so, sentence (28) also shows that headless noun phrases and the related distribution of *ah* are not confined to relative clause structures.

We conclude this discussion with a few examples taken from the 2001 translation of the New Testament (*Ng'thu-k'Thai*).⁴

- (33) [ami hnguh te ah] àihli (2:10)
 3P see past C star
 the star [that they saw]

⁴ Chapter and verse refer to the Gospel according to Matthew. As noted previously, the orthography has been supplemented to include vowel length and tone, as well as adjusted on occasion for consistency.

In (33) the focus of relativization is the object of the verb *hngu* ‘see’, understood to be *àihli* ‘star’. Since it is not the subject which is relativized, the verb appears in its stem II form *hnguh*, accompanied by the third person plural subject agreement particle *ami*. The relative clause in (33) would be (34) as an independent sentence, with the stem I form of the verb *hngu*, no preverbal person agreement particle, but the plural particle *gùì* following the non-future tense particle *ci*.

- (34) *àihli hgnu te ci gui*
 star see past NF PL
 they saw a star

Notice that the past tense particle *te*, unlike the non-future tense particle *ci*, may occur with either form of the verb. When it occurs with the stem I form, it co-occurs with *ci*.

A more elaborate structure appears in (35) which contains two relative clauses, one inside the other.

- (35) [[*ami nong-bà ah*] *hngumí k'chú-na tu bà ca*] *k'chàng* (5:32)
 3P divorce C woman marry P P NF-C man
 a man [that marries a woman [that has been divorced]]

In the larger relative clause in (35) the focus of relativization is the subject of the verb *k'chú-na* ‘marry’.⁵ Thus this verb appears in its stem I form just as in the corresponding independent sentence (36).⁶ Notice in (35) the common contraction of *ci ah* into *ca*.

- (36) *cun ah k'chàng noh [ami nong-bà ah hngumí] k'chú-na tu bà ci*
 that C man P 3P divorce C woman marry P P NF
 that man is marrying a woman that has been divorced

The smaller relative clause in (35) and (36) corresponds to the independent sentence (37).

- (37) *cun ah hngumí ami nong-bà*
 that C woman 3PL divorce
 that woman has been divorced

Sentence (37) is semantically a kind of passive. But syntactically *hngumí* ‘woman’ remains the object of the verb *nong-bà* ‘divorce’, so that the clause verb appears in its stem II form in (35) and (36).⁷

⁵ We have glossed *k'chú-na* with the single English verb ‘marry’, but it is in fact complex consisting of the noun *k'chú* ‘wife’ with the applicative suffix *-na*. It could be more literally glossed as ‘take as wife’. See Mang and Bedell (2006) for more details.

⁶ (36), as well as subsequent (37), (39), (40), (42) and (44), sound awkward with an indefinite subject. We supply the demonstrative *cun* ‘that, those’ to emphasize definiteness. This requires the particle *ah* when preceding a noun, again indicating that *ah* is not specific to the relative clause construction.

⁷ We have glossed *nong-bà* with the single English verb ‘divorce’, but it is in fact complex consisting of the verb *nong* ‘abandon, renounce’ followed by the particle *bà* ‘back, again’. It

A similar structure is seen in (38).

- (38) [[kǎ ni(k) ci ah] bii bi(k) ci ah] k'chàng gui (7:23)
 not good NF C acts do NF C man PL
 men [that do things [that are not good]]

Both relative clauses in (38) focus on the subject, and thus in both (39) and (40) the verbs appear in their stem I forms.

- (39) cun ah k'chàng gùi noh [kǎ ni(k) ci ah] bii bi(k) ci gui
 that C man PL P not good NF C act do NF PL
 those men do things [that are not good]
- (40) cun ah bii kǎ ni(k) ci
 that C act not good NF
 that act is not good

Notice that *bii* 'act' is syntactically a noun and *bi* 'do' is syntactically a verb in (38), (39) and (40).

The structures in (41) and (42) may be compared with those in (38) and (39).

- (41) [a-k'ní ah kǎ ng'theih ci ah] ghíng gui phùng (7:19)
 good C not bear NF C tree PL every
 every tree [that does not bear good things]
- (42) cun ah a-ghíng gùi a-k'ní ah kǎ ng'theih ci gui
 that C tree PL good C not bear NF PL
 those trees do not bear good things

The relevant difference is that *a-k'ní ah* 'something good' in (41) and (42) is not a relative clause like *kǎ ni(k) ci ah* 'that is not good' in (38) and (39), or *ni(k) ci (ah)* 'that is good' in (21) and (22). Rather it is a kind of nominalization parallel to the compound nouns *k'am-k'zòì* in (19) and *k'chàng-k'ní* in (23).

The structures in (43) and (44) may be compared with those in (33) and (34).

- (43) [ǎihli a ng'dáng lo-nák te ah] a-ching (2: 7)
 star 3S appear come-na past C time
 the time [that the star appeared]

could be more literally glossed as 'abandon, reverting to former (i. e. unmarried) status'. The verb *nong* does not overtly distinguish stem I and II forms. Furthermore, the 'passive' interpretation, involving a non-referential third person plural subject, is available only with a stem II verb form. Thus *nong* in (37) is stem II. See Mang (2006) for details of the occurrence of stem II verb forms in independent clauses. The verbal particle *tu* in (35) and (36) indicates that the action is done 'with others' (Jordan 1969, p. 280). Another example appears in (47) and (48).

- (44) cun ah a-ching ung àihli ng'dáng lo te ci
 that C time at star appear come past NF
 at that time a star appeared

The relevant difference is that while the relativized position in (33) is the direct object, that in (43) is not an argument but an adverbial. In such cases, in addition to the *ah* found in other relative clauses, the suffix *-na* (in its stem II form *-nák*) appears on the verb.

The structures in (45) and (46) may also be compared with those in (33) and (34).

- (45) [ami hnguh te] cuh
 3P see past D
 the one that they saw

- (46) hgnu te ci gui
 see past NF PL
 they saw it

Here the relevant difference is that in (33) the relative clause has an overt head noun but does not in (45). Similarly those in (35), (38), (41) or (43) could appear without an overt head noun so long as the context provides an interpretation for the absent head noun.

A final example is (47) and (48).

- (47) [kei on kā om tu ci] cuh (12:30)
 me with not be P NF D
 he [that is not with me]

- (48) kei on kā om tu ci
 me with not be P NF
 he is not with me

In (45) and (47) the determiner particle *cuh* 'that' has been supplied in order to emphasize the absence of a head noun. It is not grammatically required and does not appear in the earlier examples like (16), (18), (22), (28) or (30). Notice that the empty noun is interpreted differently in an independent sentence like (46) or (48) than as head of a relative clause in (45) or (47). This is also true of English 'he' in the glosses given for (47) and (48).

References

- Jordan, Marc. 1969 *Chin Dictionary and Grammar*. Mindat (privately printed).
- (Thang Ngai Om). 1999 *Caciim-k'Thài* (The New Testament). Bible Society of Myanmar.
- Bedell, George. 2000 'Agreement in k'Cho', presented to the 33rd ICSTLL, Bangkok and Trang.
- (Roman Catholic Parish of Mindat). 2001 *Ng'thu-k'Thài* (The New Testament). Mindat (privately printed).
- Nolan, Stephen. 2002 'Spelling and Alphabet in 'Cho'. *ICU Asian Cultural Studies* 28:127-138.
- Nolan, Stephen. 2003 'Verb Alternation in K'Chò', presented to 36th ICSTLL, Melbourne.
- Kee Shein Mang and George Bedell. 2006 'The Applicative Suffix *-na* in K'cho'. (not published)
- Kee Shein Mang. 2006 *An Explanation of Verb Stem Alternation in K'cho, a Chin Language*. Payap University MA Thesis.

WEST COAST BAJAU: SYNTACTICALLY ERGATIVE OR SYMMETRICAL VOICE?¹

Mark T. Miller

SIL International & University of Texas at Arlington
<mark_miller@sil.org>

1 Introduction

1.1 *The West Coast (WC) Bajau Language*

West Coast (WC) Bajau is an Austronesian language spoken by about 60,000 people who live along the western and northern coasts of Malaysian Borneo, in the state of Sabah. The variety of WC Bajau described here is spoken in Kota Belud District, which is a population and cultural center for the WC Bajau people.² WC Bajau is one of nine Sama-Bajaw [Bajau] languages, which together form a subgroup of the Western Malayo-Polynesian branch of Austronesian languages (Frawley 2003).³

1.2 *The Definition of ‘Voice’*

In this paper I provide evidence that WC Bajau has a SYMMETRICAL VOICE system.

Traditionally, ‘voice’ is understood as the derivation of an intransitive clause from an underlying transitive clause:

In the passive derivation, the underlying undergoer argument (U) becomes the derived intransitive subject and the underlying actor argument (A) is demoted to peripheral status.

In the antipassive derivation, underlying A becomes the derived intransitive subject and underlying U is demoted to peripheral status. The passive and antipassive derivations are formally marked, such as by a verbal affix (Dixon 1994:146).

These passive and antipassive derivations bear some resemblance to the so-called ‘focus’ systems of Philippine-type languages. In these languages (and western Austronesian languages in general), the morphology on the verb signals that the argument associated with a particular semantic role in the clause shall be the pivot (= nominative NP). For example, if the morphology on the verb ‘focuses’ the actor argument, the actor is the pivot of the clause; if the verb ‘focuses’ the patient argument, the patient is the pivot of

¹ Special thanks goes to Dr. Paul Kroeger, who made several helpful suggestions for revision of this paper.

² I conducted my fieldwork in Kota Belud over four visits: 1996-98, 2001, 2003-2004, and 2006, spending an estimated total of 22 months in the village area. For much of this time I was graciously hosted by a WC Bajau family. My initial research (1996-98) was conducted under the auspices of the Federal Government of Malaysia, and also under the abah State Museum. Subsequent research has been sponsored in part by SIL International.

³ Among these nine Sama-Bajaw languages, seven are spoken primarily in the southern Philippines, and one on the coasts of Sulawesi (Indonesia). Only WC Bajau has its population center on Borneo island.

the clause. In this sense, the Philippine-type ‘focus’ alternations are voice alternations, even though, as Himmelman (2002:12) notes, “it is controversial whether ‘focus’ alternations involve a reduction in valency.” I follow Himmelman in using the term ‘voice’ to refer simply to the “realignment between syntactic pivots and semantic roles” whether or not there has been a reduction in valency.

1.3 *Syntactic Ergativity or Symmetrical Voice?*

There is a general tendency in Philippine-type (and some other western Austronesian) languages for the undergoer (O) rather than the actor (A) to be the more frequent choice of syntactic pivot in transitive clauses. Many such languages have been analysed as syntactically ergative, where one of the undergoer voices is analysed as “the basic unmarked construction for transitive clauses”, and the actor voice construction is regarded as the derived antipassive form (Himmelman 2002:14). A number of the Sama-Bajaw languages have been analysed in this way,⁴ where undergoer-voice clauses are transitive (ergative) and actor-voice clauses are intransitive (antipassive).

There is an alternative to this ergative hypothesis, however, which both Himmelman (2002:14) and Ross (2002:25) identify as SYMMETRICAL VOICE. With symmetrical voice, either the actor or the undergoer of a clause is selected as pivot, and crucially, “both undergoer-voice and actor-voice clauses are transitive” (Ross 2002:24). As noted by Ross (2002:24) in his discussion of Philippine-type languages, a central issue here is whether the actor-voice clause is transitive or intransitive.

In this paper, I argue that WC Bajau has a symmetrical voice system. In WC Bajau, the undergoer argument in actor-voice clauses is treated as a core argument, not demoted to oblique status as is a defining feature of antipassive undergoers. In what follows, I provide several sources of evidence for this claim, in which I consider the morphosyntactic and referential properties of the undergoer NP in active-voice clauses, as well as the distributional properties of transitive actor-voice and undergoer-voice clauses in narrative text. First, however, I present some basic typological features of WC Bajau related to voice marking and pronominal case-marking and constituent order in the language.

2 Voice-marking and pronominal case-marking in WC Bajau

2.1 *Morphological marking on UV, AV and intransitive verbs*

Undergoer voice (UV) is the zero-marked (= non-affixed) transitive construction in WC Bajau. In the UV clause, the actor argument immediately follows the verb. The UV actor, though not the pivot of the clause, is regarded as a syntactically core argument because: (1) it is never marked as an oblique; (2) it can never be deleted from the clause; and (3) it is often highly topical. The pivot undergoer may occur either following the actor (VAU) or, more commonly, before the verb (UVA). An example of a UVA clause in the undergoer voice follows:⁵

⁴ See Trick (2006) for Sama Southern; Brainard & Behrens (2002) for Yakan; Gault (1999) for Sama Bangingi; and Walton (1986) for Sama Pangutaran.

⁵ Abbreviations: 1s.I first person singular (class I), 3s.II third person singular (class II), 3p third person plural, UV undergoer voice, DET determiner, AV actor voice, PAST past tense, PN proper noun, PERF perfect aspect, LNK linker, PASS passive, PREP preposition, PRT discourse particle, FUT future tense, NVOL nonvolitional, NEG negative.

- (1) Kui e boi ø-kerot Nisah.
 cake DET PAST UV-cut Nisah
 ‘Nisah cut the cake.’

Actor voice (AV) is marked by the nasal prefix N-. (The nasal assimilates to the place of articulation of the following consonant.) The predominant constituent order in AV main clauses is AVU:

- (2) Nisah boi ngerot kui e.
 Nisah PAST AV-cut cake DET
 ‘Nisah cut the cake.’

The morphological marking on intransitive verbs parallels that for transitive verbs. Intransitive root verbs are either zero-marked (as in UV transitive constructions) or they have the N- prefix (as in AV transitive constructions). Intransitive root verbs cannot alternate between these two forms, unlike transitive root verbs, where a choice between UV and AV is possible. Hence intransitive root verbs in WC Bajau are “obligatorily split” between zero-marked and N- marked forms, though it is unclear to what degree the split is semantically motivated.

2.2 Case-marking on pronouns

Unlike many Philippine-type languages such as Tagalog, WC Bajau lacks case marking on noun phrases. Pronouns, however, sort into two primary classes depending on their syntactic status in the clause. Class I pronouns are a mixed bag of verbal enclitics and free pronouns. Only a Class I pronoun may express the actor of an UV verb. (The same pronoun forms, when they attach to NPs, function as possessives.) Class II pronouns are free pronouns. Only a Class II pronoun may express the pivot argument of the clause, whether the undergoer of an UV clause or the actor of an AV clause. Either Class II or Class I pronouns are used to express the undergoer of an AV clause, depending on the semantic properties of the undergoer: Class II pronouns must be used for first or second person undergoers; either Class II or Class I is possible for third person undergoers; and only Class I is possible for non-human undergoers.

The set of oblique-marking pronouns is formed by the addition of the locative prefix *em-* to the Class II pronouns. (The *em-* assimilates to the place of articulation of the following consonant.) The three sets of pronouns are represented in the following table:

	Class I	Class II	<i>em-</i> + Class II (oblique set)
1st person sg.	=ku	aku	m-aku
1st person pl. (incl.) (excl.)	=-ti kami	kiti kami	eng-kiti eng-kami
2nd person sg.	=nu	kau	eng-kau
2nd person pl.	=bi	kam	eng-kam
3rd person sg.	=ni	iyō	m-iyō
3rd person pl.	gai	gai	eng-gai

Figure 1: *the West Coast Bajau pronoun system*

The UV and AV constructions may be summarized as follows:

	UV	AV
morphological marking	zero marked	nasal prefix
usual constituent order	UVA (more rarely VAU)	AVU (more rarely VAU, VUA)
pronominal expression of A	genitive (Class I)	Class II
pronominal expression of U	Class II	Class II or Class I

3 The morphosyntactic properties of the AV undergoer in WC Bajau

I will now show that the AV undergoer in WC Bajau behaves like a core argument in its morphosyntactic properties. Here we may contrast the undergoer argument in the AV construction with the (clearly oblique) actor argument in another construction in WC Bajau which I call the ‘traditional passive’.

3.1 *The actor of the ‘traditional passive’ and the actor of the UV contrasted*

WC Bajau has a ‘traditional passive’ construction marked by the verbal infix *-in-*, which is very productive on transitive verb roots. In the WC Bajau passive, the undergoer argument is the pivot just as it is in the UV clause, but the actor argument is demoted to oblique status. With verbs affixed by *-in-*, the actor argument (when it appears) is obligatorily preceded by the preposition (o)le’ ‘by’, as shown in (3):

- (3) Beluang e pan b-in-uka no [ole’ anak ni sioko]
 door DET PRT-PASS-open now by child 3s.I oldest
 ‘The door was opened by the oldest child.’ (kerabaw 032)

In fact, the demoted actor argument is optionally deleted from the clause. In contrast, the UV actor argument is never preceded by a preposition, and it can never be deleted (except that zero anaphora is possible when the UV actor argument is highly topical).

In verb-initial clauses, the le’ actor phrase has various possible orderings: it may occur just after the verb and prior to the undergoer pivot, as in (4a), but more commonly it occurs following the undergoer, as in (4b). It may even occur after an adverbial oblique, as in (4c).

- (4) a. ? Ai k-in-akan [le’ Kuzik] moto diing e dilaw.
 PERF -PASS-eat PREP Kuzik eye fish DET yesterday
 ‘The fish eye has been eaten by Kuzik.’
 b. Ai k-in-akan moto diing e [le’ Kuzik] dilaw.
 c. Ai k-in-akan moto diing e dilaw [le’ Kuzik].

Positions (4b) and (4c) are not possible for the actor argument of a basic UV clause, because in the UV clause the actor must follow directly after the verb, as shown below:

- (5) a. Ai \emptyset -kakan Kuzik moto diing e dilaw.
 PERF UV-eat PN eye fish DET yesterday
 ‘Kuzik ate the eye of the fish yesterday.’
 b. *Ai \emptyset -kakan moto diing e Kuzik dilaw.

The multiple ordering possibilities associated with the *le'* actor phrase are indicative of oblique behavior and contrast with the rigid verb-actor word order observed in the basic UV construction.

To summarize the contrast between the ‘passive’ actor and the UV actor:⁶

the ‘passive’ actor	the UV actor
is always introduced by the preposition <i>le'</i> , creating a prepositional phrase (PP)	is never introduced by a preposition
is often deleted from the clause	is never deleted from the clause except in cases of zero anaphora
follows the verb, and other clause constituents may intervene	immediately follows the verb, and no other clause constituents may intervene

The morphosyntactic properties of the ‘passive’ actor are typical of obliques, and provide evidence that in the ‘traditional passive’ construction, the actor has been demoted to oblique status.

3.2 *The actor of the ‘traditional passive’ and the undergoer of the AV contrasted*

Unlike the actor of the traditional passive, the AV undergoer is usually not marked by a preposition or other oblique marker. Where the option of an oblique undergoer in the AV does occur, it is limited to a handful of motion verbs, perception verbs, or emotion verbs where the undergoer is minimally affected. These appear to be lexically-driven processes, which in some cases involve shifts in meaning between oblique vs. non-oblique undergoers. Most WC Bajau transitive verbs, however, do not even allow the option of an oblique-marked undergoer.

Note that in some other Sama-Bajaw languages, oblique marking on the undergoer in actor voice constructions appears to be more predictable. For example, in Pangutaran Sama, oblique marking likely occurs on the actor-voice undergoer when the NP is pronominal, animate, and personal (Walton 1986:109). No such predictable criteria exist for determining oblique marking on the AV undergoer in WC Bajau.

Whereas the (oblique) actor argument of the WC Bajau passive may occur either before or after the pivot argument, the AV undergoer must appear just after the verb,⁷ and usually cannot occur to the right of any peripheral or oblique elements in the clause. This is shown in (6) and (7) below:

⁶ See the similar criteria used by Walton (1986:116-18) in his analysis of *-i-* verbs as passive in Sama Pangutaran (cited in Gault 1999:49).

⁷ An exception to this statement occurs in those cases where the actor intervenes between the AV verb and the undergoer. Most such instances that I have found come from elicited data rather than from naturally occurring text.

- (6) a. Azizy nembali kambing e kaang.
 PN AV-slaughter goat DET later
 ‘Azizy will slaughter the goat later.’
 b. ?? Azizy nembali kaang kambing e.
- (7) a. Jumel meli kuda’ e ta’ Nisah.
 PN AV-buy horse DET PREP PN
 ‘Jumel will buy the horse from Nisah.’
 b. ?? Jumel meli ta’ Nisah kuda’ e.

Note that when the passive is used in place of AV in these clauses, the actor PP constituent may occur after any peripheral or oblique element:

- (8) Kambing e akan s-in-embali kaang [le’ Azizy].
 goat DET FUT -PASS-slaughter later PREP PN
 ‘The goat will be slaughtered later by Azizy.’
- (9) Kuda’ e bineli ta’ Nisah [le’ Jumel].
 horse DET -PASS-buy PREP PN PREP PN
 ‘The horse will be bought from Nisah by Jumel.’

Finally, while the actor of a passive clause always has the option of being omitted, the undergoer of an AV clause is usually only omitted where the AV verb allows an ‘inherent argument’ interpretation, such as how the English activity verbs ‘eat’ and ‘smoke’ are often used without an explicit undergoer.

4 The semantic properties of the AV undergoer

So far we’ve been using morphosyntactic criteria to establish a distinction between ‘core’ and ‘oblique’. Even though on morphosyntactic grounds the WC Bajau AV construction could not formally be considered an antipassive, perhaps AV undergoers retain some functional characteristics that have otherwise been associated with antipassives. Cooreman (1994) identifies several “functional correlates of antipassive constructions” that have been found in a number of the world’s languages. These are:

- a. low identifiability of the O;
- b. incomplete or non-punctual aspect;
- c. low affectedness of the O.

In what follows, I draw examples mostly from WC Bajau narrative texts to show that, while the AV construction sometimes aligns with these functions, it is not limited to such functions.

4.1 ‘low identifiability of the O’

Cooreman’s first functional correlate of the antipassive is that the object (undergoer) is not highly identifiable, which typically means undergoers that are indefinite or non-referential (generic). The N- prefix in WC Bajau can be used for indefinite or non-referential undergoers. In these cases, the undergoer is sometimes deleted from the clause, as in (10):

- (10) Boi jo gai mangan, gai pan turi.
 PAST PRT 3p AV-eat 3p PRT sleep
 ‘After they ate, they went to sleep.’ (abu nawas 024)

The deletion of the object in AV clauses occurs most often with activity verbs which allow an ‘inherent argument’. This is clearly a lexically-driven process.

The AV undergoer in WC Bajau frequently encodes an indefinite but specific participant. Almost always this means that the undergoer is introduced for the first time in the discourse, whether or not it will have continuing importance in the text. Examples:

- (11) Gai moo suu’ kasa’ engko’ kendidip.
 3p AV-bring kerosene.lantern and matches
 ‘They brought a kerosene lantern and matches.’ (poon nunduk 009)

Finally, there are many text examples in which the AV undergoer is both definite and specific. Examples:

- (12) Ella=ni pan ng-endo’ kerabaw panut e.
 husband=3s.I PRT AV-take buffalo drift DET
 ‘Her husband took hold of the drifting buffalo.’ (kerabaw 075)

- (13) Dela e pan nambut iyo ta’ beluang...
 man DET PRT AV-receive 3s.II PREP door
 ‘The man welcomed him at the door...’. (asal namuk 039)

So it is apparent that the AV undergoer, while it often encodes an indefinite and/or non-specific argument, can also encode definite and specific undergoers. This contrasts with the typical antipassive construction.

As for the properties of the undergoer in UV clauses, the undergoer is only deleted when it is clearly identifiable by the context (zero anaphora). The undergoer in UV clauses is normally both definite and specific. However, there are instances where the undergoer is indefinite, as in the following examples:

- (14) ø-Ogo gai tabit ta’ kelinik kampung.
 UV-visit 3p healer PREP clinic village
 ‘They went to a health practitioner at the neighborhood clinic.’ (rupiah 012)

- (15) ... lanjang jomo ø-endo’=ni amun iyo nya’ te-kito
 rice.pot person UV-take=3s.I if 3s.II NEG NVOL-see
 ‘She took people’s rice pots when she could not be seen.’ (kerabaw 027)

In (14) the undergoer (tabit ‘healer’) is indefinite and specific. In (15) the undergoer (lanjang ‘rice pots’) is indefinite and questionably specific. From these examples it is apparent that, while UV is strongly correlated with a definite and specific undergoer, it is not an absolute correlation. The AV construction, on the other hand, typically accommodate a range of undergoers, whether highly identifiable or not identifiable.

4.2 *'incomplete or non-punctual aspect'*

Cooreman's second functional correlate of the antipassive is non-punctual or repeated action. Imperfective (durative, habitual, or iterative) aspect in WC Bajau is normally expressed by reduplication of the verb, whether AV or UV. Where non-reduplicated AV verbs do convey imperfective aspect, this is most likely due to the tendency for N- derived activities (e.g., those having an inherent argument) to be habitual or durative in their interpretation.

4.3 *'low affectedness of the O'*

Cooreman's third functional correlate of the antipassive construction is that there is no lasting effect on the object. She identifies for Chamorro a 'Demoting Antipassive', where the object is definite but morphologically marked as an oblique. She notes that only those verbs in Chamorro "which do not imply a lasting effect on the Object" allow the use of the Demoting Antipassive.

As was mentioned earlier, in WC Bajau, a handful of verbs allow for the optional demotion of a specific undergoer to oblique status in the AV, though never in the UV. These verbs, when they take the oblique argument, usually seem to entail that the undergoer is minimally affected by the action. They include the perception verb *ngenda* 'look at', the emotion verb *ngintam* 'miss; long for', and the motion verbs *nuut* 'go with; follow', *ngogo* 'go to; visit', and *meniik* 'ascend; ride'. Typically the oblique undergoer is a locative argument of some kind, and is sometimes deleted when the referent is clear from the context.

Some such AV verbs allow a choice between an oblique and a non-oblique undergoer, and for such verbs, there may be subtle or not so subtle shifts in meaning with this distinction. Consider the verb *enda*, *ngenda* which can mean either 'look at' or 'look for'. Both meanings are possible in both voices (AV and UV). However, in AV clauses the undergoer is normally oblique with the meaning 'look at', and is never oblique with the meaning 'look for'. Example:

- (16) a. *Iyang=ku ngenda' m-iyoy.*
 mother=1s.I AV-look PREP-3s.II
 'My mother looked at/ *for him.'
- b. *Iyang=ku ngenda' iyoy.*
 mother=1s.I AV-look 3s.II
 'My mother looked for / (?) at him.'

Arguably, the action of 'looking for' (with the non-oblique undergoer) has a greater potential effect on the undergoer than the action of 'looking at' (with the oblique undergoer). But it is difficult to identify meaning changes like these within some larger pattern observed for AV verbs. They seem instead to be lexical processes, rather than governed by a regular 'demoting antipassive' construction.

5 The distribution of AV vs. UV clauses in texts

In order to determine the distribution of AV vs. UV clauses in texts, a count was performed on seven WC Bajau narrative texts. For each semantically transitive clause, it was determined:

- 1) whether UV, AV, or the ‘true passive’ voice was used;
- 2) for AV clauses, whether the clause was syntactically intransitive (oblique or non-expressed undergoer) or transitive (non-oblique undergoer);⁸
- 3) for both UV and syntactically transitive AV clauses, the referential status of the undergoer.

Note that, in considering semantically transitive clauses, it was necessary to exclude cases where the AV (or UV) was grammatically determined, as often happens with dependent clauses, questions, and clefts (see Payne 1994:328-30). The percentages of clause types and AV undergoer types are compared with the “bench-mark” figures established by Cooreman (1987) for Chamarro, as reported in Givón (1994).

The results of the distribution studies of voice type, and the undergoer types of both AV and UV clauses, are indicated in Figures 2, 3, and 4 below.

voice type	number of clauses	percentage
syntactically transitive AV clauses	57	39.86
UV clauses	53	37.06
‘traditional passive’ (-in-) clauses	33	23.08
totals:	143	100.00

Figure 2: *Distribution of AV, UV, and passive clauses in the seven texts*

Figure 2 indicates that the seven texts contained nearly the same number of syntactically transitive AV clauses (57) and UV clauses (53), which clearly suggests a symmetrical voice pattern in that the two voice constructions ‘share the load’ in the expression of transitive clauses.

status of AV undergoer	number of clauses	percentage
oblique	11	12.94
no undergoer / inherent argument	17	20.00
indefinite and/or non-specific	16	18.82
definite and specific	41	48.24
totals:	85	100.00

Figure 3: *Status of AV undergoers in clauses in the seven texts*

Figure 3 shows that the number of definite and specific AV undergoers (41) is nearly one half of the total, indicating that, while AV can be used for non-specific/ generic undergoers and for newly identified (indefinite) undergoers, it is equally likely to be used for highly individuated and previously mentioned undergoers. Note that, whereas Cooreman found with the Chamarro antipassive a high percentage (61.4%) of non-anaphoric zero patients, in WC Bajau only 20% of AV undergoers were non-anaphoric zero arguments.

⁸ Cases of zero-anaphora for the AV undergoer were counted as syntactically transitive clauses. The discourse context usually made clear whether the absence of the undergoer was due to zero-anaphora or to unspecified undergoer deletion.

Taking Figures 2 and 3 together, the proportion of AV clauses in the WC Bajau texts is 85 out of 143, or 59.44%, a much higher percentage than that reported by Cooreman (1987) for Chamarro antipassives, which is only 7.7%.

status of UV undergoer	number of clauses	percentage
definite and specific	50	94.34
indefinite (newly-introduced) and specific	3	5.66
totals:	53	100.00

Figure 4: *Status of UV undergoers in clauses in the seven texts*

Figure 4 shows that the great majority (94.34%) of UV undergoers are both definite and specific. Taken together with Figure 3, we see that of the 91 tokens of definite and specific undergoers in the seven texts, AV was used 45.05% of the time (41 of 91) and UV was used 54.95% of the time (50 of 91). Thus, AV was nearly as likely as UV to express a definite and specific undergoer.

These results clearly show that, from a distributional perspective, the AV construction in WC Bajau looks very unlike the antipassive in Chamarro. Its high overall distributional frequency, as well as the likelihood of an AV verb to express a definite and specific undergoer, make the AV construction sometimes function parallel to the UV construction. The difference, of course, is that the AV construction is more versatile in being able to accommodate generic/non-specific undergoers, inherent arguments, and oblique undergoers.

6 Conclusion

It is clear from both morphosyntactic and distributional evidence that the AV construction in WC Bajau is often (though not always) used transitively. If so, then we would be hard pressed to conclude that WC Bajau is syntactically ergative, with the AV corresponding to an antipassive construction. The preferred analysis is to posit two transitive voices, UV and AV, with the AV optionally allowing, with certain verbs, for the demotion or deletion of the undergoer to derive an intransitive construction.

The symmetrical voice analysis raises several interesting questions. If both UV and AV may be transitive, what motivates the selection of one voice over another? Topical continuity may have an important role here, as has been demonstrated in other symmetrical-voice languages like Pendau (see Quick 2003) and Balinese (see Pastika 1999, Arka 2003). In fact, for both of these languages, application of Givón's (1994) and Dryer's (1994) methods for measuring topical continuity has shown that the topicality of the undergoer is a significant factor in determining voice selection. I predict that the same is true for WC Bajau, but this needs to be empirically verified.

I find it interesting that, even though the 'traditional passive' involves a syntactically demoted argument (the agent), it still occurs with a relatively high frequency (23% of all semantically transitive clauses in the present study). While the passive is usually associated with low actor topicality, the relatively high frequency of the passive in WC Bajau narrative text suggests to me that other selectional criteria may be involved.

References

- Arka, I Wayan. 2003. *Balinese morphosyntax: A lexical-functional approach*. Pacific Linguistics 547. Canberra: The Australian National University.
- Brainard, Shari and Dietlinde Behrens. 2002. *A grammar of Yakan*. Manila: Linguistic Society of the Philippines.
- Cooreman, Ann M. 1987. *Transitivity and discourse continuity in Chamarro narratives*. Berlin: Mouton de Gruyter.
- Cooreman, Ann M. 1994. A functional typology of antipassive. *Voice: form and function*, ed. by B. Fox and P. Hopper, 49-86. Amsterdam: John Benjamins.
- Dixon, R.W.M. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, Matthew. 1994. The discourse function of the Kutenai inverse. *Voice and inversion*, ed. by T. Givón, 65-99. Amsterdam, Philadelphia: John Benjamins.
- Frawley, W. (ed.). 2003. Sama-Bajaw languages. *International Encyclopedia of Linguistics 3*, Second edition, 496-8. Oxford: Oxford University Press.
- Gault, JoAnn Marie. 1999. *An ergative description of Sama Bangingi'*. Manila: Linguistic Society of the Philippines.
- Givón, Talmy. 1994. The pragmatics of de-transitive voice: Functional and typological aspects of inversion. *Voice and inversion*, ed. by T. Givón, 3-44. Amsterdam, Philadelphia: John Benjamins.
- Himmelman, Nikolaus P. 2002. Voice in western Austronesian: an update. *The history and typology of western Austronesian voice systems*, ed. by Fay Wouk and Malcom Ross, 7-16. Canberra: Pacific Linguistics.
- Pastika, Wayan. 1999. *Voice selection in Balinese narrative discourse*. PhD dissertation, The Australian National University.
- Payne, Thomas E. 1997. *Describing morphosyntax*. Cambridge: Cambridge University Press.
- Quick, Phil. 2003. *A grammar of the Pendau language*. PhD dissertation, The Australian National University.
- Ross, Malcom. 2002. The history and transitivity of western Austronesian voice and voice-marking. *The history and typology of western Austronesian voice systems*, ed. by Fay Wouk and Malcom Ross, 17-62. Canberra: Pacific Linguistics.
- Trick, Douglas, 2006, *Ergative control of syntactic processes in Sama Southern*. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines. Available from <http://www.sil.org/asia/philippines/ical/papers.html>
- Walton, Charles. 1986. Sama verbal semantics: classification, derivation and inflection. *Manila: Linguistic Society of the Philippines*.

CASES OF LINGUISTIC ASSIMILATION AND ETYMOLOGICAL DOUBLETS IN JAVANESE WORDS OF ARABIC ORIGIN

by Rochayah Machali
University of New South Wales
<r.machali@unsw.edu.au>

1. Introduction

Javanese is spoken in the island of Java, but not all who live in Java speak the language. It is spoken only in Central and East Java, with the Central Javanese variety being considered the standard. The Javanese constitutes the largest ethnic group in the Indonesian population, and it is estimated that 40 per cent of the Indonesian population speak Javanese (according to the UNESCO report in 2002, the population of Indonesia has reached 217 million). In addition, due to transmigration, Javanese is also spoken in other islands, such as in South Sumatra. Beyond Indonesia, Javanese is spoken in Surinam and New Caledonia.¹⁾

The language is a member of the Austronesian family of languages, which includes the languages of the Philippines, Malaysia, Indonesia, Madagascar, and the Pacific (Robson 2002:3). During the course of its history, the language has undergone several phases of development, reflecting specific idiomatic features and perhaps determined by the differing cultural milieus of Javanese writers (Koentjaraningrat 1985).

There are six phases in its development (Koentjaraningrat, p. 12). The first phase is the Old Javanese, which is the language of the ancient royal inscriptions of the eighth to the tenth centuries carved on stone or bronze tablets, and of ancient literary works from the tenth to fourteenth centuries. The second phase is the Old Javanese of the so-called 'Javanese-Balinese literature', which was a phase covering the 14th century of the Hindu-Javanese until the start of Islamic civilization in East Java.

The third phase is the so-called 'East-Javanese Muslim literature', which spanned over the 16th to the 17th century. It was during this time that Arabic borrowings had started. Since the spread of Islamic civilization covered the Brantas basin of East Java to the solo river of Central Java, early evidence of Arabic borrowing in Javanese would presumably be prominent in these areas as well.

The Fourth phase indicates the spread of Islam, thus the proliferation of the use of Arabic in some way. This phase is known as the 'Language of the Coastal Javanese Muslim Culture', which flourished in the religious centres of the north Javanese coastal towns. The language development in this phase happened in 17th and 18th centuries. The fifth phase of development is the 'Literary Language of Mataram', which occurred in the 18th and 19th centuries, as evidenced by the writings of the court poets of the Kingdom of Mataram, located in Central Java. The sixth phase is referred to as the 'Javanese Language of Today', which is used in the daily conversation and present-day Javanese books and newspapers.

¹⁾ Robson (2002:iii) estimated that there are about 1 million speakers of Javanese.

Paul Sidwell & Uri Tadmor, eds. *SEALSXVI: papers from the 16th meeting of the Southeast Asian Linguistics Society* (2006). Canberra, Pacific Linguistics, 2008, pp.47-63.

© Rochaya Machali

It is generally not possible to date exactly when a word enters the linguistic inventory of a language, but in the case of Arabic loanwords, they have certainly been in the Javanese language for a long time. Although the words have made their way into Javanese through the spread of Islam, many of the words are not necessarily of Islamic religion. An examination of two recent Javanese articles on the Internet reveals that even in an article on ‘erotism’, three per cent of the words used are Arabic loanwords while one on ‘Sufism’ has seven per cent words of Arabic origin.²⁾

It is important to note here that when the phrase ‘Arabic language’ is used, it refers to the classical Arabic, not to its colloquial variety. As to the Javanese, only words of the ‘ngoko’ level are used as entries, since this level is considered neutral and does not contain interpersonal component of meanings as do other levels.³⁾

2. Loanwords: Javanese words of Arabic origin

Loanword has been defined as ‘a word adopted, often with some modification of its form, from one language into another’ (Collins English Dictionary, 1979). On his examination of Javanese loanwords in the Indonesian language, Poedjosoedarmo (1982) has found that borrowings may take the form of direct borrowing, loan-blends, loan-shifts, or loan translation. Similar categories have been found in Arabic loanwords in Javanese, in varying degrees and rates. However, since the sound systems of Indonesian and Javanese languages differ significantly, borrowings take different patterns of assimilation although they are of similar categories as those in Indonesian. It is the assimilation that becomes the focus of discussion in the paper. So, the aim of this paper is twofold: (1) sound assimilation in loanwords and its patterns; (2) lexical adaptation in loanwords and its pattern.

Of the four categories mentioned above, direct borrowing constitutes the largest in number and has been assimilated into Javanese in varying degrees, which is an indication of different ‘chronological layers’ (Campbell, 1996). As a result, ‘etymological doublets’ may occur. However, Campbell’s observation may not necessarily be true for loanwords in Javanese, since not all etymological doublets are cognates. While they may have the same etymon, they denote distinct concepts that uniquely Javanese.

Before the linguistic assimilation is analysed, it is necessary to explain first of all how the list of loanwords have been obtained (see Methodology below). Once a list has been obtained, patterns of sound assimilation and morphological adaptation are identified and analysed.

3. Methodology

The method of finding loanwords was carried out in the following procedures. The first procedure involved the arduous task of scanning the Javanese-English dictionary (Horne 1974) and the (more) practical Javanese-Indonesian Dictionary (Partaatmadja 1992). A

²⁾ “Pamomong” section of the “Suara Merdeka” newspaper: 11/01/2004 and 16/11/2003 editions respectively.

³⁾ The Javanese language has three speech levels (called social styles by Horne). The highest level is called *Krama*, the middle *Madya*, and the lowest (also considered the neutral) is called *ngoko*. The vast majority of words are neutral (ie of *ngoko*) level, but a thousand or so words are restricted to particular situations defined by the relationship between speakers and the people they are talking about. Thus the word *aku* ‘I’ has a formal counterpart (ie *Krama form*) *kula* which also means ‘I’.

multistage culling followed this first procedure. The main purpose for using both dictionaries was mainly for checking any discrepancy of entries that may have occurred in the two periods of the publications.

The scanning was carried out in a rather complicated way. The process takes the following steps: (1) scanning the Horne dictionary for the loanwords; (2) at the same time, the loanwords were glossed alongside the words. The listing and glossing follow the system used by Jones (1984, 1996) in listing Arabic loanwords in Indonesian (in 1996 I was involved in updating Jones' list against the most recent edition of *Kamus Besar*, the Indonesian Dictionary). The (Jones) list has the following numbered elements for each entry: (1) the Arabic loanword itself; (2) its meaning in English; (3) the transliteration as it appears in Wehr; (4) page reference in Wehr; (5) the phonemic representation of the Arabic word as it appears in Wehr. The dictionaries (Javanese and Arabic) are described below.

3.1 Reference dictionaries and List

(A) The Horne dictionary: Javanese-English (henceforth 'JDI')

The Javanese-English dictionary was compiled by Horne and was published in 1974. According to Horne (pp. ix-x), it is intended to be a general-purpose dictionary of Javanese as used by the educated urban Javanese speakers from Central Java, considered as the area of the standard language. There are two lists that become the source materials for the dictionary: one representing the pre-Revolution lexicon and the other represents the post-Revolution.

Horne compiled the first list from entries in two Javanese dictionaries: by Pigeaud (1938) and by Poerwadarminta (1939). The second list came from materials that she collected: (a) oral materials (monologues and dialogues) tape-recorded by Javanese speakers; (b) books and periodicals published in Java for the Javanese on a variety subjects. The materials, wrote Horne, totalled nearly a million words of running text, which were then processed by computer into concordances, and the final list was obtained. Horne estimated that the two lists have eighty per cent overlap.

Horne explained further that both lists were examined by at least three Javanese speakers. Using the second list as a check, words considered not common use in the second list were dropped; new words and new usages for old words were compiled in the second list. In cases of disagreement among speakers, the word or meaning in question was accepted after consulting other speakers, due to the fact that we cannot expect every speaker to know every word and every meaning.

(B) The Partaatmadja dictionary: Javanese-Indonesian practical dictionary (henceforth referred to as 'JDII')

Like Horne, Partaatmadja (1992) also based his entries on the dictionary by Poerwadarminto. However, while Horne was very explicit in her methodology of creating the entries, not so with Partaatmadja. Apart from reference to the earlier dictionary, no other materials sources have been mentioned. The overlap between the two Javanese dictionaries used in this paper is approximately five per cent, with more recent words included in JD II non-existent in JDI, for example *ambal* 'carpet', perhaps pertinent to 'modern' living in Java,

As JDII is meant to be for non-Javanese Indonesians, it is designed for general readership with brief, less complicated entries than JDI. Also, there is no specific mention of it being based on the 'standard' Central Javanese variety. In this way, entries maybe expected to

originate from East Java as well. For this reason, entries in JDII are meant to complement those obtained from JDI.

(C) The Wehr dictionary: Modern Written Arabic (henceforth ‘the Wehr’)

The dictionary (the 1980 edition) presents the vocabulary and words of modern written Arabic. It is based on the language form found in books, newspapers, periodicals and letters, a form that is also used in formal public address, radio and television, and in religious ceremonies. According to Wehr, the morphology and syntax of written Arabic are essentially the same in all Arab countries, with vocabulary differences limited mainly in the specialised vocabulary.

The vocabulary, therefore, is standardised and not limited in its scope and time (Wehr, p. ix). As such, the dictionary is useful for checking or double-checking meaning found in the Jones List.

3.2 Listing and Culling Procedures

It has been stated in the introduction that Arabic borrowing dates back to the distant past, as early as the end of the 14th century.⁴⁾ The long history of borrowing can result in varying degrees of assimilation, for example there are etymological doublets (even triplets) that share the same meaning: *sekabat*, *sobat* and *sahabat* all refer to the idea of ‘(close) friend’. A decision needs to be made for cases like this, ie (a) by selecting the main entry to be on the list, and (b) by placing the rest as cognates. This was why a listing procedure was administered to the list of loanwords obtained from the scanning. The listing procedure involves the following:

1. Scanning JDI to create a tentative list; only the *ngoko* form is included. So, while the word *kitan* ‘circumsicion’ is listed, its honorific (*krama*) form *selam* is culled. Derivatives were placed in a separate column, eg the verb form *nyalati* was placed alongside the base form *salat* ‘prayer’. Names were excluded from the list, eg *Darul Islam* (name of an organization in Indonesia).
2. The tentative list was then checked against JDII, mainly for updating the list, for example *waswas* ‘anxiety’ was added to the list, a word not found in JDI. The possible explanation for its absence may be due to the fact that such word is more used in East Java.
3. The resulting list was checked again to separate etymological doublets or triplets that are not cognates.
4. A final list of 565 entries was obtained and was then categorized and analysed for patterns of sound assimilation and morphological adaptation, presented in Section 4 below. Before this categories and patterns are presented, the phonetic framework for both Javanese and Arabic need to be compared.

3.3 Phonetic Framework

Javanese and Arabic sound systems differ greatly in terms of the point and manner of articulation. While Javanese has 20 distinct consonants in its inventory, Arabic has 30

⁴⁾ According to Sneddon (2003:74) Arabic borrowing in Indonesia even dates back to the end of the 13th century, when Islam was spread in the west of the archipelago. Evidence of such spread can be found in the Trengganu inscription from the Malay Peninsula and the Minye Tujuh gravestone from Sumatra in the 14th century.

distinct ones (Uhlenbeck 1978; Newman 2002). Differing from Uhlenbeck, who lists 21 consonants, only 20 are listed here for reasons explained below.

3.3.1 The sounds of Javanese

The following Table is adapted from Uhlenbeck (1978:24).

Table 1: *Javanese consonant inventory*

		labial	palato-alveolar	dento-alveolar	palatal	velar
closing sounds	intensive	p	T	t	c	k
	not intensive	b	D	d	j	g
nasal		m		n	n+	N
sibilant				s		
liquids				l, r		
semi-vowels		w			y	
aspirant						h

Differing from Uhlenbeck, the [t] and [d] with raised tongue tip are here placed under ‘palato-alveolar’ rather than as dental. Likewise, [q] is deleted from the inventory, because the data (from the two Javanese dictionaries examined) do not indicate it as an independent phoneme. In fact, Uhlenbeck himself was not sure whether or not [k] and [q] are independent phonemes. The only occurrence of [q] is ‘Qur’an’ but even this has a co-entry ‘Kuran’ in the dictionaries.

As to the vowel inventory, Uhlenbeck lists six distinct vowels, as listed below.

A	O	E
U	I	e+

So, there are A,O,E,U,I and a schwa-vowel, represented as /e+/ here. The first five has two allophones each, which are categorized as those with ‘grave accent’ and those with ‘acute accent’. It would suffice here to mention just this at this stage. They will only be referred to where necessary in Section 4 below.

3.3.2 The sounds of Arabic

Newman (2002) lists 30 phonetically distinct consonant segments in classical Arabic, as presented in Table 2 below. He also lists six vowels in a separate Table. He has warned readers that vowel inventory of colloquial Arabic varieties would differ considerably. However, for purposes of comparison with Javanese vowel system, it would suffice here to use the classical vowel system for contrasting it with the Javanese. In the table below, the pharyngeal signs are changed from Newman’s, which is mainly due to availability of computer fonts.

Table 2: *Arabic consonant inventory*

	bi-labial	labio-dental	dental	dento-alveolar	post-alveolar	palatal	velar	uvular	pharyngeal	glottal
plosive	p		t, d			j	k	q		/
nasal	m			n						
trill				r						
tap				R						
fricative		f	T, D	s, z	S			X, “	H, +	h
approximant						y	w			
lateral				l						

The above phonemes, plus: Pharyngealised consonants: t+, d+, s+, D+, l+ and the affricate dZ

Table 3: *The vowels of Arabic*

	Front	Central	Back
Close	ī, ĩ		u, ū
Open		a, ä	

As can be seen here, the vowels of Arabic are distinguishable in terms of a six-vowel system, with three different vowel qualities for each of the vowel. The vowels are all voiced and oral, and no diphthongs are found.

3.3.3 Arabic sounds non-existent in the Javanese sound system

On comparing Tables 1 and 2, we can see that there are a few striking differences between the two systems. Firstly, in terms of the Point of Articulation, the Arabic phonemes non-existent in the Javanese consonant system are: f, T, D, R, S, y, w, q, X, “, /, the Pharyngeal consonants as well as the affricate /dZ/. The most important implication of such difference is that words borrowed from Arabic would be assimilated into the Javanese consonant system. Patterns of assimilation are presented below in Section 4.

In terms of the vowel system, there are three vowels that correspond in both Arabic and Javanese: a, i, u. However, the Arabic sound quality inherent in each of the three vowel phonemes is non-existent in the Javanese vowel system. In table 3, it can be seen that the sound quality is realized as intensification or lengthening of the vowel. Although a similar case of vowel intensification occurs in Javanese, it is a semantic case rather than phonetic. Therefore, loanwords from Arabic having the three intensified vowels would be assimilated in the Javanese vowel system. In some cases, the three sounds are weaker and not as oral, depending on the sound structures of the loanword. Generally, the Arabic intensive vowels are pronounced as short and non-intensive sounds in Javanese, as in *adan* for [adhän] ‘call for prayer’, *hurup* for [h.uruf] ‘alphabet’, and *iklim* for [iqlim] ‘climate’.

It should be noted, however, that the non-existence of Arabic sounds in Javanese may also result in direct borrowing, thus unassimilated sound. Furthermore, although there are sound correspondences in Javanese and Arabic, their environments of occurrence would differ in the two. For example, certain consonant clusters may occur finally in Arabic, but not in Javanese. Thus while we can say [dhikr] in Arabic, it will be [dikir] in Javanese (see Transliteration system in Appendix B). So, ideally, there should be a

comparison of sound environments, for example, which sound structures can occur initially, medially or finally in the two languages. Also, sound environments are discussed for explaining (radical) changes of word forms when they occur.

4. Sound Assimilation in direct borrowing

There are two parts in this section: 4.1 sound assimilation involving consonant changes; and 4.2 sound assimilation involving vowel changes. In the Tables on consonant changes (4.1), vowel changes are also indicated as notes; then the general patterns of vowel change are presented in 4.2.

4.1 Consonant changes

The largest part of the borrowing takes the form of direct borrowing, where words have been borrowed without change of meaning or with slight change of meaning nuances. Although many of the words are directly borrowed, they often have to be phonologically or morphologically recast to suit the Javanese sound system. In what follows, the focus is on phonetic recast in the loanwords.

As mentioned earlier, the consonant inventories between the two languages differ greatly in terms of the pharyngeal sounds and certain other consonants as well as in terms of the Arabic vowel quality that are non-existent in Javanese.

Below are examples from the data of how non-existent Arabic sounds may be realised in the words (Transliteration of sounds is presented in the Appendix; here it is square-bracketed). The phonemes selected are not in particular order, and where possible all three occurrences (initial, medial, final) are shown in the examples.

4.1.1 The uvular sounds

The Arabic uvular fricatives /X/, /ʕ/ and the uvular plosive /q/ are commonly pronounced as [k] in Javanese. All occurrences of /q/, except for that in ‘Qur’an’, are realised as [k] sound in Javanese. Their occurrence in the word can be initial, medial or final. As for /X/, it is sometimes also pronounced as [h], and there is also a case of full phonetic recast in the case of /ʕ/, as in *setiyar* from [ikhtiyar] ‘own option’. Note: → means changes into.

Table 4.1.1: Sound changes involving uvular sounds

Sound change	Loanword	Original (in Wehr)	gloss	Notes
/X/ → [h]	haram	[h.aräm]	‘forbidden’	Thick and long vowel in the original
	rahmat	[rah.ma]	‘god’s mercy’	Also involves vowel change [o] → [a] and adding [t]; as well, meaning is more specific than its original
	sah	[s.ah.h.]	‘legal’	
/X/ → [k]	kurmat	[h.urmat]	‘respect’	
	mokal	[muh.äl]	‘impossible’	Also vowel change (see Section 4.2)

/ʔ/ → [k],	kotbah	[khut.ba]	‘sermon’	Also involves vowel change and another consonant change grave [t] → light [t] Also the generic change of [q] → [k] in the final position Also involves the vowel change [u] → [a] and the addition of [h] in the loanword
	makluk	[makhlüq]	‘creature’	
	naskah	[nuskha]	‘writing; copy’	
/q/ → [k]	kamus	[qämüs]	‘dictionary’	Also accompanied by vowel changes (see vowel change in Section 4.2)
	akal	[‘aq̣l]	‘intelligence’	
	mutlak	[mut.laq̣]	‘absolut’	

The pharyngeal fricative /G/ is predominantly pronounced as [g], except for bihal’ mule ’when /G/ is pronounced as [h]. Some examples:

Loanword	Original (in Wehr)	gloss	Notes
galib	[ghälib]	‘usual’	
magrib	[maghrib]	‘sunset’	
mubaligh	[muballigh]	‘preacher’	The thick ‘l’ in the original becomes lighter in the loanword

4.1.3 The glottal sounds

The glottal plosive /// in Arabic is realised in most interesting ways in Javanese (transliterated as [‘]). The most common way is to recast the sound as fronted vowels in Javanese, as in adil ‘just’, ijajil ‘devil’ (both occur initially). In other cases, however, the sound is recast as velar nasal /N/ in Javanese, as in:

Loanword	Original (in Wehr)	gloss	Notes
ngahad	[ah.ad]	‘Sunday’	
Dulkangidah	[dhü,l-qa’da]	‘the Islamic 11th month’	

4.1.4 The dental sounds

The dental fricative /D/ and the pharyngealised /D+ /and /d+ / are most commonly realised as [d] in Javanese, but also as [l] in some cases, as shown in the examples in Table 4.1.4.

Table 4.1.4: Sound changes involving dental sounds

Sound change	Loanword	Original (in Wehr)	gloss	Notes
/D/ → [d]	dikir	[dhikr]	‘invocation of God’	Also involves vowel change (see Section 4.2)
	adan	[adhän]	‘call to prayer’	
/D+ / → [d],	kadi	[qäd.ī]	‘judge’	
/d+ / → [l]	taklim	[ta’z.īm]	‘great respect’	
	luhur	[z.uhr]	‘midday’	

In this case, recasting the sound as [l] becomes compulsory when /d+/ occurs in final position, as in *lapal* and *aral* below, as such (voiced and thick) sound would not normally occur in final position in Javanese, as in [lapal] ‘spoken word’; [lila] ‘sincere’; [aral] ‘hindrance’.

4.1.5 Sound changes involving the fricatives

The post alveolar fricative /S/ and the pharyngealised fricative /s+/ are ‘neutralised’ as [s] or by retaining the accented fricative sound as [sy] in Javanese. The fricative dental phoneme /T/ is also pronounced [s] in the few loanwords found. as in:

Table 4.1.5: *Sound changes involving the fricatives*

Sound change	Loanword	Original (in Wehr)	gloss	Notes
/S/ → [s]	sahid	[shahīd]	‘religious martyr’	Also involves vowel change (Section 4.2)
	mesgul Kures	[mashghül] [Quraish]	‘sad’ ‘name of tribe in Arabia’	
/s+/ → [s]	sabar	[s.abr]	‘patience’	
	tasawup	[tas.awwuf]	‘sufism’	
	kusus	[khus.üs.]	‘special’	
/T/ → [s],	Selasa	[thalätha]	‘Tuesday’	Meaning is more specific meaning than its original Also the consonant change [h.] → [k]
	isbat	[ithbät]	‘assert with parables’	
	kadas	[h.adath]	‘ritual impurity’	
/S/ → [sy],	masya Allah	[mä shä,a ,lläh]	‘Good Lord!’	Also involves vowel change (Section 4.2)
	syarif	[sharif]	‘descendant of Mohamad’	
	absyah	[afs.ah]	‘legal by Isl. law’	(differing from absah ‘valid’); also involves the change of [f] → [b], see 4.1.6 below.

4.1.6 A note on the the labio-dental fricative

The labio-dental fricative /f/ is commonly pronounced as [p], except for a few words such as *fajar* ‘dawn’, *wafat* ‘pass away’, *maaf* ‘pardon’. In fact, no entry under [f] is found in JDII. Their entry in JDI is most probably influenced by the fact that they co-occur in Indonesian and are probably used as ‘Indonesian borrowing’ in Javanese.

The above patterns of assimilation are rather simplified, aimed mainly to show how non-existent sounds are compensated or recast in the borrowed word. In fact, even existent sounds are slightly different in pronunciation, for example while the Arabic /h/ is glottal, it is velar in the Javanese sound system.

On examining the above examples, it can be seen that many of the non-existent sounds can, in principle, occur in any position in the word although rarely in the final

position. This maybe due to the ‘voiced’ nature of the Arabic sounds, which in Javanese would be voiceless. This explains why the pharyngealised /d+/, commonly pronounced as [d] becomes [l] in final position, as in the case of aral ‘hindrance’ above. In other words, the position of the phoneme in the words may cause a variation in the form of the assimilated sound.

4.2 Vowel change: general patterns

It has been mentioned earlier that non-existent vowels also undergo assimilation in the loanword. There are commonly two kinds of assimilation: by pronouncing the Arabic long vowels as short, as in hurup for [h.uruf] ‘alphabet’ or by pronouncing them as weak sounds. The weak sounds may take the form of the Javanese schwa vowel or its variation, for example the /i/ phoneme [aqīqa] ‘shaving head’ becomes [kekah]. However, even for vowels that exist in Javanese, there can be a change in pronunciation, particularly when back phoneme /u/ occurs in the first syllable of the Arabic word. It is often pronounced as [o], as in kotbah for [khut.ba] ‘sermon’, sokur for [shukr] ‘thanks’, donga for [du’ä] ‘prayer’. The general patterns of vowel change are explained further in below.

Sound changes involving vowel recast generally take the following patterns (originals are square-bracketed):

- a) Long vowel → short vowel, as in the cases of [ghālib] → galib (Table 4.1.2); [mashghül] → mesgul and [shahīd] → sahid (Table 4.1.5).
- b) Changes involving the close back vowel [u] is typically recast as the back close-mid [o], as in [muh.äl] → mokal; [khut.ba] → kotbah. However, there is also the untypical change of this close high vowel into open-backed [a], particularly when the it’s followed by velar sound, as in [nuskha] → naskah.
- c) When fricatives occur with the open vowel [a] in Arabic, there is a tendency to pronounce this vowel as a schwa in Javanese, as in Table 4.1.5: [mashghül] → mesgul; [thalātha] → Selasa.
- d) When a consonant cluster occurs in the Arabic word, there is a tendency to insert a vowel in between, as in [‘aq] → akal (Table 4.1.1); [s.br] → sabar (Table 4.1.5) and [dhikr] → dikir (Table 4.1.4). The type of vowel being inserted depends on the vowel preceding or following the cluster.

Another vowel recast of different nature is concerned with the Arabic velarised [a] (see Table 3, referred to open-central). In this case, there are two options: (1) changing it to open front vowel [a] in Javanese, or, most commonly it’s recast as the Javanese nasal [N] as in [ah.ad] → ngahad (see table 4.1.3).

5. Semantic Recast in Direct Borrowing

It is interesting to note that loanwords in the category of direct borrowing are those mainly concerned with philosophical concepts and religious practices. In the examples of Table 5, the original Arabic meaning and the Javanese meaning are presented, mainly to see where semantic recast has occurred together with the phonetic recast.

The examples in the previous Tables also indicate that ‘direct borrowing’ very rarely changes meaning or very slightly if any (eg from general meaning to specific). In fact, this tendency of meaning-specification seems to be a general tendency in many of the

loanwords of the ‘direct borrowing’ category. Another change of meaning that also generally occur would be those along the singular-plural dimension of nouns, for example *kewan* ‘animal’ comes from [khayawān] that is plural in Arabic. The change along this dimension is very common in a language such as Javanese (and Indonesian and other ethnic languages of Indonesia such as Madurese and Sundanese), where plurality is not syntactic but lexical, through word reduplication.

Last but not least, sound changes involving the [-at] or [-ah] ending that relates to [taa] *marbouta* and gender marker in Arabic are discussed in Section 6 below.

Table 5: *Examples of semantic recast*

Loanword	Arabic meaning	Javanese meaning
Hikmah Arabic [h.ikma]	‘underlying wisdom/power’	supernatural power; supernatural wisdom
Ilham Arabic [ilhām]	‘inspiration’	heavenly-sent inspiration
Isim Arabic [ʾs.m]	‘safeguard; protect’	A written Arabic phrase used for warding off danger/illness
Jabur Arabic [jbr]	‘gratify’	refreshment served at a collective prayer gathering during the evening of the fasting month
Kajat Arabic [khāja]	‘necessity; desire; wish’	to hold a party; ceremonial event
kalimasada Arabic [kalimatussyahāda]	‘the Muslim Credo’	magical book having the power of resurrection
Sekaten Arabic [syahādaitain]	‘creed’	(1) an important court festival held during Mulud; (2) the gong music accompanying the festival.

A brief examination of the examples above will immediately show us that in general the Javanese meaning is very specialised compared to its original. Of the seven examples above, the last two deserve further explanation (please see Section 6 below). While they both use the same etymon as the base word (ie [syahāda] ‘the Muslim credo’), they are obviously not cognates. This phenomenon is explained further in Section 7 on etymological doublets.

6. Some morphological recast

It has been mentioned above that plurality is grammatical in Arabic, while in Javanese it is lexical through repetition of the noun (see also the note on ‘*kewan*’ above, where the plural form is used as singular in Javanese). There is also a case of gender marker [-at] or [-ah] ending in Arabic. According to Campbell (1996) the distribution of the feminine endings -at and -ah in Indonesian and Malay reflects the distribution of these endings in Persian and Hindustani (Urdu), at least for the older period. In his view -at is the earlier form, which correlates with Persian and Urdu/Hindustani. The alternative ending -ah represents a later re-arabization. However, while this feature is commonly found in Indonesian/Malay, it is not so in Javanese, at least not in the JDI entries.

Ideally, the discussion on morphological recast in this Section should cover grammatical and lexical devices such as above and how they can be used productively in loanwords. However, in this paper the focus on morphological recast is only on two

categories: the so-called ‘loan blends’ (Section 6.1) and ‘loan translations’ (Section 6.2). Then the processes involved in these two categories will be discussed.

Apart from these two categories, however, there is also the recast through word ellipsis, while retaining the meaning. For example, *rijal* ‘noises heard in the night (invisible)’ is an elided form of [*rijalul’gaib*]; the same goes for [*sunnat-ar-rasul*], a euphemism for ‘circumcision’, has been borrowed in the Javanese inventory as *sunat*.

6.1 *Loan-blends*

Loan-blends occur when new idioms are partly formed from Arabic and partly from Javanese elements. Differing from what Poedjosoedarmo has suggested for Javanese loanwords in Indonesian, loan-blends can also be formed from Arabic words, either for creating new meaning or for providing emphasis. It should be noted that when this occurs, either of the Arabic words would be attached with Javanese meaning. Another kind of loan-blend is when an Arabic word is morphologically changed using Javanese affixes to form new meaning. Therefore, the word ‘idiom’ can take the form of a word or more (Fernando, 1996:2)

The main purpose of a loan-blend is for meaning specialisation: it is for making meaning more specific for particular contexts. For example, *abdi* ‘servant’ can be further specialised as *abdi dalem* ‘one who serves in the royal palace/court’. Table 6 provides some more examples of loan-blends.

Table 6.1: *Loan-blends*

Notes: A= Arabic origin; J= Javanese origin

loan-blends	meaning	Word compounds
<i>asal mula</i>	‘place of origin; source’	A: <i>asal</i> (from Arabic [<i>as.li</i>]) ‘original’ J: <i>mula</i> ‘beginning’
<i>ilmu kedjawen</i>	‘mystical Javanism’	A: <i>ilmu</i> (from Arabic [<i>’ilm</i>] ‘science; knowledge’ J: <i>kedjawen</i> ‘Javanese-ism’ [<i>ke-an</i>] affixes in this case means ‘concerned with’.
<i>ratu adil</i>	‘the messiah’	J: <i>ratu</i> ‘queen’ A: <i>adil</i> (from Arabic [<i>’ādil</i>] ‘just’
<i>kejiman</i>	‘being (inadvertently) possessed (by the genie)’	A: <i>jim</i> (from Arabic [<i>jinn</i>] ‘genie’ J: [<i>ke-an</i>] affixes in Javanese which means ‘[done] inadvertently’ (Horne, p.xxii)

It is important to note that the [*ke-*] and [*-an*] affixes are just two of the very frequently used grammatical devices for use in a productive way in loanwords of this kind. In some cases, the device can be used for creating a ‘new’ loanword formed from the same etymon, thus producing etymological doublets which may not be cognates. An obvious example of this concerns the word *slamet* [*salāma*] ‘well; safe and sound’. Once the Javanese verb-forming device [*-an*] is applied to the loanword (in this case in the word *slametan*), the meaning can change dramatically, one of which would be *slametan* ‘to hold ceremony on someone’s behalf’ (see Section 7 on etymological doublets).

Therefore, the elements in loan blends can be of different nature, lexical and/or morphological, as can be seen in table 6.1 above.

6.2 Loan translation

Loan translations occur when new words are constructed from Javanese forms parallel to Arabic models. Loanwords in this category seem to be generally related to spirituality and mysticism. In the examples that follow, they are presented as etymological doublets, even triplets, as they appear as entries in JDI, along with their meaning.

Table 6.2: *Examples of loan translations*

	Loanword & and origin	loan translation	meaning
1	berkah Arabic [baraka] 'blessing'	berkat	food, blessed by a religious official, taken home from a ritual ceremony by the guests after they have eaten a portion of it.
2	ilmu Arabic ['ilm] 'knowledge; science'	elmu	abstract knowledge; supernormal skill; mystic philosophical lore
3	sariah Arabic [sharī'a] 'cannon law; lawfulness'	serengat	fulfilment of religious duty according to Muslim law
4	tarekah Arabic[t.arīqa] 'Derwish order; spiritual path, esp. mystical'	tarekat	religious acts that bring one closer to God, according to Islamic principle (eg meditation, devotion to the ethical principles)
		tirakat	to deny oneself food & sleep as a self sacrificial act in order to be granted one's desire

Campbell (1996) looked at cases like this in Malay/Indonesian and recognized that the loanwords in Arabic should be classified into several layers, derived from different sources and stemming from different periods.

The most recent layer of Arabic loanwords in Malay/Indonesian is relatively easy to recognize because it contains many words that were borrowed in an exclusively written way (for example direct borrowing). According to Campbell (1996), in this most recent stage some words that had already been borrowed in an earlier period were rearabized, which explains the presence of some lexical pairs in the language, such as *sahabat/sobat* 'friend' (both are from Arabic [sahāba] 'friendship' or the latter from [suhba] 'friendship'), *fardu* 'moral obligation'/*perlu* 'must, need' (both are from Arabic [fard] 'moral obligation'), *mungkin/mungkin* 'possible' (both are from Arabic [mumkin]).

While this explanation on chronological layers maybe true for Indonesian and Malay, especially for explaining cases of cognates, it seems to be different for Javanese, since the 'new' word gains new meaning, not cognate. It is, therefore, necessary here to provide an explanation for this phenomenon, and to use a different conceptual framework. In this context, the case of 'loan translation' above is explainable in terms of what Gonda (1991) calls 'psychological rejection' of a loanword by creating a 'new' one from the already existing loan and by giving it new meaning (see also Section 7 below).

7. Etymological doublets: cases of non-cognates having the same etymons

As can be seen in the examples of Table 5, the Islamic credo *syahadat* appears twice, both of which have very different meaning in Javanese. In fact, the same concept also appears as two

other entries in JDI: *sahadat* and *klimah sahadat*, with the latter being glossed as ‘phrase/expression containing the two Arabic phrases basic to Islam: believing in Allah and believing in Mohamad as prophet’. As it is, these last two would fall under the category of ‘direct borrowing’ because no change of meaning has occurred.

Thus we have:

Table 7: *Phono-morphological and semantic syncretism*

Loanword,	Origin and gloss	Javanese meaning
<i>kalimasada</i>	Arabic [<i>kalimatussyyahāda</i>] ‘the Muslim Credo’	magical book having the power of resurrection
<i>Sekaten</i>	Arabic [<i>syahādaitain</i>] ‘creed’	(1)an important court festival held during Mulud; (2) the gong music accompanying the festival.
<i>klimah sahadat</i>	Arabic [<i>kalimatussyyahāda</i>] ‘the Muslim Credo’	Same meaning (thus cognates)

The most probable explanation would be that the two loans in the Table are older than the latter two. They must have made their way into Javanese during the early years of Islamic civilization in Java. Both *kalimasada* and *sekaten* have Hindu-Bhuddist elements in them and contain an element of ‘syncretism’ central in the Javanese belief system.⁵⁾ The word *kalimasada*, for example, appears in an old writing by Mpu Sedah and Panuluh from the Old Javanese period (Zoetmoelder, 1974:261). As mentioned in the introduction, the Old Javanese period ended when Islam came, so the word must have made its way into Javanese in the transitional period.

Another important point to make concerns the doublet *slamat* and *slamet*. While the first entry is directly borrowed, it is not so with the second. It has been explained earlier that the meaning shifts dramatically in the blended word *slametan*. Besides shifts in meaning, JDI also provides types of *slametans*, among which are: *slametan brokohan* ‘christening ceremony at childbirth’, *slametan jenang abang* ‘ceremony celebrating a circumcision or wedding’, *slametan kol* ‘annual ceremony commemorating a death anniversary’, etc. These, of course, would fall under the category of loan-blends, as explained further below. It suffices here to say that such elaborate entries in the dictionary indicate the centrality and the importance of the *slametan* concept in the Javanese belief system (Geertz, 1960:11). Reasons such as this seem to be the most probable cause of loan shifts.

The examples in Table 7 show doublets or triplets that enter the dictionary: one is the loanword with original Arabic meaning and one that is a loan translation. In each case, the loan translation retains an element of meaning from the loanword, for example the semantic element ‘blessing’ that is found in the loanword is retained in the loan translation. As it was with loan translation, the meaning of the newly formed words tends to be more specialized than its loan counterpart.

⁵⁾ Syncretic Islam is referred to as the *Agami Jawi*, “a manifestation of Javanese Islam that represents an extensive complex of mystically inclined Hindu-Buddhistic beliefs and concepts, syncretically integrated in an Islamic frame of reference”. (Koentjaraningrat, 1985:317).

The ‘new’ loans presented here reflect spiritual and mystical practices in Java, which are considered inherently Javanese. Of the four groups of examples presented here, three deserve further explanation and will be used to explain how the loans interrelate. They are: tarekat, tirakat, and (ng)elmu.

According to Koentjaraningrat (p. 409ff), the people who practice Islamic mysticism in Java are usually organised into mystical movements called tarekat under the charismatic leadership of a teacher called mursid, aiming to be closer to God. In their attempts to achieve their aim (to be closer to God), they often undertake the so-called tirakat, ie denying themselves food and sleep as self-sacrificial act.

This description of mystical path is also inherent in the activity called ngelmu (the verb form of elmu ‘mystical knowledge/science’), a path for obtaining a ‘supernormal skill’; the Javanese considers elmu as highly intertwined with religion (in this case Islam). Probably that was why a different entry is necessary beside the loanword ilmu ‘knowledge’ (Koentjaraningrat, p. 410). In this case the Javanese has ‘translated’ the original Arabic loanword and redefine it in the context of its own belief system and world-view, and thereby creating a ‘new’ word, that in some way represent what Gonda refers to as ‘psychological rejection’ of the loan word.

8 Concluding remarks

Javanese words of Arabic origin have entered the inventory in different ways and can be categorized differently. In this paper, the main focus has been on the linguistic assimilation and recast of the (Arabic) words. The largest part of loan category, ie the direct borrowing, has undergone certain linguistic assimilation and recast. Patterns of sound assimilation have been identified, based on the Arabic phonemes that are non-existent in Javanese. Variations to such patterns have also been indicated.

The data from Arabic loanwords in Javanese also indicate the occurrence of etymological doublets that are not cognates. Since they are not cognates, a different explanatory framework has to be presented, one that is different from that proposed by Campbell in his postulate of chronological layers in loans in Malay/Indonesian. At first glance, doublets (or triplets) in Javanese may appear as cognates, and an untrained eye may retain one of them as the main entry, thus deleting the other (and think of it/them as cognates/s). If this happens, the heterogeneous world-view of the Javanese is reduced into the homogeneous; while the Arabic loanword in question prevails, it would fail to ‘capture’ the nuances of meaning contained in the (Javanese) facts and practices entailed in the word(s).

As far as sounds assimilation is concerned, sound patterns were presented, although it should be noted here that these patterns are rather simplified. A more detailed analysis is needed as to how these sound patterns look in different syllable patterns. This is beyond the aim of the paper.

As far as data is concerned, it would have been ideal to include materials from recent Javanese articles as they appear in Javanese magazines and newspapers, materials on the Internet, etc just so that we can see whether Arabic borrowing still continues in the more recent Javanese, considering the renewed interest on Islamic studies in Indonesia. Also it would be necessary to go beyond Central Java, to see whether Arabic borrowing shows different overall picture. As it is, the Horne dictionary, on which the data for this

paper was heavily based, is more on Javanese used in Central Java (as Horne herself mentioned).

Appendix: Arabic transliteration system used here

The transliteration is based on the Weh'r "Modern Written Arabic" dictionary, with little adaptation, due to the complex interaction with special fonts. So, instead of [d] with a dot underneath for 'dhaat', [d.] is used instead. Jones has used the same system in "kamus etimologi" for Indonesian.

No	Names of Arabic alphabet	transliterated as (sound)	example in Arabic	meaning
1	alif	[a]	[abjad]	'alphabet'
2	ba'	[b]	[bah.r]	'sea'
3	ta'	[t]	[tafsīr]	'commentary'
4	tha'	[th]	[thalj]	'snow'
5	jim	[j]	[jasad]	'body'
6	kha'	[h.]	[h.ād.ir]	'to be present'
7	kho'	[kh]	[khut.ba]	'sermon'
8	dal	[d]	[daftar]	'register'
9	dzal	[dh]	[dhikr]	'chanting'
10	raa'	[r]	[rah.ma]	'mercy'
11	zay	[z]	[zakāh]	'alms'
12	siin	[s]	[sih.r]	'sorcery'
13	shiin	[sh]	[shahīd]	'martyr'
14	shad	[s.]	[s.abr]	'patience'
15	dhaat	[d.]	[d.arūra]	'emergency'
16	taa'	[t.]	[t.abīb]	'medicine man'
17	zaa'	[z.]	[z.ālim]	'unjust'
18	ayn	[ʿ]	[ʿāshiq]	'absorbed'
19	ghayn	[gh]	[ghaib]	'invisible'
20	faa'	[f]	[fulūs]	'money'
21	qaaf	[q]	[qaum]	'people'
22	kaaf	[k]	[kātib]	'writer'
23	laam	[l]	[la'na]	'accursed'
24	miim	[m]	[mu'jiza]	'miracle'
25	nuun	[n]	[nabī]	'prophet'
26	waaw	1. [w], 2. [ū]	[wāsīt.]	'referee'
27	haa	[h]	[haiba]	'tremendous'
28	hamzah	[,]	[Al-jazā,ir]	'Algeria'
29	yaa	1. [y], 2. [ī]	[yaum]	'day'

Other transliteration: alif maqshuura [ā]; fatkha [a]; damma [u]; kasra [i]; tak marbuuta [a]; tanwiin [an, in, un]

References

- Campbell, Stuart. 1996. *The Arabic Element In Indonesian-What Do Students Need to Know About It?* <http://intranet.usc.edu.au/wacana/4/campbell.html> (The University of the Sunshine Coast, Queensland, Australia)
- Echols, John M and Hasan Shadily. 1989. *Kamus Indonesia-Inggris = An Indonesian-English dictionary* (third edition, with revision by John U. Wolff and James T. Collins) Jakarta: Gramedia.
- Fernando, Chitra. 1996. *Idioms and Idiomaticity*. London: Oxford University Press,
- Geertz, Clifford. 1964. *The religion of Java*. New York: Free Press.
- Gonda, J. 1991. 'Peminjaman di Asia Tenggara'. In Harumurti Kridalaksana (ed.) *Masa Lampau Bahasa Indonesia: sebuah bunga rampai*. Yogyakarta: Kanisius.
- Hanks, Patrick (ed.). 1986. *Collins dictionary of the English language : an extensive coverage of contemporary international and Australian English* (2nd Edition), Sydney: Collins.
- Horne, Elinor Clark. 1974. *Javanese-English dictionary*. New Haven: Yale University Press.
- Jones, Russell. 1984. Loan words in contemporary Indonesian. In J.W.M Verhaar (Ed.) *Towards a description of contemporary Indonesian: Preliminary Studies*, Part 2. NUSA Linguistic Studies in Indonesian and Languages in Indonesia, Volume 19, p. 1-38.
- KBBI, Kamus Besar Bahasa Indonesia (Indonesian Monolingual Dictionary published by the National Office of Education), 3rd Edition, Jakarta: Balai Pustaka, Kridalaksana, Harimurti (ed), 1991. *Masa Lampau Bahasa Indonesia: sebuah bunga rampai*. Yogyakarta: Kanisius.
- Koentjaraningrat. 1985. *Javanese Culture*. Singapore, New York : Oxford University Press.
- Newman, Daniel L. 2002. The phonetic status of Arabic within the world's languages: the uniqueness of the lu Al-d+Aad. In *Antwerp papers in Linguistics* (APIL), Universiteit Antwerpen, <http://www.uia.ac.be/apil/>
- Partaaatmadja. 1992. *Kamus Praktis Jawa-Indonesia: A practical Javanese-Indonesian dictionary*. Surabaya: Djojobojo.
- Poedjosoedarmo, S. 1982. *Javanese Influence on Indonesian*. Canberra: Pacific Linguistics (D-38).
- Robson, S. 2002. *Javanese Grammar for students*. Monash Asia Institute, Monash University.
- Sneddon, James N. 2003. *The Indonesian language: its history and role in modern society*. Sydney: UNSW Press.
- Suara Merdeka, an on-line news paper containing Javanese articles under the section entitled "Pamomong", <http://www.suamerdeka.com>
- Uhlenbeck, E. M. 1983. *Javanese linguistics : a retrospect and some prospects*. Koninklijk Instituut voor Taal-, Land- en Volkenkunde, Dordrecht: Foris Publications.
- Uhlenbeck, E. M. 1978. *Studies in Javanese morphology*. The Hague: Martinus Nijhoff.
- Verhaar, J.W.M (Ed.) 1984. *Towards a description of contemporary Indonesian: Preliminary Studies*, Part 2. NUSA Linguistic Studies in Indonesian and Languages in Indonesia, Volume 19.
- Wehr, Hans 1979. *A dictionary of modern written Arabic*. (J. Milton Cowan, ed.), Wiesbaden: Harrassowitz.
- Zoetmulder, P. J. 1974. *Kalangwan: a survey of old Javanese literature*. The Hague: Martinus Nijhoff.

MALAY NUMERAL CLASSIFIERS: SKETCHING CONCEPTUAL REPRESENTATION FROM A NATIVE SPEAKER'S PERSPECTIVE

Khazriyati Salehuddin
MARCS Auditory Laboratories
<k.salehuddin@uws.edu.au>

Heather Winskel
University of Western Sydney
<h.winskel@uws.edu.au>

Abstract

Numeral classifiers are defined as a syntactic-semantic category denoting “some salient perceived or imputed characteristic” of a particular count noun (Allan 1977: 298). Similar to 36 other East and Southeast Asian languages, Malay has a rich numeral classifier system (Adams & Conklin 1973). Numeral classifiers in Malay syntactically exist in a noun phrase together with a numeral preceding it and an obligatory noun following it. Malay numeral classifiers are semantically significant as they provide information about the physical and functional properties of count nouns to the speakers within the Malay speech community. The semantic criteria for dividing the system of each classifier category appears to be complex and opaque, with mixed semantic criteria involved in classifying members of a given category. Though the motivation may be obvious in cases like “dua **orang** kanak-kanak”, “dua **ekor** kucing” and “dua **batang** pokok”, (literally in English ‘two **human** children’, ‘two **tail** cats’, and ‘two **stem** trees’ respectively), the motivation seems somewhat arbitrary in cases like “dua **buah** kereta”, “dua **patah** perkataan” and “dua **bentuk** cincin” (literally ‘two **fruit** cars’, ‘two **broken** words’ and ‘two **shape** rings’ respectively). Other than the classifiers for human, animals and trees – which are rather straight forward in their usage – the acquisition of the other Malay classifiers seems to involve a lot of memorisation in determining which classifier should co-occur with a particular noun; as there seem to be many exceptions or non-prototypical examples which need to be learnt, for example, “biji” (the classifier that literally means ‘seed’ for all types of fruits, eggs, some household items) and “kaki” (the classifier that literally means ‘leg’ for umbrellas and some kinds of plants) (Asmah Haji Omar & Rama Subbiah, 1995: 23). Despite sharing some universal characteristics with other classifier languages such as Thai, Japanese and Korean (Carpenter, 1991; Uchida & Imai, 1999; and Yamamoto & Keil, 2000), the subcategorisation of the Malay classifiers has characteristics that are specific to Malay. This paper aims to illustrate the conceptual system of Malay classifiers from a native speaker’s perspective and discuss its classification vis-à-vis Allan’s (1977) description of other classifier languages of the world.

Numeral Classifiers

Numeral classifiers form one syntactic-semantic category that is common in Sino-Tibetan and Austronesian languages – so common that 37 of the Southeast Asian and East Asian

languages have them as an obligatory category in their linguistic structure. They are also described as a closed class lexical category as their membership is so exclusive that new members are not added to the existing list of numeral classifiers (Uchida & Imai 1999: 51). Syntactically, numeral classifiers form part of a noun phrase within which, this syntactic-semantic category must be present together with a numeral and a head noun. Numeral classifier examples in Thai, Japanese and Chinese are illustrated as follows.

Thai	N	Adj	Num	CL	
	Dek	naa-rak	soong	khon	
	Child	lovable	two	CL	
		'Two cute children' (Simpsons, n.d.)			
Japanese	Num	CL	N		
	Ni	mai	kami		
	Two	CL	paper		
	'two sheets of paper' (Yamamoto & Keil 2000)				
Chinese	Num	CL	N		
	Saam1	go3	gaalze1		
	Three	CL	older sister		
	'three older sisters' (Wei & Lee 2001)				

Semantically, numeral classifiers provide information about the physical and functional properties, and to a certain extent, the social status of a noun, to the speakers within and across a particular speech community. Because of this, classifiers are actually a grammatical system that reflects the manner countable objects are counted, quantified and categorized by the particular speech community (Yamamoto & Keil 2000).

Though structurally their grammatical role seems somewhat similar to quantifiers in English such as 'a piece of' and 'a portion of', classifiers are different from the English quantifiers due to the fact that rather than "unitizing" them, classifiers categorize nouns into categories that carry semantic information. In other words, instead of quantifying mass nouns in measuring units, classifiers qualify all count nouns based on their permanent qualities. In addition, classifiers are stable because unlike English quantifiers, substitutions of classifiers results in the ungrammatical use of language.

Malay, an Austronesian language, is said to be one of the languages that has an extensive system of numeral classifiers (Richards et al. 1985). The system of Malay numeral classifiers is very complex and has to be formally taught in schools. This is indicated by numerous publications of reference books solely focusing on Malay classifiers. The Malay numeral classifiers are also described as coefficients and numeral coefficients (Asmah Haji Omar & Rama Subbiah 1995: 23) and its absence makes the structure of any formal Malay sentence ungrammatical. Structurally, Malay numeral classifiers are somewhat similar to Japanese and Chinese, as Malay numeral classifiers are preceded by a numeral and followed by a head noun.

Num CL N
Tiga orang kanak-kanak
 Three CL child
 ‘three children’

Num CL N
Empat ekor arnab
 Four CL rabbit -pl
 ‘four rabbits’

As illustrated above, when numeral classifiers are used in Malay, the noun that forms the head of the noun phrase appears in a singular form. In situations where numbers are not used to indicate plurality, numeral classifiers are obligatorily dropped (Asmah Haji Omar & Rama Subbiah 1995).

N -pl	N -pl
<i>Rumah-rumah</i>	<i>Buku-buku</i>
House -pl	Book -pl
‘houses’	‘books’

In isolated cases, instead of appearing to the left of the head noun, classifiers can be used to the right of it, following a numeral.

N Num CL
anak dua orang (DBP Corpus: B00403)
 Child two CL
 ‘Two children’

N Num CL
Kerbau balar tujuh ekor (DBP Corpus: B00455)
 Buffalo albino seven CL
 ‘Seven albino buffalos’

Despite being syntactically different from the norm, what is most significant in the two examples above is that the choice of classifiers is maintained and consistent for each type of noun, i.e. the noun ‘children’ co-occurs with the {animate; human} “orang” classifier while the noun ‘buffalos’ co-occurs with the {animate; animal} “ekor” classifier.

Imai and Gentner described numeral classifiers as similar to modifiers that unitize uncountable objects for quantifying purposes (i.e. unitizing modifiers) such as ‘a spoonful of sugar’, ‘a bowl of rice’ and ‘a cup of coffee’ (1997: 174). This however does not truly describe Malay classifiers as while unitizing modifiers are used to quantify mass nouns, numeral classifiers are used to quantify all count nouns. Imai and Gentner’s unitizing modifiers are indeed used in Malay expression such as “sesudu gula”, “semangkuk nasi” and “secawan kopi”, each means ‘a spoon of sugar’, ‘a bowl of rice’ and ‘a cup of coffee’ respectively. Unitizing modifiers, however, should not be equated with numeral classifiers also because as illustrated above, unlike unitizing modifiers, numeral classifiers are stable. To illustrate, the numeral classifier “orang”, may only be used for human beings and

substitutions with any other classifier will result in ungrammatical use of language. On the other hand, unitizing modifiers are unstable in nature. To illustrate, ‘sugar’ can collocate with a variety of expressions like ‘a pinch of’, ‘a spoon of’, ‘a cup of’, ‘a bowl of’, ‘a sack of’ etc.

At a superficial level, the selection of numeral classifiers in classifying Malay nouns seems to be arbitrary. Although its usage in “dua orang kanak-kanak”, “dua ekor kucing” and “dua batang pokok” for ‘two CL children’, ‘two CL cats’ and ‘two CL trees’ seems rather straightforward, its usage in “dua buah kereta”, “dua biji cawan” and “dua kaki payung” for ‘two CL cars’, ‘two CL cups’ and ‘two CL umbrella’ is rather difficult to explain.

Apart from the numeral classifiers for human (“orang” – literally ‘human’), animal (“ekor” – literally ‘tail’) and trees (“batang” – literally ‘stem’) that seem to be relatively semantically transparent, the acquisition of other Malay classifiers appears to involve a more cognitively demanding task. There appear to be many exceptions to the rule in the selection of some classifiers, e.g. the classifiers “buah” that literally means ‘fruit’ for cars, houses, concepts but not fruit, the classifier “biji” that literally means ‘seed’ for cups, stones but not seeds, and the classifier “kaki” that literally means ‘leg’ for umbrellas, and some plants, but not legs. This kind of arbitrariness in the majority of Malay numeral classifiers means that Malay classifiers are often regarded as loose and opaque which presumably results in greater difficulty on the part of the learner in acquiring them.

Numeral Classifiers and their Cognition

The study on numeral classifiers is significant in psycholinguistics as its usage reflects the manner a particular speech community categorizes countable objects (Yamamoto & Keil 2000).

Numeral classifiers are said to be a unit that classifies nouns based on the characteristics of the nouns in question. Generally, numeral classifier system shares some universal features among different classifier languages such as the distinction between animate and inanimate objects.

However, the classifier system also exhibits the manner an individual language categorizes objects which is specific to that particular language. The way classifiers divide the semantic space gives psycholinguistic researchers the opportunity to carry out research on the relationship between language and cognition, such as the Whorfian hypothesis and whether linguistic categories affect the way humans think or perceive of the world, or in turn to what extent does the way that humans think or perceive influence the linguistic categories formed in languages of the world. A less deterministic version of the Whorfian hypothesis argues that language may affect classification during the process of thinking-for-speaking (Slobin 1996, 2003). Children focus their attention on the language-specific characteristics of their language from a very early age. These aspects may affect online thinking-for-speaking or may have longer term consequences on non-linguistic cognition.

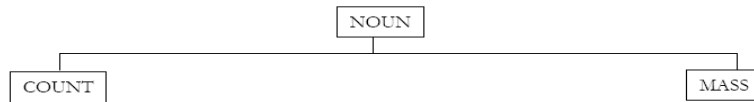
Though classifiers are an important syntactic-semantic element of Malay linguistics, the process involved in the categorization of Malay classifiers that takes place in the cognitive system of its native speakers has never been studied.

The collocation between Malay numeral classifiers with particular nouns appears to occur in both a systematic and fairly arbitrarily manner. Furthermore, there are cases where

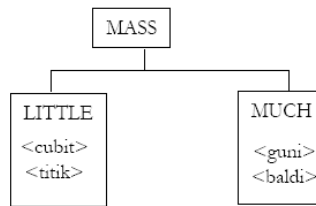
mixed semantic criteria are used in object categorization. These mixed semantic criteria involve biological features, size, and form of objects.

This paper aims to illustrate the conceptual system of Malay numeral classifiers from a native speaker’s perspective and discuss its classification vis-à-vis Allan’s (1977) description of other classifier languages of the world.

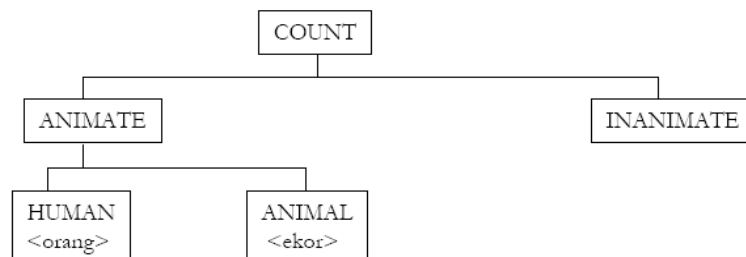
Similar to other languages of the world, Malay nouns are subcategorized into “count nouns” and “mass nouns”.



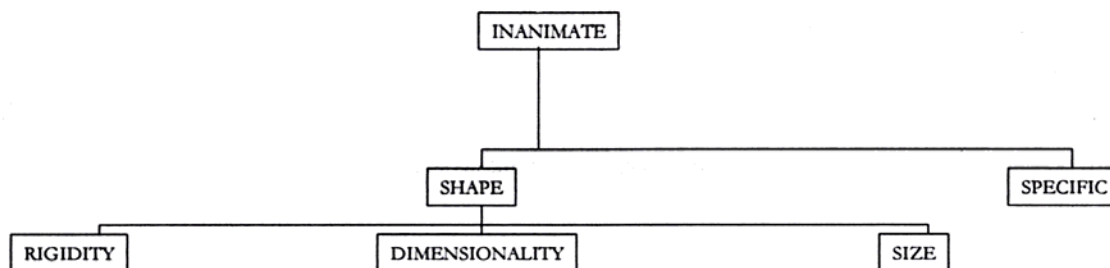
The Malay mass nouns do not co-occur with any classifier; instead, they co-occur with unitizing modifiers based on the quantity of the mass nouns that is needed to modify the nouns. To illustrate, mass nouns occurring in a small quantity can be modified by quantifiers “cubit” (a pinch of) and “titik” (a drop of), while those occurring in a large quantity may be modified by other quantifiers such as “guni” (a sack of) and “baldi” (a pail of).



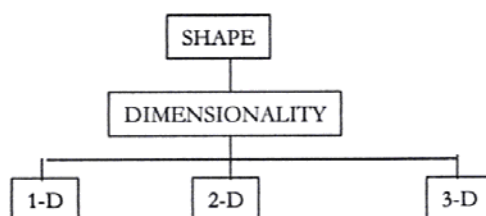
Malay count nouns can further be categorized by the Malay numeral classifier system into two semantic categories on the basis of animacy, i.e. animate versus inanimate. The animate nouns are then subcategorized into two subcategories namely human and animal which are qualified by the classifiers “orang” and “ekor” respectively.



Inanimate objects on the other hand are subsequently subcategorized into two subcategories i.e. shape and the specific categories. The shape category is then further subcategorized into rigidity, dimensionality and size.



Further analysis on the organization of Malay classifiers shows that to a large extent Malay classifiers can be described in a similar manner as the shape-based classifiers described by Allan (1977). This is because, like most classifier languages that Allan described, shape-based classifiers of Malay can also be subcategorized based on the dimensionality of the objects, i.e. one-dimensional, two-dimensional and three-dimensional classifiers.



The division of the shape-based classifier category however, is not clear-cut as its partitioning is not mono-referential but poly-referential in nature. This subcategory of classifier involves a combination of different subcategories, i.e. with mixed semantic criteria used to qualify a particular noun due to the dependency of secondary parameters on primary parameters.

Each of the one- and two-dimensional categories, for example, is not independent as each of them has to be combined with a different subcategory, i.e. 'rigidity' - depending on whether the object is rigid or flexible.

Thus, one-dimensional nouns that are rigid like a pen, a ruler, a tree and a cane, co-occur with the classifier "batang" while one-dimensional nouns that are flexible like a chain, a strand of thread, and a bracelet co-occur with the classifier "utas". Two-dimensional nouns, on the other hand, if they are rigid like a plank of wood, a slice of cake, and a photograph, co-occur with the classifier "keping" while those that are flexible like a piece of cloth, a shirt, a piece of paper and a pair of trousers, co-occur with the classifier "helai".

Similar to the description of the shape-based classifiers for the majority of classifier languages, Malay's three-dimensional subcategory also does not occur independently. Its subcategorization has to be combined with the 'size' subcategory. However, unlike the other classifier languages described by Allan (1977), the size range for size classification is from 'fine', 'small', 'medium' and 'big'.

Thus, for three-dimensional objects that are fine like pebbles, precious stones, rice and sand, the classifier "butir" is used while those that are small e.g. fruits, cups, and bulbs,

the classifier “biji” is used instead. Three-dimensional objects that are medium in size like chunks of chocolate, fist-sized stones and lumps of clay co-occur with “ketul”, while those that are big e.g. bags, computers, cars and houses co-occur with the classifier “buah”. The shape-based classifiers mentioned above are summarized below.

Table 1: Malay shape-based numeral classifiers

Dimensionality Other References	One-Dimensional	Two-Dimensional	Three-Dimensional
+rigid	<u>Batang</u> [pencil, road, ruler, needle, rod]	<u>Keping</u> [plank, sliced bread / cake, cookies]	(not applicable)
-rigid	<u>Utas</u> [chain, bracelet, thread, hair]	<u>Helai</u> [leaf, paper, shirt, pants, cloth]	(not applicable)
Big	(not applicable)	(not applicable)	<u>Buah</u> [book, computer, cupboard, vehicle, house]
Medium	(not applicable)	(not applicable)	<u>Ketul</u> [chunks of chocolate, fist-sized soil]
Small	(not applicable)	(not applicable)	<u>Biji</u> [stone, fruit, cup, candy, egg]
Fine	(not applicable)	(not applicable)	<u>Butir</u> [rice, sand, precious stones, pebbles]

Similar to Micronesian languages, Malay does not have function based classifiers (Adams & Conklin 1973). Instead, specific classifiers exist in Malay language. The specific sub-category however is even more complex than the other classifier sub-categories since unlike those of shape, rigidity and size, the mapping of specific classifiers is not based on perceptual saliency. Specific categories are unique in their use as their usage is only limited to the nouns they collocate with. Flowers, for example, co-occur only with “kuntum” while firearms only with “laras”. The following illustrates the possible Malay noun phrases with their respective classifier.

Num	CL	N
Tiga	orang	guru
Three	CL {animate; human}	teacher
	‘three teachers’	
Num	CL	N
Tiga	ekor	kucing
Three	CL {animate; animal}	cat
	‘three cats’	
Num	CL	N
Tiga	batang	pensel
Three	CL {1D,+rigid}	pencil
	‘three pencils’	

Num	CL	N
Tiga	utas	rantai
Three	CL {1D, -rigid}	chain
	'three chains'	
Num	CL	N
Tiga	keping	papan
Three	CL {2D, +rigid}	plank
	'three planks'	
Num	CL	N
Tiga	helai	sarung
Three	CL {2D, -rigid}	sarung
	'three sarungs'	
Num	CL	N
Tiga	butir	pasir
Three	CL {3D, fine}	sand
	'three grains of sand'	
Num	CL	N
Tiga	biji	cawan
Three	CL {3D, small}	cup
	'three cups'	
Num	CL	N
Tiga	ketul	coklat
Three	CL {3D, medium}	chocolate
	'three chunks of chocolate'	
Num	CL	N
Tiga	buah	rumah
Three	CL {3D, big}	house
	'three houses'	
Num	CL	N
Tiga	buah	projek
Three	CL {specific; abstract}	project
	'three projects'	
Num	CL	N
Tiga	kuntum	bunga
Three	CL {specific; flower}	flower
	'three flowers'	
Num	CL	N
Tiga	laras	pistol
Three	CL {specific; firearms}	pistol
	'three pistols'	

Num	CL	N
Tiga	bilah	pisau
Three	CL {specific; knife}	knife
'three knives'		

Overall, the Malay Numeral Classifier system may be sketched as illustrated in Figure 1 by the majority of its adult native speakers.

Conclusion

This sketch on the conceptual representation of the numeral classifier system is only the starting point in identifying the manner Malays think and manifest what they perceive. It is hoped that this paper will stimulate further investigations into the issues raised in the language and cognition field.

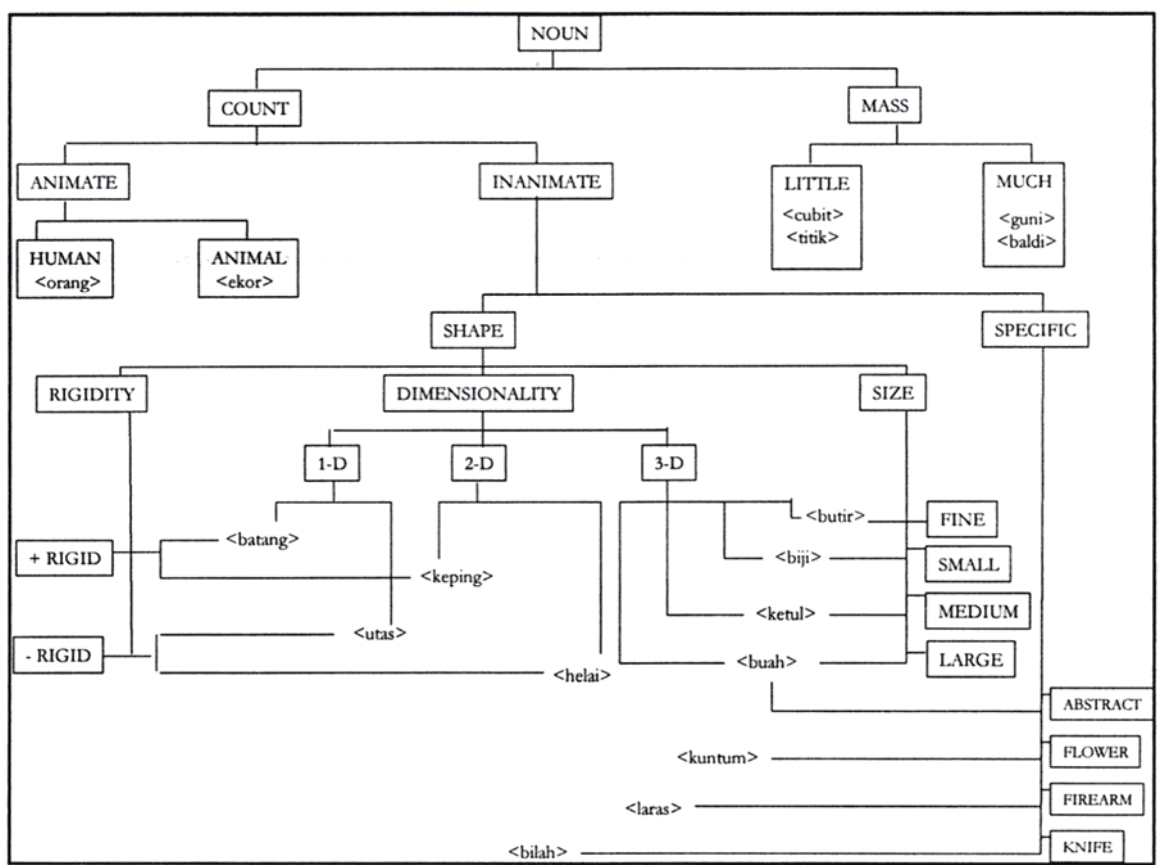


Figure 1: The Malay numeral classifier system

References

- Adams, Karen L., & Conklin, Nancy Faires. 1973. Towards a theory of natural classification. In *Papers from the Ninth Regional Meeting Chicago Linguistic Society*, Claudia Corum, T. Smith-Stark and Ann Weiser (eds.). Chicago: Linguistic Society. pp.1-10.
- Allan, Keith. 1977. Classifiers. *Language* 53(2), 285-311.
- Asmah Haji Omar & Rama Subbiah. 1995. *An Introduction to Malay Grammar*. Kuala Lumpur: Dewan Bahasa dan Pustaka
- Carpenter, Kathie. 1991. Later rather than sooner: Extralinguistic categories in the acquisition of Thai classifiers. *Journal of Child Language* 18: 93:113.
- Dewan Bahasa dan Pustaka Corpus.
- Imai, Mutsumi & Gentner, Dedre. 1997. A Cross-Linguistic Study of Early Word Meaning: Universal Ontology and Linguistic Influence. *Cognition* 62. 169-200.
- Richards, Jack C, Platt, John & Weber, Heidi. 1985. *Longman dictionary of Applied Linguistics*. Essex: Longman Group Limited.
- Simpson, Andrew. (n.d.). Classifiers and DP structure in Southeast Asia. *The Oxford handbook of Comparative Syntax*, ed. by Guglielmo Cinque & Richard Kayne, 806-838. Oxford: Oxford University Press.
- Slobin, Dan I. 1996. From “Thought and Language” to “Thinking for Speaking”. *Rethinking Linguistic Relativity*, ed. by J.J. Gumperz and S.C. Levinson, Cambridge: Cambridge University Press.
- Uchida, Nobuko & Imai, Mitsumi. 1999. *Japanese Psychological Research* 41. 1 (50-69) [Special Issues: Lexical Development: Evaluation of Word-Learning Constants]
- Wei, Li. & Lee, Sherman. 2001. The use of Cantonese classifiers and quantifiers by young British-born Chinese in Tyneside. *International Journal of Bilingual Education and Bilingualism*, 14(6), 359-382.
- Yamamoto, Kasumi & Frank Keil. 2000. The Acquisition of Japanese Numeral Classifiers: Linkage between Grammatical Forms and Conceptual Categories. *Journal of East Asian Linguistics* 9. 379-409.

FATALUKU AS A TONE LANGUAGE

Ruben Stoel
Leiden University
<r.b.stoel@let.leidenuniv.nl>

1. Introduction

Fataluku is a Papuan (or ‘non-Austronesian’) language spoken in the Lautem district of East Timor, with about 30,000 speakers. According to Hull (2005) there are five dialects. One of these dialects is the Loré dialect, which was described by Campagnolo in his doctoral thesis (Campagnolo 1973). Probably the most remarkable feature of this dialect is its complex prosodic system. In the present paper I will first give an outline of Campagnolo’s analysis of some parts of this system, and then I will present an alternative analysis of my own. Whereas Campagnolo uses the notion of accent, I will analyze Fataluku as a tone language, in which each syllable either has a H (high) tone or no tone.

2. Campagnolo’s analysis

2.1 Accent classes

According to Campagnolo (1973), each Fataluku word belongs to one of seven accent classes, depending on the number of syllables (native words have from 1 to 4 syllables). In case of monosyllabic words, there is only one accent class, while there are two classes for words of two, three, and four syllables each. In each class, the accent may fall on specific syllables only. For example, words of class 4 can have an accent on the second, third, or fourth syllable, or on both the second and fourth syllables, while words of class 4’ can have an accent on the first, third, or fourth syllable, or on both the first and fourth syllables. The actual accentuation pattern depends on the context in which the word is used.

This is illustrated by the example in (1), which shows the four possible accentuation patterns of the class 4 word *eceremu* ‘think’.

- | | | | |
|-----|----------------------|--------------------|---------------------|
| (1) | <i>ecéremu</i> | ‘(he) thinks’ | (2nd. syll.) |
| | <i>háí ecerému</i> | ‘(he) thought’ | (3th. syll.) |
| | <i>háí eceremú i</i> | ‘(he) thought so!’ | (4th. syll.) |
| | <i>eceremú nara</i> | ‘if (he) thinks’ | (2nd. + 4th. syll.) |

The distinction between classes 4 and 4’ depends on whether the first syllable can get an accent or not, and the same is true for the two classes of three-syllabic words. If used in isolation, words of classes 3’ and 4’ have an accent on the first syllable, while words of classes 3 and 4 are accented on the second syllable.

The two-syllabic word classes, however, are distinguished by vowel length, rather than accent position. In isolation, words of class 2’ have a ‘long accent’ on the first syllable, while words of class 2 have a ‘short accent’ on the first syllable (this is actually a distinction in vowel length, rather than accent). This is illustrated by the minimal pair in example (2).

- (2) lóre ‘sow’ (class 2)
 ló:re ‘leave’ (class 2’)

Note that the opposition between short and long vowels only occurs in the first syllable of bisyllabic words, and only if this syllable is accented. Monosyllabic (class 1) words always have a long vowel, if used in isolation.

- (3) le ‘house’ [le:]

Classes 1, 2’, 3’, and 4’ are marked, while classes 2, 3, and 4 are unmarked (Campagnolo 1973:72). The number of lexemes that belong to the unmarked classes is greater than those that belong to the marked classes. Marked and unmarked words have different reduplication patterns. In unmarked words, the first two syllables are reduplicated, while in marked words only the first syllable is reduplicated, as in example (4). Note that markedness is indicated by an accent sign on the first syllable of a marked word.

- (4) tapule ‘buy’ tapu-tapule ‘buy repeatedly’
 fúlehe ‘return’ fu-fulehe ‘come and go back’

2.2 Accents in two-word phrases

If a word is followed by another word in the same phrase, then either word may lose its accent, or the accent of the second word may shift to the penultimate syllable. It is also possible that the first word obtains an additional accent on its final syllable. According to Campagnolo (1973:75-77), there are four possible accentuation patterns for two consecutive words, for which he uses the symbols ‘+’, ‘x’, ‘□’, and ‘<’. The meaning of these symbols is summarized in table 1 (I use the notion ‘default accent’ to refer to an accent on the first syllable for marked words, or on the second syllable for unmarked words).

Campagnolo (1973: 77) distinguishes between ‘junction’ and ‘disjunction’, which determine which of the four possible accentuation patterns are actually used in a given situation. As table 2 shows, this depends on the classes of the two words in the phrase.

Table 1: *Accentuation in two-word phrases*

type	word 1	word 2
+	default accent	penultimate accent or no accent
x	default accent	default accent
□	no accent	default accent
<	default accent and final accent	penultimate accent or no accent

Table 2: *Accentuation patterns in junction and disjunction.*

word 1	word 2	junction	disjunction
class 1, 2	any class	+	□
class 2', 3, 3', 4, 4'	class 1	+	<
	class 2, 2', 3, 3', 4, 4'	x	<

From the examples given by Campagnolo on p. 78-79, it appears that the distinction junction - disjunction corresponds to a distinction in usage as follows:

junction:

- 3rd person possessive pronoun *i* + noun
- compounds
- object + verb (in declarative sentences)

disjunction:

- 2nd person plural possessive pronoun *i* + noun
- possessive constructions (with non-pronoun possessor)
- subject + verb
- object + verb (in imperative sentences)

For example, the phrase *tava átane* 'he asks' consists of the class 2 word *tava* (3.sg) followed by the class 3' word *átane* 'ask'. Since this is a subject-verb construction, we have a disjunction, and hence type '□'. Thus the first word loses its accent, while the second word keeps its default accent. This results in a single accent on the first syllable of the second word, i.e. *tava Atane*.

3. A tonal analysis

3.1 Accent or tone?

My reanalysis is motivated by the idea that the analysis of Campagnolo is too complex. The distinction between four accentuation types in combination with a distinction between junction and disjunction is too difficult to use in practice. Moreover, the junction - disjunction distinction does not make much sense for possessive pronouns, since a 3rd person possessor results in a junction, and a 2nd person plural possessor in a disjunction. Clearly this has no semantic or syntactic motivation.

Campagnolo analyses accent as a property of words. However, since at least Nespor and Vogel (1986), phonologists have recognized that there are prosodic constituents which are partly independent of the syntax. Some of the accents that are recognized by Campagnolo can better be analyzed as not belonging to the word, but to a larger prosodic constituent. This includes word-final accents, such as the second accent in the phrase *ecéremú nara* 'if (he) thinks'. I propose that these two accents must be analyzed as a lexical tone and a boundary tone, respectively, in which the lexical tone is associated with a word, and the boundary tone with a phrase.

Campagnolo analyses Fataluku as a language in which accents can move to different syllables of the word (cf. example (1) above). However, ‘typical’ accentual languages, such as Japanese, do not allow the pitch accent to move around within a word. So it is doubtful whether Fataluku should be called an accentual language. Moreover, Pulleyblank (1986) argues that the so-called ‘accentual languages’ are better analyzed as tone languages, in which each syllable either has a H (high) tone or no tone (see also Yip 2002:258). I propose a similar analysis for Fataluku here. From a prosodic point of view, Fataluku is then rather different from most other Southeast Asian languages, and more resembles some of the Bantu languages spoken in Southeast Africa (cf. Odden 1988).

3.2 Lexical tones

I will make a distinction between lexical tones and boundary tones. A lexical tone is a tone that is associated with a word, and is given in the lexicon. A boundary tone is associated with the end of a prosodic domain, and is inserted by the phonology (see Ladd 1996 for a discussion of boundary tones within the model of intonational phonology). Instead of the seven word classes distinguished by Campagnolo, I make a distinction between only three classes, which differ in their associated lexical tones as follows:

- class 1 words have a lexical H tone associated with the first syllable
- class 2 words have a lexical H tone associated with the second syllable
- class 0 words do not have an associated lexical H tone

All full words (i.e. nouns and verbs) are either class 1 or class 2 words, whereas grammatical words can be of any class. For example, the third person possessive pronoun *í* is a class 1 word, while the second person plural possessive pronoun *i* is a class 0 word.

Under certain conditions, a lexical tone may be deleted or move to another syllable. If a word is spoken in isolation, that is, if it is used as a one-word sentence, then the word keeps its lexical tone, except if it is associated with the final syllable of the word. I assume that Fataluku has a constraint that forbids a lexical H tone to be associated with the final syllable of an utterance. An explanation for this constraint will be presented below. Any lexical H tone that is associated with the final syllable moves to the preceding syllable, if there is such as syllable.

Thus, in a word of three syllables or more, the lexical H tone is maintained (in the following examples, H tones are written after the respective syllables).

- (5) *u.pu^H.ru* ‘fly’ → *u.pu^H.ru* (2nd. syll.)
 lá^H.pu.sa ‘spider’ → *la^H.pu.sa* (1st. syll.)

In words of two syllables with a lexical H tone associated with the second syllable, the H tone moves to the first syllable due to the ‘no final lexical H tone’ constraint. The lexical H tone does not move if it is associated with the first syllable, but the first syllable is lengthened instead. I assume that vowel length is predictable, and that there is no phonemic distinction between short and long vowels.

- (6) *po.si^H* ‘cat’ → *po^H.si* (tone moves)
 pá^H.lu ‘father’ → *pa^H.lu* (lengthening)

In monosyllabic words, the lexical H tone cannot move to another syllable, so the ‘no final lexical H tone’ constraint is violated, as there must be at least one lexical H tone in an utterance. Again, the syllable is lengthened.

(7) le^H ‘house’ → le:^H (lengthening)

This analysis leads to three questions:

- what motivates the ‘no final lexical H tone’ constraint?
- why assume that bisyllabic words without a long vowel have a lexical H tone associated with the second syllable?
- how to explain vowel lengthening?

As for the first question, a lexical H tone cannot be realized on the final syllable of an utterance, because it appears that this syllable is reserved for intonational tones. This includes a H tone that can turn a statement into a question, as in example (8).

(8) po^H.si ‘a cat’ (statement)
po^H.si^H ‘a cat?’ (question)

It would not be possible to distinguish this intonational H tone from a lexical H tone if both could be associated with a final syllable, at least if a lexical H tone and an intonational H tone are identical from a phonetic point of view (this is a topic for further research). Alternatively, one could assume that there is a phonological constraint that forbids that more than one H tone is associated with the same syllable.

So the ‘no final lexical H tone’ constraint motivates the movement of the lexical H tone from the second to the first syllable in certain bisyllabic words. But why do I assume that a lexical H tone is associated with the second syllable in words without a long vowel, and with the first syllable in words with a long vowel? Firstly, in this way, all words have the same tonal structure, since words of three or four syllables also have a lexical H tone on either the first or second syllable. But the real evidence that words with a long vowel have a H tone associated with the second syllable comes from compounding.

If the second word of a compound has only one syllable, then it loses its tone, and only the tone of the first word remains.

(9) ká^H.kal ‘sago palm’ → ka^H.kal (in isolation)
ká^H.kal + to^H ‘bowl’ → ka^H.kal.to ‘k.o. bowl’

If the first word of the compound is a bisyllabic word with a final H tone, then the H tone does not move to the first syllable, as it does in isolation. Since the H tone is not final, the ‘no final H tone’ constraint is not violated, and the H tone stays on the second syllable, as predicted by our analysis.

(10) va.ta^H ‘coconut’ → va^H.ta (in isolation)
va.ta^H + to^H ‘bowl’ → va.ta^H.to ‘k.o. bowl’

Finally, the question remains how to explain vowel lengthening. According to Campagnolo, the vowel of monosyllabic (full) words is lengthened if the word is pronounced in isolation.

(11) le^H 'house' → $[le:H]$

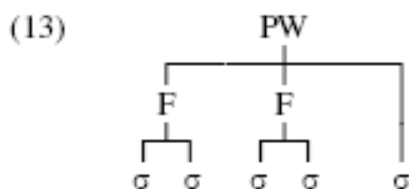
Note that the 'no final lexical H tone' constraint appears to be violated here, since the H tone cannot shift to another syllable, and, furthermore, each (full) word must have a H tone. One solution would be to assume that monosyllabic long-vowel words are actually bisyllabic, and that the H tone associates with the penultimate syllable (i.e. $[leH.e]$). But this does not explain why the vowel in bisyllabic words with an initial H tone is lengthened:

(12) $pa^H.lu$ 'father' → $[pa:H.lu]$

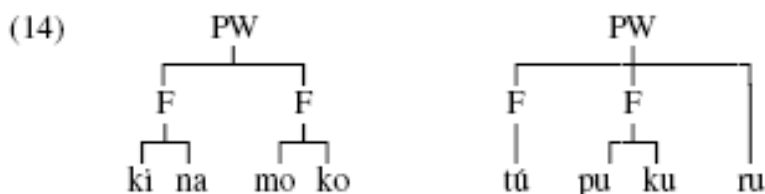
Campagnolo (1973:65) suggests that vowel lengthening in these words occurs to avoid ambiguity, but from a phonological point this cannot be considered a satisfying explanation. Another possibility is that vowel lengthening can be predicted from the prosodic structure. This is a possibility that I will develop in the next section.

3.3 The prosodic word

In this section, I present a model of the prosodic word in Fataluku. A prosodic word (PW) is a constituent within the prosodic structure, which also includes the foot (F) and the syllable (σ). A prosodic word typically corresponds to a syntactic word, although there may be exceptions. For example, a clitic may be integrated into a following or preceding prosodic word. As for Fataluku, I assume that in principle a foot consists of two syllables, but that if a syllable has an associated H tone, then it must be the last syllable of the foot. Thus there are three possible types of feet, i.e., $(\sigma.\sigma)$, $(\sigma.\sigma^H)$, and (σ^H) , while $*(\sigma^H.\sigma)$ is impossible. A prosodic word contains either one or two feet, and, optionally, an unparsed (or extrametrical) syllable. A model of the prosodic word is given in (13).



In this model, words with a H tone associated with the second syllable do not have a final extrametrical syllable, while words with an initial H tone do. Compare, for example, the words *kinamoko* 'child' and *túpukuru* 'owl' in (14).

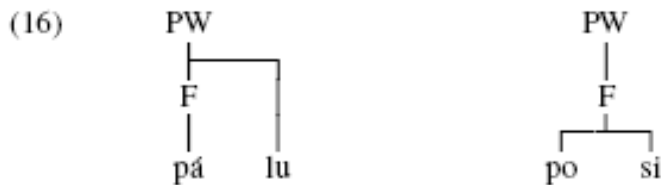


This model of prosodic structure is motivated by the following observations. First, it can explain the maximal word length in Fataluku. According to Campagnolo (1973:84), Fataluku words have one to four syllables, but loanwords may have up to five syllables, provided that the H tone is associated with the second syllable. This is in accordance with our model, if we assume that a syntactic word must correspond to a prosodic word.

Secondly, the model explains the form of the reduplicant. The reduplicant consists of two syllables if the word has a H tone associated with the second syllable, and of one syllable if the H tone is associated with the first syllable (cf. section 2.1). By using the model, these two rules can be simplified to just one rule, i.e. the reduplicant is the foot.

- (15) $i.pi^H.le$ ‘steal’ $ipi-ipile$ ‘steal repeatedly’
 $fú^H.le.he$ ‘return’ $fu-fulehe$ ‘come and go back’

Thirdly, the model makes it possible to formulate a rule for vowel lengthening. If a word has only one parsed syllable (i.e. a syllable that is part of a foot), then the vowel of this syllable is lengthened. Thus, *pálu* ‘father’ has a long vowel, as it has only one parsed syllable, whereas there is no lengthening in *posi* ‘cat’, since both syllables are parsed, as is shown in (16).



4. Tones in two-word phrases

4.1 Lexical tones

In section 2.2, I gave an overview of Campagnolo’s analysis of accents in two-word phrases. I will now present my own analysis. I will distinguish two processes: the movement or deletion of lexical tones, which is the topic of the current section, and the insertion of boundary tones, which I will discuss in section 4.2.

In two-word phrases, the lexical H tone of the first word does not change, but the H tone of the second word may be moved or deleted, according to a rule which I will call the ‘phrase rule’. This rule only applies in one of the following conditions:

- the first word has only one syllable
- the first word has two syllables with a H tone associated with the second syllable
- the second word has only one syllable

In other cases, nothing happens. The phrase rule is as follows:

- If the lexical H tone of the second word is associated with the last foot, then this H tone is deleted, otherwise it is moved to the penultimate syllable.

Now, let’s look at a few examples. In example (17), the first word has two syllables, but the H tone is associated with the first syllable, so nothing happens.

- (17) $i^H.kar$ ‘knife’ + $la^H.fai$ ‘big’ →
 $i^H.kar.la^H.fai$ ‘machete’

In example (18), the first word also has two syllables, but now the H tone is associated with the second syllable. So the phrase rule applies, and the H tone of the second word is deleted, since it is associated with the last foot (in fact there is only one foot).

- (18) $i.a^H$ ‘foot’ + $fu.ka^H$ ‘finger’ →
 $i.a^H.fu.ka$ ‘toe’

The phrase rule also applies in example (19), since the first word has only one syllable. Note that the second word has only one foot, because the last syllable is unparsed (see section 3.3). As the H tone is associated with the last (and only) foot, it is deleted.

- (19) i^H 3.poss + $la.ri^H.nu$ ‘root’ →
 $i^H.la.ri.nu$ ‘its root’

In example (20), the second word consists of two feet. Since the H tone is not associated with the last foot, it is not deleted, but moved to the penultimate syllable.

- (20) i^H 3.poss + $sa.pu^H.ra.ki$ ‘orange’ →
 $i^H.sa.pu.ra^H.ki$ ‘his orange’

4.2 Boundary tones

A boundary tone is a H tone associated with the final syllable of the last word of a prosodic phrase. It marks the end of the phrase, and indicates that another phrase is about to follow. The H tone that may appear at the end of an utterance does not mark a boundary, but is used to form a question (see section 3.2). What Campagnolo calls a disjunction is actually a boundary between two prosodic phrases, and a junction is the absence of a boundary (cf. section 2.2). In Fataluku, an object and a following verb form a single prosodic phrase, so no boundary tone appears. A subject and a verb, on the other hand, form two prosodic phrases, separated by a boundary tone at the end of the subject. Since Fataluku is a SOV language, it is only prosody that disambiguates between a subject reading and an object reading in a sentence consisting of a noun phrase and a verb.

This is illustrated by the examples in (21) and (22). The sentence in (21) is an object-verb construction, so it consists of only one prosodic phrase, and there is no boundary tone. Note that the phrase rule (discussed in section 4.1) does not apply, because both words have more than two syllables. In sentence (22), on the other hand, there are two prosodic phrases, corresponding to the subject and the verb, and the first one ends in a boundary tone.

- (21) $tú^H.pu.ku.ru$ $e.ce^H.re.mu.$
 owl remember
 ‘(he) remembers the owl.’ (OV)

- (22) tú^H.pu.ku.ru^H e.ce.re^H.mu.
 ‘the owl remembers (him).’ (SV)

Note that the H tone of the verb in (22) is moved to the penultimate syllable. This can be explained by assuming that prosodic words may have only one H tone. Since the last syllable of the subject carries an additional H tone, it must form a prosodic word of its own. This monosyllabic word and the following verb then undergo the phrase rule. As the H tone of the verb is not associated with the last foot, it is moved to the penultimate syllable.

Another object-verb vs. subject-verb pair is given below. In (23) the phrase rule does not apply, because its conditions are not satisfied. Note that *toto* ‘see’ is a class 2 word, but its H tone shifts from the second to the first syllable of the word, because of the ‘no final lexical H tone’ constraint (cf. section 3.2). In (24), a boundary tone is inserted, so the phrase rule applies, and the H tone of *toto* is deleted, since it is associated with the last foot of the word.

- (23) tú^H.pu.ku.ru to^H.to.
 owl see
 ‘(he) sees the owl.’ (OV)

- (24) tú^H.pu.ku.ru^H to.to.
 ‘the owl sees (him).’ (SV)

A H boundary tone cannot be added to a word that already has a lexical H tone on the final syllable. In that case, the lexical H tone is deleted instead. Thus the boundary tone is actually a ‘polar tone’ that changes a toneless syllable into a syllable with a H tone, and vice versa.

An example of such a tone is given in second example below. The object-verb construction in (25) consists of a single phrase. The word *posi* ‘cat’ has a H tone associated with the second syllable, so the phrase rule applies, and the H tone of *toto* ‘see’ is deleted, since it is associated with the last foot. The subject-verb construction in (26) does not have a boundary tone, because the last syllable of the subject already carries a H tone, and tone is deleted instead. This polar tone marks a phrase boundary, so the phrase rule does not apply. As in example (23), the H tone of *toto* in (26) moves to the first syllable because of the ‘no final lexical H tone’ constraint.

- (25) po.si^H to.to.
 cat see
 ‘(he) sees the cat.’ (OV)

- (26) po.si to^H.to.
 ‘the cat sees (him).’ (SV)

5. Conclusion

Fataluku can be analyzed as a tone language, in which each syllable either has a H tone or no tone. A distinction must be made between two types of H tones: lexical tones, which are specified in the lexicon, and boundary tones, which mark the end of a prosodic phrase. Under certain conditions, a lexical tone may be deleted or moved to another syllable. A boundary

tone cannot be inserted into a word with a final lexical tone, and, instead, the lexical tone is deleted to mark the boundary. These rules correctly predict the different tone patterns of object-verb and subject verb sentences.

References

- Campagnolo, Henri. 1973. *La langue des Fataluku de Lórehe (Timor Portugais)*. Paris: Université René Descartes.
- Hull, Geoffrey. 2005. Fataluku. *East Timor language profile no. 7*. Dili: Instituto Nacional de Linguística.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- Nespor, Marina and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Odden, David. 1988. Predictable tone systems in Bantu. *Autosegmental studies on pitch accent*, ed. by H. van der Hulst and N. Smith, 225-251. Dordrecht: Foris.
- Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. Dordrecht: D. Reidel.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

THE LANGUAGE CHOICE BEHAVIORS OF PAIWAN PEOPLE IN THE TAIPEI METROPOLITAN AREA

Kai-lin Sung

National Chengchi University Graduate Institute

<Kay0711@yahoo.com.tw>

Abstract

This paper was intended to probe into the language choice behaviors of Paiwan people in the Taipei metropolitan area. In this research, 67 subjects were studied through survey questionnaire based on Fishman's domain analysis. 7 subjects were further interviewed in order to find some possible explanations for the findings. The purpose of this study, in addition to proceeding a general investigation on native language proficiency, was to explore Paiwan people's use of language from the viewpoint of conversation participants, topics, and places. Moreover, this paper was aimed at confirming the correlation of language use and two determinate variables, age and education level. With the awareness of the loss of Paiwan language, this investigation may help visualize the seriousness of the problem by presenting statistics with contrastive analysis.

0. Introduction

With abundant working opportunities, the Taipei metropolitan area has the most immigrants among all the cities in Taiwan. In Taipei, in addition to the official language, Mandarin, other Chinese dialects and aboriginal languages are used at different degrees by native speakers. However, in contrast to Chinese dialects, aboriginal languages are found rather abandoned or not passed down to the next generation. Although the influence of number of speakers on language use is self-evident, which may account for the situation of aboriginal languages in metropolitan area, still other factors are considered to play key roles as well. Several researchers have studied the situation of aboriginal languages in the aspects of language proficiency and language use. Tsao(1997) reported that aboriginal speakers generally have superior abilities in native languages to Chinese Mandarin. Han (1996) had similar findings from the case of Bunun. Liu (2002) observed that Bunun speakers' proficiency in Mandarin had been kept pace with their native language. However, little research has been done on the language proficiency and language use of Paiwan people. The purpose of this study was to investigate the language choice behaviors of Paiwan people, focusing on residents in the Taipei metropolitan area. By means of survey questionnaire and interview, this paper was aimed at presenting Paiwan aboriginals language abilities in Mandarin and their native language. Another main purpose was to probe into Paiwan aboriginals' use of language from the viewpoints of conversation participants, topics, and places. In this study, both age and education level were hypothesized to be determinant variables and were confirmed with contrastive analysis. Results of this study may provide some insights into the problem of language loss and references for further research.

1. General Information on Paiwan

Paiwan is the third largest aboriginal group spreading in Taitung and Pingtung counties, with a population of around 66,000 people. According to Li (1990), Paiwanic is one of the three subgroupings of Formosan languages in addition to Atayalic and Tsouic. Opinions on dialects of Paiwan language are found to be widely divided. Some consider the division from a geographical point of view, which separates Paiwan language into four dialects, namely Northern, Southern, Central, and Eastern Paiwan. Nevertheless, Ho (1995) argues that Paiwan language is composed of five tribal dialects, including Ravare, Paumauma, Pavuavua, Palilalilave, Paqaluqalu.

2. Literature Review

Language choice behaviors have been studied from different aspects. The anthropological approach focuses on external factors that influence language choice and predicts possible results. On the contrary, the socio-psychological emphasizes inner reasons, such as psychological needs, that stimulate people to select certain language. The sociological approach, by contrast, aims at subcategorizing people's language choice behaviors, which meets the needs of this paper and thus is applied for analysis.

This study was conducted by questionnaire based on Chan's (1994) and Liu's (2002) surveys, which employed Fishman's domain analysis (1964). In domain analysis, it is proposed that one language appears to be more appropriate than another in certain domains, which are considered to be constellations of factors such as topic, location, and participants. For example, a home domain is made up by family members talking about everyday life at home. Domain analysis is often adopted to explain for the phenomenon of diglossia. Generally speaking, high language is selected for formal domains, such as education and government, while low language is used in informal domains, such as home and neighborhood. In a domain-oriented survey, subjects are often asked to indicate which language goes with that domain by rating the frequency of use.

3. Method

3.1 Overview

This study was made up of 67 native Paiwan speakers in the Taipei metropolitan area. Both survey and interview were employed to investigate the language choice behaviors of Paiwan aboriginals in Taipei and discover possible explanations.

3.2 Subjects and Sampling

In view of accuracy, subjects were selected from ten different places in Taipei downtown area, Taipei suburbs, and Taipei county. Subjects were Paiwan aboriginals who had dwelled in Taipei for more than ten years. What follows are the sampling areas:

- Downtown Taipei: Chong-cheng area, Ta-ann area, Ta-tung area, Wan-hua area
- Taipei Suburbs: Shih-lin area, Nei-hu area, Wen-shan area
- Taipei County: Hsin-tian city, Hsi-chih city, Tan-tsui town

For language proficiency is concerned, no special criterion was set but subjects were assured to be native Paiwan speakers. Since age and education level were hypothesized to be key

factors, different age and education groups were included on purpose when distributing survey questionnaires so that an uneven distribution can be prevented. In addition, significant difference of proficiency in Paiwan language was found to fall at the age of 40 and the education level of senior high school. Therefore, subjects were divided into older age (40 and over), younger age (under 40), high education (senior high school or above, and low education (junior high school or below) groups. Table 3.1 summarizes the information background of subjects.

Table 3.1: *Subjects' Background Information*

Main Groups	Subgroups	Number of Subjects
Age	Older Age	30
	Younger Age	37
Education	High Education	37
	Low Education	30
Religion	Christian	46
	Catholic	13
	No Religious Belief	8
Place of Growing Up	Taitung	37
	Pintung	14
	Taipei	15
	Hualian	1

3.3 Materials

The survey questionnaire was composed of three main parts, including basic information in open-ended question form, and language proficiency as well as language use in close-ended question form. In terms of basic background information, subjects were required to give sex, age, education level, place of growing up, and religion. As for language proficiency, a 5-point Likert Scale format was used to indicate subjects' proficiency in Paiwan language and Chinese Mandarin. The investigation on language use was operated by 3-point Likert Scale. Subjects were asked to choose a nearest point to represent their frequency of using Paiwan language and Chinese Mandarin in various domains. The survey questionnaire is provided in appendix 1.

A total number of 80 survey questionnaires were distributed, and 67 respondents were effective, with an effective response rate of 83.75%. Statistics were calculated using the T-Test and ANOVA. A $P < 0.05$ was considered as statistically significant.

4. Results and Discussion

This investigation was conducted to present the language proficiency and language use situation of Paiwan aboriginals in the Taipei metropolitan area. The previous study by Liu (2002) led us to infer that age and education level both play key roles in aboriginal speakers' language performance and language choice. In the following analysis, findings are first summed as whole and followed by crossed analysis by age and education level respectively.

4.1 Language Proficiency

Agreeing with the original hypothesis, Table 3.1 indicates that Paiwan aboriginals in Taipei have better abilities in Mandarin and the difference between languages is significant.

Comparing with Tsao's findings (1997), which says that aboriginals generally have similar proficiency in their native language and Mandarin, the results show that Mandarin has been well popularized in recent years. According to an aboriginal language activist, Hung-ming Po, this phenomenon can be attributed to some possible factors. First, aboriginals are obedient natured. Their belief in abiding by the rule of law, namely the Guoyu Zheng-ce practiced in the fifties, caused the high efficiency of popularizing Mandarin and brought a great impact on native language maintenance. Second, in terms of the Taipei metropolitan area, the shortage of native speakers may lower the chances of using Paiwan language. Third, as for Paiwan language is concerned, differences in dialect accents, which sometimes brings inconvenience in communication, prompt native speakers to use a unified language, Mandarin.

Table 4.1: *Language Proficiency of Paiwan People in Taipei Metropolitan Area*

Language Subject	Paiwan	Mandarin	T-Test
All	3.43	4.73	.000*

Like many other aboriginal languages, Paiwan is facing the crisis of language recession. Researchers and Paiwan language activists observe a correlation between age and Paiwan language proficiency. Table 3.2 shows that Paiwan 40 and over possess communicative ability in Paiwan, while Paiwan under 40 are found to communicate with difficulty and their production is limited to individual lexical items. By contrast, their ability in Mandarin is nearly equivalent. As for the difference between Paiwan and Mandarin abilities, not only in the lower age group but also in the higher age group, Mandarin ability is significantly superior to Paiwan, which implies that language shift has already taken place in both groups in terms of Paiwan people in the Taipei metropolitan area. The previous study on other aboriginal languages led us to assume that higher age group does not present a significant difference between the language abilities. The findings, however, go against the observation made several years ago. A possible explanation is that the Taipei metropolitan area is not a positive Paiwan speaking environment, and therefore language recession is faster than the Paiwan speaking environment, such as Taitung. However, the attribution to environmental factor should be confirmed with further contrastive study on the native speakers in Taipei and Taitung and will not be furthered discussed here.

Table 4.2: *Language Proficiency V.S Age*

Language Education	Paiwan	Mandarin	T-Test
High Education	3.14	4.76	.000*
Low Education	3.80	4.70	.001
ANOVA	.042	.633	

As for Paiwan language is concerned, education level plays another key role in the process of language recession. The reason is that the competing language, namely Mandarin, was used forcedly when the Guoyu Zheng-ce was practiced. Even nowadays,

Mandarin is still the main language used in school education. Therefore, speakers with higher educational background should be immersed in the Mandarin-speaking environment longer and possess better Mandarin ability. Table 3.3 supports the hypothesis that Mandarin of the higher education group is superior to that of the lower education group. Nevertheless, the number does not present a significant difference, which may suggest that Mandarin is widely used besides the location of school so that both groups can possess similar proficiency. On the other hand, Paiwan aboriginals' ability in their native tongue appears to be correlated with education. Paiwan people at low education level possess better proficiency in their native tongue than the high education group, and the difference is significant. Similar to the result of Table 4.2, both education levels present a significant difference between the abilities of Paiwan language and Mandarin. Subjects at high and low education levels have considerably superior proficiency in Mandarin, which again confirms the recession of Paiwan language in the Taipei metropolitan area.

Table 4.3: *Language Proficiency V.S Education Level*

Language \ Education	Paiwan	Mandarin	T-Test
High Education	3.14	4.76	.000*
Low Education	3.80	4.70	.001
ANOVA	.042	.633	

4.2 Language Use

Following Fishman's domain analysis, language use issue can be analyzed in three aspects, including conversation participants, conversation topics, and conversation locations.

4.2.1 Conversation Participants

The investigation on language use among different conversation participants was aimed to detect Paiwan language's symbolic functions of showing intimacy and solidarity. Intimacy regards conversation participants, such as family, colleagues, schoolmates, and strangers, while solidarity concerns the communication with people of the same and different tribes.

4.2.1.1 Intimacy

4.2.1.1.1 Family

Table 4.4 shows a contrastive result of language use from the viewpoint of conversation participants' generations. In terms of Paiwan, the frequency increases gradually when generations are older. On the contrary, Mandarin is used more and more when generations are younger. The differences in the use of Paiwan and Mandarin among generations both present significant differences. Referring back to the fact that age is correlated to proficiency in Paiwan language, it is possible that the choice of language is confined owing to limited proficiency in Paiwan language. For example, when communicating with children, Paiwan speakers have to use Mandarin because children have limited or no proficiency in Paiwan language. Comparing the use of Paiwan language and Mandarin among different generations, Mandarin is chosen more frequently than Paiwan language except for the situation of talking to grandparents, which marks the trace of language shift from Paiwan to Mandarin beginning at the parental generation. As for the difference between the use of

Paiwan language and Mandarin, the results of parents, siblings/spouse, and children are found to be significant, while the number of grandparents does not reach significance. Echoing to the finding that age is correlated to Paiwan proficiency but not Mandarin proficiency, it is likely that grandparents have communicative abilities both in Paiwan language and Mandarin, so language choice is freer. However, in parental generation and below, Paiwan aboriginals do not speak Paiwan language fluently, and thus Paiwan language is far less chosen when having conversation with them.

Table 4.4: *Language Use V.S Participants (Family)*

Language	Paiwan	Mandarin	T-Test
Family			
Grandparents	2.35	2.17	.158
Parents	2.19	2.50	.007
Siblings and Spouse	1.94	2.76	.000*
Children	1.61	2.90	.000*
ANOVA	.000*	.000*	

To further analyze the relationship between language use and age, subjects were again divided into higher and lower age groups, and the results were presented in Table 4.5. As for Grandparents and parents are concerned, Paiwan language is chosen more often than Mandarin by the older age group, while Mandarin is used more frequently by younger age group. Such differences in both groups are significant. Moreover, in terms of the choice of Paiwan language, the difference between older and younger age groups is significant and the situation overlaps with the choice of Mandarin. One possible reason is that Paiwan 40 and over have to use Paiwan language when talking to their grandparents and parents because of grandparents' and parents' limited proficiency in Mandarin. By contrast, speakers at younger age would use Mandarin much more frequently owing to their incompetence in Paiwan language. Considering conversation with siblings and spouses, both older and younger age groups use Mandarin more than Paiwan language, yet the difference is significant only in the younger age group. In terms of correlation of age and language choice, the use of Paiwan language presents a significant difference between older and younger age groups, but no significant difference is found in the use of Mandarin. As for conversation with children, Mandarin is still used at a larger degree than Paiwan language by both groups and the differences are significant. However, judging by the insignificant difference between the use of Paiwan language and Mandarin, it is likely that both age groups have similar choice behaviors when conversation participants are children. To sum up, when talking to family members, Paiwan 40 and over use Paiwan language more often than Mandarin, while Paiwan under 40 select Mandarin more frequently instead, and the differences are significant in both groups. Besides, the differences between two groups are also significant in terms of the use of Pawan language and Mandarin. That is to say speakers' age is an important factor influencing language choice behaviors when conversation participants are family members.

Table 4.5: *Language Use V.S Participants (Family) V.S Age*

Family	Language Age	Paiwan	Mandarin	T-Test
Grandparents	Older Age	2.92	1.42	.000*
	Younger Age	1.91	2.49	.000*
	ANOVA	.000*	.000*	
Parents	Older Age	2.86	1.84	.000*
	Younger Age	1.69	2.80	.000*
	ANOVA	.000*	.000*	
Siblings /Spouse	Older Age	2.39	2.68	.052
	Younger Age	1.51	2.82	.000*
	ANOVA	.000*	.198	
Children	Older Age	1.83	2.92	.000*
	Younger Age	1.40	2.88	.000*
	ANOVA	.076	.738	
All	Older Age	2.63	2.30	.000*
	Younger Age	1.68	2.73	.000*
	ANOVA	.000*	.000*	

The results of language use among family members in relation to education are shown in Table 4.6. When communicating with grandparents, Paiwan language is chosen at a higher rate by both high and low education groups; however, the differences between choice of Paiwan language and Mandarin are not significant. The findings suggest that education level does not play a key role in language choice when conversation subjects are grandparents. A further comparison of the differences between high and low education groups, only the choice of Paiwan language reaches significant difference, while the use of Mandarin is at a similar rate in both groups. As for parents are concerned, Paiwan language is more frequently adopted by lower education group, but a significant difference between language choices is found only in higher education group. The results reveal that when having conversation with parents, education influences the choice of Paiwan language but not Mandarin, which might be caused by the same factor in former analysis that speakers at high education level possess limited proficiency in Paiwan language. Regarding siblings and spouse as conversation participants, Mandarin is used at a larger degree than Paiwan language by both groups and the differences are significant. As for the differences between language choices, education is still correlated to the choice of Paiwan language but not Mandarin because only the choice of Paiwan language is significantly different between high and low education groups. An important finding is that when having conversation with children, Mandarin is far more used than Paiwan language by both groups, and no significant difference is found, which may imply that education level is not an important factor in language choices when conversation participants are children. To make a small conclusion, when conversation participants are family members, Mandarin is used at a larger degree than Paiwan language by both education groups, and the differences are significant. In terms of language choices, education level only plays an important role in the choice of Paiwan language, while Mandarin is used at a similar degree by both groups.

Table 4.6: *Language Use V.S Participants (Family) V.S Education Level*

Family	Language Education		Paiwan	Mandarin	T-Test
	Grandparents	High Education		2.15	2.11
Low Education			2.54	2.25	.100
		ANOVA	.008	.499	
Parents	High Education		1.97	2.55	.000*
	Low Education		2.44	2.42	.900
		ANOVA	.002	.390	
Siblings/Spouse	High Education		1.67	2.79	.000*
	Low Education		2.23	2.71	.003
		ANOVA	.000*	.526	
Children	High Education		1.5	2.9	.000*
	Low Education		1.75	2.9	.000*
		ANOVA	.320	1	
All	High Education		1.89	2.57	.000*
	Low Education		2.36	2.55	.029
		ANOVA	.000*	.858	

After analyzing the situation of language use among family members, an extensive comparison is made in a larger scope, namely intimacy. Table 4.7 indicates the use of Paiwan language and Mandarin from the most intimate, i.e. family members, to the least intimate, i.e. strangers. The findings reveal that Mandarin is more frequently adopted than Paiwan language in each intimacy level, and the differences are all significant. The results imply that Mandarin has successfully penetrated to the core of intimacy, family, where native language is believed to be most frequently used. Furthermore, the frequency of using Paiwan lowers with the decrease of intimacy to the standard of significant difference, except for the pairs of schoolmates/colleagues and strangers ($P=.150$). The findings suggest that Paiwan language symbolizes intimacy only to the intimacy level of schoolmates/colleagues. A possible explanation is that Mandarin is the main language used at school and work. Adding that most of schoolmates and colleagues are non-Paiwan speakers, Mandarin is therefore more frequently chosen.

As for Mandarin is concerned, it is used at a similar degree among friends, schoolmates/colleagues, and strangers; no significant difference is found in paired analysis. A lower frequency of using Mandarin among family members draws a clear distinction from the other three groups of participants. That is to say Mandarin is used at a considerably larger degree beyond the intimacy level of family.

Table 4.7: *Language Use V.S Intimacy*

Intimacy \ Language	Paiwan	Mandarin	T-Test
Family Members	2.11	2.56	.000*
Friends	1.82	2.93	.000*
Schoolmates Colleagues	1.35	2.92	.000*
Strangers	1.18	2.95	.000*
ANOVA	.000*	.000*	

Table 4.8 shows the correlation of language use, intimacy, and age. In terms of younger age group, Mandarin is consistently used more frequently than Paiwan in every intimacy level. Although older age group has the same preference for Mandarin when speaking to friends, schoolmates/colleagues, and strangers, Paiwan language is adopted more frequently when Paiwan 40 and over have conversation with the most intimate persons. All the differences between choices of language are significant in both groups in each level of intimacy. On the other hand, comparing the frequency of using Paiwan and Mandarin between older and younger age groups, the former surpasses the latter in using Paiwan, but the situation is opposite in using Mandarin. The differences are significant both in Paiwan language and Mandarin in each intimacy level, while an exception is found in using Mandarin when conversation participants are strangers. The results suggest that age is correlated to the use of Paiwan language at all intimacy levels, which may be attributed to younger age groups' limited proficiency in Paiwan language. As for Mandarin, it is used at a larger degree than Paiwan language in each level of intimacy, and the differences are significant from family to schoolmates/colleagues, while the difference is insignificant in the level of strangers. Overall speaking, Mandarin is used more frequently than Paiwan language by both groups and the differences are significant. Besides, age is an important factor influencing language choice when conversation participants are taken into account.

According to the interview, most Paiwan 40 and over Taipei residents grew up in Taitung, where Paiwan language was used frequently and often acquired as first language. Although immigration to the Taipei area decreases the frequency of using native language, Paiwan language is still frequently used when talking with family or friends over 40. By contrast, most Paiwan under 40 left their hometown, Taitung, during schooling or were even born in Taipei. Their social worlds are full of non-Paiwan speakers, and thus Paiwan language is far less chosen in daily communication, which indirectly influences their native language ability.

Table 4.8: *Language Use V.S Participants (Intimacy) V.S Age*

Persons	Language	Paiwan	Mandarin	T-Test
	Age			
Family	Older Age	2.63	2.30	.000*
	Younger Age	1.68	2.73	.000*
	ANOVA	.000*	.000*	
Friends	Older Age	2.30	2.82	.007
	Younger Age	1.41	3	.000*
	ANOVA	.000*	.009	
Schoolmates /Colleagues	Older Age	1.79	2.82	.000*
	Younger Age	1.12	2.99	.000*
	ANOVA	.000*	.000*	
Strangers	Older Age	1.5	2.88	.000*
	Younger Age	1.04	3	.000*
	ANOVA	.006	.129	
All	Older Age	2.37	2.63	.000*
	Younger Age	1.45	2.88	.000*
	ANOVA	.000*	.000*	

Considering the correlation of education and language use, Mandarin is used more than Paiwan language by both education groups in every intimacy level and the differences are all significant. Comparing the use of Paiwan language and Mandarin, the former is used more frequently by low education group in every level of intimacy and the frequency falls with the decrease of intimacy in both groups. By contrast, Mandarin is used at a larger degree by high education group in every level of intimacy, yet a regular distribution is not found and significant difference is only observed in the level of schoolmates and colleagues. In terms of Paiwan language, the differences are all significant; however, only the level of schoolmates/colleagues shows significant difference in the use of Mandarin. Responding to the fact that education level is correlated to proficiency in Paiwan language, it is likely that high education group can not use Paiwan language as fluently as low education group, which lowers their willingness in choosing Paiwan language for communication.

Table 4.9: *Language Use V.S Intimacy V.S Education Level*

Persons	Language Education	Paiwan	Mandarin	T-Test
	Family	High Education	1.89	2.57
Low Education		2.36	2.55	.029
ANOVA		.000*	.857	
Friends	High Education	1.54	2.97	.000*
	Low Education	2.18	2.87	.000*
	ANOVA	.004	.158	
Schoolmates /Colleagues	High Education	1.21	2.98	.000*
	Low Education	1.63	2.83	.000*
	ANOVA	.000*	.000*	
Strangers	High Education	1.05	3	.000*
	Low Education	1.38	2.89	.000*
	ANOVA	.035	.157	
All	High Education	1.60	2.81	.000*
	Low Education	2.13	2.72	.000*
	ANOVA	.000*	.034	

3.2.1.2 Solidarity

The second subgroup of conversation participants is concerning solidarity, which is composed of tribespeople, non-tribespeople, and presence of both tribespeople and non-tribespeople. The findings reveal that Mandarin is used more frequently than Paiwan language in each level of solidarity, and the differences are all significant. Furthermore, the frequency of using Paiwan language decreases progressively with the lowering of solidarity. A significant difference is found in the pair of tribespeople and presence of both tribespeople and non-tribespeople ($P=.000$). On the contrary, the frequency of using Mandarin increases by degrees, and differences between each pair are significant ($P=.000$, $P=.022$). The findings suggest that Paiwan language symbolizes solidarity in the core level, judging by the fact that the difference between tribespeople and presence of both tribespeople and non-tribespeople is significant. However, the competing language, Mandarin, has reached its popularity, which brings a great impact on the use of Paiwan language. Besides, according to the interview, Paiwan aboriginals prefer to speak Mandarin so that misunderstandings caused by different accents can be prevented. Only when ascertaining that the other party has the same accent, Paiwan language would be taken into consideration.

Table 4.10: *Language Use V.S Solidarity*

Language Solidarity	Paiwan	Mandarin	T-Test
Tribespeople	2.05	2.59	.000*
Presence of both tribespeople and non-tribespeople	1.57	2.79	.000*
Non-tribespeople	1.44	2.95	.000*
All	1.75	2.79	.000*
T-Test	.000*	.001	

Table 4.11 is a further analysis of the correlation of language use, solidarity and age. The results show that Paiwan language is used more frequently than Mandarin by Paiwan 40 and over when conversation subjects are tribespeople, but such difference between using Paiwan and Mandarin is not significant. In the other levels of solidarity, Mandarin is used at a larger degree than Paiwan language by both groups, and the differences are all significant. The findings imply that Paiwan language's symbolic function of solidarity is mainly maintained by older age group who speaks native language more often during the communication with tribespeople. From the perspective of age, it is correlated with language choice in inner levels of solidarity. Generally speaking, Paiwan language is used more often by older age group, while Mandarin is spoken more frequently by younger age group. The differences reach significance in each paired analysis. However, when conversation subjects are non-tribespeople, age is not determinant in language choice. In both groups, Mandarin is used far more frequently than Paiwan language, and the differences are significant. In addition, the choice of Paiwan language is similarly low in both groups, while that of Mandarin is similarly high. No significant difference is observed. Generally speaking, Mandarin is used more often than Paiwan language in both groups and the differences are significant. The findings may suggest that the convenience of communication in a popular language has overridden the purpose of expressing solidarity. As for the role of age, concluding by the significant differences between groups, it is an important factor that influences the frequency of using Paiwan language as well as Mandarin in terms of solidarity. Echoing back the fact that Paiwan under 40 have limited proficiency in their native language, it is fairly likely that their choice of language is confined and thus expands the difference between groups.

Table 4.11: *Language Use V.S Solidarity V.S Age*

Solidarity	Language	Paiwan	Mandarin	T-Test
	Age			
Tribespeople	Older Age	2.48	2.29	.382
	Younger Age	1.72	2.74	.000*
	ANOVA	.000*	.017	
Presence of both tribespeople and noon-tribespeople	Older Age	2	2.59	.002*
	Younger Age	1.33	2.97	.000*
	ANOVA	.001*	.000*	
Non-tribespeople	Older Age	1.73	2.88	.000*
	Younger Age	1.32	3	.000*
	ANOVA	.084	.114	
All	Older Age	2.18	2.63	.000*
	Younger Age	1.48	2.91	.000*
	ANOVA	.000*	.000*	

In the view of education, as presented in Table 4.12, speakers of both groups use more Mandarin than Paiwan language in each level of solidarity, and the differences are all significant except for low education group in the level of tribespeople. Although the former analysis says that the difference between low education group's proficiency in Paiwan and Mandarin is significant, the result points out that Paiwan language is still used at a similar degree with Mandarin in the core level of solidarity. Nevertheless, beyond the intimacy level of tribespeople, Paiwan language is spoken far less than Mandarin by both groups. A possible explanation is that Paiwan language, although not mastered even by low education group, still plays a key role in expressing solidarity. Therefore, when having conversation with tribespeople, low education group speaks Paiwan language almost as frequently as Mandarin. However, high education group is limited in language choice owing to their infamiliarity with Paiwan language, which draws a clear distinction in their use of languages. An interesting finding is that although the frequency of speaking Paiwan between groups is not significantly different in each level of solidarity, an overall average reveals that education is still a determinant influencing the frequency of speaking Paiwan language in terms of solidarity. By contrast, no significant difference is found in Mandarin in each level of solidarity as well as the total average.

Table 4.12: *Language Use V.S Solidarity V.S Education Level*

Solidarity \ Education		Language	Paiwan	Mandarin	T-Test
		Education			
Tribespeople	High Education		1.89	2.60	.000*
	Low Education		2.25	2.56	.187
	ANOVA		.062	.846	
Presence of both tribespeople and noon-tribespeople	High Education		1.55	2.87	.000*
	Low Education		1.63	2.71	.000*
	ANOVA		.720	.158	
Non-tribespeople	High Education		1.39	3	.000*
	Low Education		1.54	2.88	.000*
	ANOVA		.523	.142	
All	High Education		1.64	2.84	.000*
	Low Education		1.91	2.74	.000*
	ANOVA		.034	.194	

After the analysis by levels of intimacy and solidarity, the following discussion focuses on a general comparison of these two main factors in relation to language use. The results in Table 4.13 indicate that Mandarin is used more frequently than Paiwan language, and the differences are significant in terms of both factors. Besides, neither the use of Paiwan nor that of Mandarin shows a significant difference in intimacy and solidarity.

Table 4.13: *Language Use V.S Affectional Factors*

Language \ Factors		Paiwan	Mandarin	T-test
Intimacy		1.83	2.77	.000*
Solidarity		1.75	2.79	.000*
T-Test		.295	.651	

A further analysis of the role of age is presented in Table 4.14. The findings reveal that Mandarin is used more frequently than Paiwan language by both groups in terms of intimacy as well as solidarity, which might imply that symbolic functions performed by low language can not override the requirement on communication efficiency brought by high language in the Taipei metropolitan area. In addition, in view of language choice, age is still a key factor, judging by the significant difference in each situation. Finally, comparing the frequency of using Paiwan language, intimacy is likely to be more stimulative for older age group than solidarity, and younger age group's use of Paiwan language is slightly higher in solidarity than in intimacy. However, a further paired analysis does not show significant difference.

Table 4.14: *Language Use V.S Affective Factors V.S Age*

Factors	Language		Paiwan	Mandarin	T-Test
	Age				
Intimacy	Older Age		2.37	2.63	.000*
	Younger Age		1.45	2.88	.000*
	ANOVA		.000*	.000*	
Solidarity	Older Age		2.18	2.63	.000*
	Younger Age		1.48	2.91	.000*
	ANOVA		.000*	.000*	

Similar to the findings of age, Table 4.15 shows that Mandarin is used more frequently than Paiwan language by both education groups in view of intimacy and solidarity, and the differences are all significant. For the use of Paiwan language is concerned, low education group speaks Paiwan language more often than high education group in both factors. By contrast, high education group use Mandarin more frequently, but significant difference is found only in the factor of intimacy.

Table 4.15: *Language Use V.S Affective Factors V.S Education*

Factor	Language		Paiwan	Mandarin	T-Test
	Education				
Intimacy	High Education		1.60	2.81	.000*
	Low Education		2.13	2.72	.000*
	ANOVA		.000*	.034	
Solidarity	High Education		1.64	2.84	.000*
	Low Education		1.91	2.74	.000*
	ANOVA		.034	.194	

4.2.2 Conversation Topic

This section discusses the situation of language use in various conversation topics and its correlation with age and education.

The overall results are shown in Table 4.16. Generally speaking, Mandarin is used at a larger scale than Paiwan language in every conversation topic, and the differences are significant. In view of language choice, Paiwan language is adopted the most frequently when talking about daily life, but the least when the topic is regarding professional knowledge. As for Mandarin, the frequency is high in every topic, but the climax falls on professional knowledge. The results suggest that the frequency of using Paiwan language decreases when conversation topic requires special terms that Paiwan lexicon is lack of. In fact, the interviewees also express that the shortage of special terminology in Paiwan language raises ties in communication. Many modern inventions, proper names, or academic terms are not translated into Paiwan language. Adding that much information, such as news or professional knowledge, is released in Chinese, it is much easier for them to communicate in Mandarin.

Table 4.16: *Language Use V.S Topic*

Topic \ Language	Paiwan	Mandarin	T-test
Daily Life	1.76	2.83	.000*
Social Issues	1.44	2.86	.000*
Professional Knowledge	1.30	2.89	.000*
All	1.52	2.86	.000*

The correlation of age, topic, and language use is presented in Table 4.17. In every topic, Mandarin is used more frequently than Paiwan language by both groups, and the differences are significant. Overall speaking, Paiwan language is used more by Paiwan 40 and over, while Mandarin by Paiwan under 40, and the differences of language use of both groups are significant. Besides, in both groups, the frequency of using Paiwan language decreases progressively while that of Mandarin increases by degrees, which overlaps with the distribution in the findings presented in Table 4.16. Echoing the previous analysis, age is a determinant both in the use of Paiwan language and Mandarin, for the differences between groups are significant in every situation.

Table 4.17: *Language Use V.S Topic V.S Age*

Topic \ Language \ Age	Paiwan	Mandarin	T-Test	
Daily Life	Older Age	2.14	2.63	.019
	Younger Age	1.46	2.94	.000*
	ANOVA	.002	.012	
Social Issues	Older Age	1.82	2.71	.000*
	Younger Age	1.19	2.97	.000*
	ANOVA	.003	.004	
Professional Knowledge	Older Age	1.67	2.74	.000*
	Younger Age	1.08	3	.000*
	ANOVA	.000*	.001	
All	Older Age	1.91	2.70	.000*
	Younger Age	1.25	2.97	.000*
	ANOVA	.000*	.000*	

In contrast to age, the role of education level appears to be indeterminate in its relation to conversation topic. As shown in Table 4.18, Mandarin is spoken more frequently than Paiwan language by both groups regardless of conversation topics, and all the differences are significant. From the perspective of language, Paiwan language is adopted at a similarly low degree by both groups, while Mandarin is used at a high degree. No significant difference is found between the groups in the use of both languages. Similar to the results in Table 4.17, the frequency of using Paiwan language decreases as conversation topics require special terminology, yet Mandarin shows increase instead.

However, an interesting finding is that although education level is not a key factor in view of each topic, the overall average reveals positive correlation of education level and topic in both language uses. A possible explanation is the increase and decrease of frequency is not progressive in most situations. To be more specific, high education group, owing to their limitation in language proficiency, shows a sudden drop of frequency when conversation topic changes from daily life to social issues. Supporting by high education group's similarly soaring frequency of adopting Mandarin in the discussion of social issues and professional knowledge, it is likely that the difficulty of speaking native language raises abruptly for high education group from the degree of social issues. Because of the unprogressive distribution, the overall average indicates a contrary result, i.e. significant differences, to the individual analysis.

Table 4.18: *Language Use V.S Topic V.S Education*

Topic	Language Education		Paiwan	Mandarin	T-Test
	Daily Life	High Education		1.61	2.87
Low Education			1.95	2.76	.000*
ANOVA			.126	.375	
Social Issues	High Education		1.28	2.94	.000*
	Low Education		1.67	2.76	.000*
	ANOVA			.074	.051
Professional Knowledge	High Education		1.17	2.94	.000*
	Low Education		1.5	2.83	.000*
	ANOVA			.066	.171
All	High Education		1.36	2.92	.000*
	Low Education		1.73	2.78	.000*
	ANOVA			.004	.017

4.2.3 Conversation Location

The final analysis emphasizes the correlation of locations and language use. Locations are divided into home, tribe (such as religious or festival gathering), school, working place, public places (such as governmental or private institutes), and business occasions (such as department store or market). The following table indicates that Mandarin is used more often than Paiwan language in every location, and the differences are all significant. The findings point out that Mandarin has brought a serious impact on Paiwan language in terms of conversation locations. Especially in locations like tribe and home, which suggest solidarity and intimacy respectively, the communication tool is not low language but high language being used at a much larger degree.

In terms of the use of Paiwan language, the frequency decreases in the order of Tribe > Home > School > Working Place > Business > Public Places. According to the interview, it is a custom for tribal gathering, such as meeting or festival, to be conducted with an opening speech in Paiwan language; therefore, tribal gathering is considered to be the occasion where Paiwan language is used the most frequently. As for Mandarin, the

frequency increases in the order of Tribe < Home < Working Place < School < Business Occasions < Public Places. The inconsistency of these two sequences shows an interesting finding: both Paiwan language and Mandarin are spoken more frequently at school than at work, which implies that language choice is freer at school than at work.

Table 4.19: *Language Use V.S Conversation Locations*

Language Locations	Paiwan	Mandarin	T-Test
Tribe	1.97	2.66	.000*
Home	1.92	2.80	.000*
Working Places	1.47	2.89	.000*
School	1.54	2.92	.000*
Public Places	1.16	2.96	.000*
Business Occasions	1.17	2.95	.000*
All	1.62	2.86	.000*

Similar to the findings shown in Table 4.19, the paired analysis of Table 4.20 indicates that Mandarin is used more frequently than Paiwan language by both age groups in every occasion, and the differences are entirely significant. From the perspective of language, Paiwan language is used more often by older age group and Mandarin is spoken more frequently by younger age group. The frequency of using Paiwan language by older age group is Home > Tribe > Working Place > School > Public Places > Business Occasions, while that by younger age group is Home > Tribe > School > Working Place > Business Occasions > Public Places. Significant differences between groups are found in the occasions of home, tribe, working place, and school and the total average of using Paiwan language. In terms of Mandarin, it is used by older age group in an increasing degree by the order of Tribe < Home < Working Place < School < Business Occasions < Public Places, while that for younger age group is Home < Tribe < Working Place < School = Public Places = Business Occasions; significant differences are found only in the occasions of tribe and working place. Generally speaking, age does not differentiate the use of language in business occasions and public places because the differences between groups are not obvious. By contrast, home as and school are locations where the use of Paiwan language is significantly different between groups. Among all locations, only the use of language in tribes and working places reveal significant differences. To explain for the results, business occasions and public places are full of non-Paiwan aboriginals, and therefore the frequency of using Paiwan language is low and that of Mandarin is relatively high. As for home is concerned, it seems that age plays an important role only in the use of Paiwan language, which overlaps with correlation of age and proficiency in both languages. As for school is concerned, age still plays a key role in the use of Paiwan language, while the influence on the use of Mandarin is obscure. Finally, in view of tribe and working place, age influences the use of both languages. That is to say the language choice behaviors of both groups are the most dissimilar when conversation takes place in tribes and at work.

Table 4.20: *Language Use V.S. Conversation Locations V.S Age*

Locations	Language Age	Paiwan	Mandarin	T-Test
Home	Older Age	2.23	2.76	.000*
	Younger Age	1.68	2.82	.010
	ANOVA	.006	.684	
Tribe	Older Age	2.06	2.60	.000*
	Younger Age	1.46	2.88	.000*
	ANOVA	.000*	.021	
School	Older Age	1.65	2.88	.000*
	Younger Age	1.31	3	.000*
	ANOVA	.034	.059	
Working Place	Older Age	1.83	2.79	.000*
	Younger Age	1.22	2.97	.000*
	ANOVA	.003	.048	
Public Places	Older Age	1.18	2.95	.000*
	Younger Age	1.12	3	.000*
	ANOVA	.584	.207	
Business Occasions	Older Age	1.15	2.93	.000*
	Younger Age	1.20	3	.000*
	ANOVA	.670	.118	
All	Older Age	1.74	2.79	.000*
	Younger Age	1.50	2.91	.000*
	ANOVA	.000*	.000*	

Finally, the correlation of locations, language use, and education level is presented in Table 4.21. In every occasion, Mandarin is used at a larger degree than Paiwan language by both groups, and the differences are all significant. Generally speaking, Paiwan language is spoken more frequently by low education group and Mandarin is used more often by high education group. In terms of the sequence of frequency, Paiwan language is used at a decreasing degree by high education group in the order of Tribe > Home > School > Working Place > Business Occasions > Public Places, while the order for low education group is Home > Tribe > School > Working Place > Public Places > Business Occasions. As for Mandarin, it is used at an increasing degree by both groups in the same order of Tribe < Home < Working Place < School < Business Occasions < Public Places. Accordingly, high and low education groups tend to use Mandarin in different occasions in a similar way; however, their preferences for Paiwan-speaking occasions are dissimilar. Furthermore, from the perspective of locations, the use of Paiwan language in the occasions of home and school are significantly different between groups, while no significant difference is found in the other situations. In other words, people of different education levels have similar language choice behaviors except for the using Paiwan language at home and school. A possible explanation is that low education group possesses better proficiency in Paiwan language and therefore their frequency in using Paiwan language is naturally high in the occasions where it is mostly spoken. As for school, since

Mandarin is the mainstream language, the length of schooling may enlarge the difference between groups.

Table 4.21: *Language Use V.S. Conversation Locations V.S Education Level*

Locations	Language Education	Paiwan	Mandarin	T-Test
Home	High Education	1.64	2.83	.000*
	Low Education	2.26	2.75	.014
	ANOVA	.002	.598	
School	High Education	1.42	2.95	.000*
	Low Education	1.72	2.86	.000*
	ANOVA	.045	.154	
Working Places	High Education	1.36	2.94	.000*
	Low Education	1.60	2.82	.000*
	ANOVA	.254	.191	
Public Places	High Education	1.15	2.98	.000*
	Low Education	1.19	2.93	.000*
	ANOVA	.687	.151	
Business Occasions	High Education	1.23	2.96	.000*
	Low Education	1.07	2.94	.000*
	ANOVA	.162	.788	
Tribe	High Education	1.86	2.62	.000*
	Low Education	2.10	2.72	.000*
	ANOVA	.050	.321	
All	High Education	1.51	2.86	.000*
	Low Education	1.77	2.85	.000*
	ANOVA	.000*	.627	

From the viewpoint of age, Paiwan language is used more frequently by older age group, and Mandarin is more often chosen by lower age group. The differences are all significant in terms of conversation participants, topics, and locations. Judging by the frequency of using Paiwan language, it seems that Paiwan 40 and over are more confined by conversation locations, while for Paiwan under 40, conversation topic is more determinant. Such limitations in using Paiwan language would limit Paiwan speakers' language choices and therefore directly increases the frequency of using mainstream language. The results are summed in Table 4.22 below.

Table 4.22: *Language Use V.S. Factors V.S Age*

Factor \ Language Age	Paiwan	Mandarin	T-Test	
Persons	Older Age	2.34	2.63	.000*
	Younger Age	1.45	2.89	.000*
	ANOVA	.000*	.000*	
Topics	Older Age	1.91	2.70	.000*
	Younger Age	1.25	2.97	.000*
	ANOVA	.000*	.000*	
Places	Older Age	1.74	2.79	.000*
	Younger Age	1.50	2.91	.000*
	ANOVA	.000*	.001*	

Similar to the findings by age, both education groups use Mandarin more than Paiwan language, and the differences are significant. Education level seems to be a key factor in the use of Paiwan language, because the differences between groups are significant in each factor. By contrast, in the use of Mandarin, education level is only determinate in terms of persons and topics. The results are summed in Table 4.23.

Table 4.23: *Language Use V.S. Factors V.S Education Level*

Factor \ Language Education	Paiwan	Mandarin	T-Test	
Persons	High Education	1.61	2.82	.000*
	Low Education	2.09	2.73	.000*
	ANOVA	.000*	.014	
Topics	High Education	1.36	2.92	.000*
	Low Education	1.73	2.78	.000*
	ANOVA	.004	.017	
Places	High Education	1.51	2.86	.000*
	Low Education	1.77	2.85	.000*
	ANOVA	.000*	.627	

5. Conclusion

The aim of this paper is to probe into the language choice behaviors of Paiwan people in the Taipei metropolitan area. Based on survey and interview, the analysis first begins with a language proficiency investigation, followed by exploration of various influential factors under three main categories, namely persons, topic, and places. In this study, age and education are considered to be the focal variables, and their correlation with language use is discussed case by case.

In view of language proficiency, Paiwan people in the Taipei metropolitan area have better ability in Mandarin, and the difference between Paiwan and Mandarin is significant. Generally speaking, Paiwan 40 and over and those with low education

background possesses better fluency in Paiwan language, while Paiwan aboriginals' difference in the ability of Mandarin is not distinguishable either by age or education.

To make a conclusion, according to the study, the use of Paiwan language in the Taipei metropolitan area is facing a serious threatening of the mainstream language, which is spoken by both age and education groups at a considerably high degree in nearly every situation regarding persons, topics, and places. A clear distinction between age and education groups is found in the use of Paiwan language, which suggests the loss of Paiwan language in different dimensions. According to Dorian (1980), the three symptoms of language death are fewer speakers, fewer domains of use, and structural simplification, which is proved to be the extremity of situation that Paiwan language is encountering now.

References

- Dorian, Nancy C. 1980. "Language shift in community and individual: The phenomenon of the laggard semi-speaker." *International Journal of the Sociology of Language* 25:85-94.
- Dorian, Nancy C. 1981. *Language death: The life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.
- Fasold, Ralph. 1984. *The Sociolinguistics of Society*. Oxford: Blackwell.
- 1990. *The Sociolinguistics of Language*. Oxford: Blackwell.
- Gumperz, John. 1982. *Discourse Strategies*. New York: Cambridge University Press.
- Hall, Stuart, and de Gay, Paul. 1996. *Questions of Cultural Identity*. London: Sage.
- Heller, Monica. 1988. Strategic ambiguity: codeswitching in the management of conflict. In Heller, Monica (ed.) *Codeswitching: Anthropological and Sociolinguistic Perspectives*, pp.77-97. New York: Mouton de Gruyter.
- Ho, Dan-An. (1995). Paiwan Dictionary. In *Comparative Austronesian Dictionary*. D. Tryon, ed. Pp. 307-314. Berlin: Mouton de Gruyter.
- Ho, Szu-yin and Liu, I-chou. 2002. "The Taiwanese/Chinese Identity of the Taiwan People in the 1990s," *American Asian Review*, Vol. XX, No. 2. Summer, pp. 29-75.
- Huang, Shuanfan. 2000. 'Language, identity and conflict: a Taiwan study', *International Journal of Sociology of Language*, 143:139-149.
- Li, Wei. 1998. "The 'Why' and 'How' Questions in The Analysis of Conversational Codeswitching". In *Codeswitching in Conversation: Language, Interaction and Identity*, pp. 156-179, (ed.) by Peter Auer, New York: Routledge.
- Information on Indigenous Peoples. 2002. *Council of Indigenous Peoples, Executive Yuan*. Retrieved December 10th, 2005. from <http://gina.nat.gov.tw/rdec1/agent/A-P/a00000108-apc-news/node2/rule1/1.htm>
- Myers-Scotton, Carol. 1983. "The negotiation of identities in conversation: a theory of Markedness and code choice", *International Journal of Sociology of Language*, 44:115-136.
- 1993. *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon (Oxford University Press).[Reprint edition 1995].
- 1998. A theoretical introduction to the Markedness model. In C. Myers-Scotton, ed., *Codes and consequences: Choosing linguistic varieties*, 18-38. New York: Oxford University Press.
- 1999. "Explaining the role of norms and rationality in codeswitching", *Journal of Pragmatics* 32:1259-1271.

- Stroud, Christopher. 1992. 'The Problem of intention and meaning in codeswitching,' Text12 (1)
- Tse, John Kwock-Ping. 2000. "Language and A Rising Identity in Taiwan," International Journal of Sociology of Language.
- Trudgill, Peter. 2000. *Sociolinguistics: An introduction to language and society*. London: Penguin Books.
- 王思斌 (1996)。中國人際關係初級化與社會變遷 《管理世界》第3期
- 張秀絹 (2000)。排灣語參考語法。台北:遠流。
- 劉秋雲 (2002)。台灣地區原住民母語教育政策之探討: 布農族為例。國立政治大學語言學研究所碩士論文(未出版)。

Appendix 1: Survey Questionnaire

I. Personal Information: please check in the proper box

Sex: M___, F___

Age: _____

Education Level: Self-study (Never attend schooling)_____, elementary school____, junior high school____, senior high school____, university and above_____.

Occupation: Government Employees ____, Teaching____, Church staff____, Civil____, Business____, farming/fishing____, military____, students____, service____, free lancer____, unemployed____, others _____ (please specify)

Tribe _____

Religion: Catholic____, Christian ____, Buddhist____, Dao____, None____, Others _____ (Please specify)

Mother's tribal background _____

Father's tribal background _____

Spouse's tribal background _____ (skip if unmarried)

Language Proficiency:

Fluency Language	Fluent	Good	Fair	Weak	None
Paiwan					
Mandarin					

II. Language Use

What are your choices of languages in the following situations?

Please note that both languages in every situation have to be marked.

3= often, 2= sometimes, 1= seldom or never

Situations	Paiwan			Mandarin		
1. Talking to grandfather	3	2	1	3	2	1
2. Talking to grandmother	3	2	1	3	2	1
3. Talking to father	3	2	1	3	2	1
4. Talking to mother	3	2	1	3	2	1
5. Talking to siblings	3	2	1	3	2	1
6. Talking to spouse	3	2	1	3	2	1
7. Talking to children	3	2	1	3	2	1
8. Talking to friends	3	2	1	3	2	1
9. Talking to supervisors	3	2	1	3	2	1

10. Talking to colleagues	3	2	1	3	2	1
11. Talking to subordinates	3	2	1	3	2	1
12. Talking to teachers	3	2	1	3	2	1
13. Talking to classmates	3	2	1	3	2	1
14. Talking to strangers	3	2	1	3	2	1
15. Talking to tribespeople	3	2	1	3	2	1
16. Talking to non-tribespeople	3	2	1	3	2	1
17. Talking when both tribespeople and non-tribespeople are present	3	2	1	3	2	1
18. Talking about daily topics	3	2	1	3	2	1
19. Talking about social issues	3	2	1	3	2	1
20. Talking about professional knowledge	3	2	1	3	2	1
21. Talking at home	3	2	1	3	2	1
22. Talking at school	3	2	1	3	2	1
23. Talking after class	3	2	1	3	2	1
24. Talking at work	3	2	1	3	2	1
25. Talking at public institutes	3	2	1	3	2	1
26. Talking at private institutes	3	2	1	3	2	1
27. Talking at department stores	3	2	1	3	2	1
28. Talking at markets	3	2	1	3	2	1
29. Talking at churches or temples	3	2	1	3	2	1
30. Talking at tribal meeting	3	2	1	3	2	1
31. Talking at tribal ceremonies	3	2	1	3	2	1

Appendix 2: Interview Questions

A. Basic Information

How old are you?

Where did you grow up?

Which tribe do you belong to?

What's your religion?

Are both of your parents Paiwan aboriginals? Do they teach you Paiwan language?

If parents are in a mixed marriage, which language do you use more often?

Do you speak Paiwan language well? Do you understand Paiwan language?

B. Language Use

Do you often talk in Paiwan language at home? (to grandfather / grandmother / father / other / siblings)

Do you talk in Paiwan language at school or at work? (to eachers/classmates/colleagues)

Do you talk too your good friends in Paiwan language often?

What language do you use when talking to strangers?

What language do you use when talking to tribespeople?

What language do you use when talking to non-tribespeople?

What language do you use when you talk about daily topics/homework/problems at work?

What language do you use when shopping at groceries?

What language do you use at churches or temples?

What language do you use at tribal ceremonies?

A DESCRIPTIVE STUDY OF COMPARATIVE CONSTRUCTIONS IN KAVALAN AND AMIS*

Li-May Sung
National Taiwan University
limay@ntu.edu.tw

Cheng-chuen Kuo
National Taiwan University
kuo.jona@msa.hinet.net

1. Introduction

In many languages, comparatives are based on specialized morphology and syntax. Languages such as English make use of a degree modifier *more/-er* and a special word *than* to introduce a comparative complement which could be either a phrase consisting of the standard of comparison alone or a clause containing the standard as a subconstituent, as illustrated in (1-2).¹

- 1) John is taller than Tom (is).
- 2) John has written more books than Tom (has).

While comparatives in Indo-European languages have attracted a fair amount of attention, related issues are barely addressed in the Austronesian literature. Our goal in this paper aims to explore this linguistic aspect starting with two neighbouring Formosan (Austronesian) languages, Kavalan and Amis, spoken in the east coast of Taiwan.² In this work our purpose is twofold: firstly, to provide a brief overview of the major devices for expressing comparison in Kavalan and Amis, and secondly, to examine how the essential components in a comparison, namely the comparee, the standard and the gradable predicate, interact with each other in terms of predication, case assignment and word order. We will start with Kavalan.

* This work is supported by the NSC project (NSC 94-2411-H-002-012), Taiwan, granted to the first author. We are grateful to our Kavalan and Amis consultants (Nengi, Buya, Abas, Ofad, Panay, and Usay) for their patience and passion.

¹ Comparison of inequality will be central to our discussion here. Others such as superlatives (e.g. *John is the tallest of the children*), equatives (e.g. *John is as tall as Tom*) or less typical comparatives involving with two different gradable predicates (e.g. *I'm more cunning than brave*) are beyond the scope of this paper. We will leave these for further research in the future.

² Amis data in this study are from the Haiian dialect (Coastal Amis), one of the Central dialects spoken in Changpin, Taitung County. Kavalan data are from the Hsinshe village, Hualien County. Both are predicate-initial languages.

2. Kavalan Comparative Constructions

Two types of comparatives are found in Kavalan, a ‘juxtaposed’ type and an ‘addity’ one.³

2.1. The ‘Juxtaposed’ Type

This is the commonest construction used in Kavalan comparatives. This type of comparatives is expressed by two straightforwardly parallel clauses to indicate comparison as shown in (3).⁴

- (3) a. ibabaw ti-abas, ilibeng ti-utay
 AF.top NCM-PN AF.below NCM-PN
 ‘Abas is taller than Utay.’
 (Lit: ‘Abas is tall; Utay is short.’)
- b. nengi ti-abas, mai nengi ti-utay
 good NCM-PN NEG good NCM-PN
 ‘Abas is better than Utay.’
 (Lit: ‘Abas is good; Utay is not.’)

Each comparative above contains two gradable predicates which mark a contrast by means of a pair of lexical items standing in polar opposition *tall-short* as in (3a), or showing a positive-negative polarity *good-not good* as in (3b). The distribution of word order, verbal voices, nominal case marking in this type of comparatives is no difference from non-comparative constructions as discussed in Kavalan (Li 1996, 1997; Chang 1997, 2000; Lee 1997).

2.2. The ‘Addity’ Type

In this type, Kavalan mainly employs a strategy of addity of a standard NP without resorting to any degree modifier or *than* particle. The comparative construction as exemplified in (4) does not differ from a non-comparative counterpart (5) except for one added standard argument.

- (4) a. m-sanem (ya) singsi ‘nay tu pataqsian ‘nay
 AF-smart NOM teacherthat OBL student that
 ‘That teacher is smarter than that student.’
- b. Rangazi ti-utay ti-abas(-an)
 AF.rich NCM-PN NCM-PN(-LOC)

³ Kavalan also makes use of a fixed expression *wiya=ti* ‘leave-PFV’ to indicate a self-changing comparison meaning *getting more/getting -er*. This is illustrated in (i).

(i) wiya=ti misi ti-utay
 leave=PFV AF.fat NCM-PN
 ‘Utay is getting thinner.’

The object of comparison, *Utay*, in (i) is compared with himself, the standard. Since the standard argument is typically not expressed in such a construction, we deliberately leave it aside in this paper.

⁴ The following abbreviations are used in this paper: 1SG: 1st person singular; 2SG: 2nd person singular; 3SG: 3rd person singular; ACC: accusative; AF: agent focus/voice; CAU: causative; CONJ: conjunction; GEN: genitive; IF: instrumental focus/voice; LF: locative focus/voice; LNK: linker; LOC: locative; NCM: noun-class marker; NAF: non-agent focus/voice; NEG: negative marker; NOM: nominative; OBL: oblique; PF: patient focus/voice; PFV: perfective; PN: proper noun; RED: reduplication.

- ‘Utay is richer than Abas.’
- (5) a. m-sanem (ya) singsi ‘nay
AF-smart NOM teacher that
‘That teacher is smart.’
- b. Rangazi ti-utay
AF.rich NCM-PN
‘Utay is rich.’

The ‘added’ standard argument *pataqsian* ‘student’ in (4a) is syntactically placed as the direct object, in parallel to a transitive structure predicated with a verb such as hit in (6).

- (6) p<m>ukun (ya) singsi ‘nay tu pataqsian ‘nay
<AF>hit NOM teacher that OBL student that
‘That teacher hit that student.’

As shown in (4), the gradable predicate *msanem* ‘smart’ or *Rangazi* ‘rich’ does not undergo any morphological or syntactic processes as in English, where a degree modifier *more/-er* and a comparative complement introducer *than* are needed. There is, however, one exception when the nominal kinship terms *qaqa/suani* ‘older sibling/younger sibling’ are manipulated in comparing ages. They must undergo a morphological process, i.e. Ca reduplication, to derive a comparative meaning. Examples are given below:

- (7) a. qaqa ti-abas
older.sibling NCM-PN
‘Abas is the older brother.’
- b. suani ti-utay
younger.sibling NCM-PN
‘Utay is the younger brother.’
- (8) a. qa-qaqa ti-abas tu pataqsian ‘nay
RED-older.sibling NCM-PN OBL student that
a’. *qaqa ti-abas tu pataqsian ‘nay
older.sibling NCM-PN OBL student that
‘Abas is older than that student.’
- b. sa-suani ti-utay tu pataqsian ‘nay
RED-younger.sibling NCM-PN OBL student that
- b’. *suani ti-utay tu pataqsian ‘nay
younger.sibling NCM-PN OBL student that
‘Utay is younger than that student.’

Most examples of the ‘addity’ comparatives are instantiated with agent voices (AF); non-agent voices such as LF are perfectly acceptable as long as the example is in the right context. This is shown in (9).

- (9) Q: ibabaw qaqa-su timaiku
AF.tall older.brother-2SG.GEN 1SG.OBL
A: en. (qa-)ibabaw-an-na aisu
Yeah QA-tall-LF-3SG.GEN 2SG.NOM

Q: 'Is your brother taller than me?'

A: 'Yeah, he is taller than you.'

Various word order patterns such as VSO, SVO, or VOS are attested in AF comparative constructions if case markers are overtly present, in accordance with the observation made in Chang (1997: 24-27) for non-comparative constructions in Kavalan.⁵

- (10) a. misi ti-angaw ti-utay(-an)
 AF.fat NCM-PN NCM-PN(-LOC)
 b. ti-angaw misi ti-utay(-an)
 NCM-PN AF.fat NCM-PN(-LOC)
 c. misi ti-utay*(-an) ti-angaw
 AF.fat NCM-PN*(-LOC) NCM-PN
 'Angaw is fatter than Utay.'

This strategy of adding one standard argument in most cases is confined to one-place gradable predicates such as *fat*, *smart*, *tall* etc. Such an additivity strategy does occasionally apply to two-place predicates such as *like* in (11); it is quite restricted, however, and it is somehow lexically idiosyncratic. With other predicates such as *angry* in (12), the arguments are often interpreted not as a comparison relationship, but as two conjoined constituents.

- (11) m-rizaq=iku ti-abas ti-awun
 AF-happy=1SG.NOM NCM-PN NCM-PN
 'I like Abas more than (I like) Awun.'
- (12) qunut=iku ti-awun ti-abas
 <AF>angry=1SG.NOM NCM-PN NCM-PN
 'I am angry at Awun and Abas.'
 (*'I am angry at Awun more than (I am) at Abas.')

3. Amis Comparative Constructions

We now turn to examine the comparative constructions in Amis. In Amis, there are four types of comparatives: juxtaposed type, nonimal type, -ki- type and ikaka/isafa type. Each will be discussed with respect to their morphosyntactic characteristics.

3.1. The 'Juxtaposed' Type

As in Kavalan, Amis also uses a juxtaposition to indicate a comparison, where the comparee argument and the standard are located separately in two clauses. The gradable predicate occurs as the main predicate of one clause while the other clause is predicated either with the antonymous or negative form. This is shown in (13):

- (13) a. fangcal kaku, tati'ih kisu
 AF.good 1SG.NOM AF.bad 2SG.NOM
 b. fangcal kaku, ca'ay ka-fangcal kisu
 AF.good 1SG.NOM NEG KA-good 2SG.NOM
 'I am better than you.'

⁵ Without case-markers, subjects are required to precede objects.

The distribution of word order, verbal voices, nominal case marking in this type of comparatives do not differ from non-comparative constructions as discussed in Amis (Huang 1995; Wu 1995, 2000; Liu 1999; Liu 2003).⁶

3.2 Nominal Comparatives

In the nominal comparatives, a comparison is instantiated by means of an equational construction where the comparee NP always occurs sentence-initially as the nominal predicate of the proposition, followed by a headless relativized subject NP in which contains a gradable predicate and the standard NP. This is illustrated in (14).

- (14) [NP u kaka aku] [NP k-u
 NCM older.sibling 1SG.GEN NOM-NCM
 fangcal⁷ tisuwan/tisuwanan/atu kisu]
 good 2SG.ACC/2SG.LOC/CONJ 2SG.NOM
 ‘My brother is better than you.’
 (Lit: ‘The one who is better than you is my brother.’)

This type of comparatives allows only a fixed word order: comparee-[predicate-standard]. With respect to case-marking, the comparee NP in a sentence-initial non-argument position can only take a caseless nominal marker *u*.⁸ In (14), the standard argument is marked with an accusative case, a locative case, or even a nominative case with the presence of a conjunction *atu*.⁹

⁶ The word order of AF expressions in Amis is mainly VSO and VOS, while SVO is acceptable but less common. The word order for Non-AF expressions (NAF) is limited to VOS.

⁷ In many of Amis literature, it has been mentioned that a suffix *-ay* often appears in AF verbs or *ma-* type verbs when either these verbs function as a noun modifier (i) or occur in the relative clauses (ii).

- (i) a. *ma-su'su'-ay* (a) *tamdaw* (Wu 2003: 64)
 MA-fat-FAC LNK person
 b.* *ma-su'su'* (a) *tamdaw*
 MA-fat LNK person
 ‘fat person’
- (ii) *ma-patay tu k-u-ya mi-kalat-ay ci*
 MA-die ASP NOM-NCM-that AV-bite-FAC NCM
aki-an (a) wacu (Wu 2003: 64)
 Aki-DAT LNK dog
 ‘The dog that bit Aki is dead.’

However, the suffix *-ay* does not occur in relative clauses such as (14), whenever a comparative interpretation is implied. And though the suffix *-ay* has been treated traditionally as a nominalizer (e.g. Lin 1995; Wu 1995, 2000; Liu 1999), we agree with Wu (2003) and Liu (2003) in analyzing *-ay* as an epistemic or evidential marker.

⁸ In this paper we follow Liu (2003) in assuming that the morpheme *u* is a nominal marker, whose claim is different from a ‘neutral case’ analysis or a ‘common noun marker’ analysis (Liu 1999). We also agree with Wu (2003) in that the nominative case marker for personal proper nouns is a zero form (e.g. *ø-ci*).

⁹ There are two conjunctions *atu* and *a* in Amis, which are used typically in conjoining NP categories. The same morpheme *a* is also traditionally viewed as a ‘linker’ in the Formosan literature. As a ‘linker’, it optionally intervenes either between two verbal predicates (indicating a control relationship, a resultative/descriptive expression, a manner adverbial expression, etc.),

The comparison can appear with a more complex form of NP. Example (15) below shows how a relativized NP (with appearance of the head noun) fits well either in a comparee NP or a standard NP position.

- (15) u ni-‘aca-an aku a futing
 NCM PFV-buy-PF 1SG.GEN LNK fish
 k-u adihay t-u ni-‘aca-an isu a afar
 NOM-NCM many OBL-NCM PFV-buy-PF 2SG.GEN LNK shrimp
 ‘The fish I bought are more than the shrimp you bought.’

3.3 -ki- Comparatives

In this comparative construction, a verbal affix *-ki-* is incorporated with the gradable predicate, deriving a “transitive-like” verbal complex with the meaning of ‘exceed’ along the line of Stassen’s (1985) typological study.¹⁰ Voice/Focus marking is obligatory in this type of comparison. The construction thus can be further categorized into an AF and a PF with respect to the selection of the grammatical subject.

3.3.1 AF -ki- Comparatives

While the form of the focus/voice affix in non-comparative sentences may differ with respect to different types of verbal predicates (i.e. *mi-*, \emptyset , <um>, *ma-* for AF), only one AF marker *mi-* can be used for *-ki-* comparatives and the main predicate is represented as ‘*mi-ki-V*’. The case-marking pattern is identical to that of a non-comparative AF construction: that is, in AF *-ki-* comparatives the comparee NP gets the nominative case and the standard NP gets the locative case. Compare a comparative (16) with a non-comparative counterpart (17):

or between different kinds of modifiers (e.g. possessives, numeral, attributive adjectival predicates, etc.) and their head nouns. Readers are referred to Liu (2003) for her analysis of the second *a* also as a conjunction rather than a linker. In this paper, we simply follow the tradition in terming the morpheme *a* as a linker without taking a stance.

¹⁰ Three further subtypes of ‘exceed’ has been classified in Stassen (1985:43). The verbal predicate ‘exceed’ could be placed after the gradable predicate in a so-called ‘serial verb’ construction with either a finite form (i) or an infinite form (ii). The ‘exceed’ could also appear as the main predicate while degrading the gradable predicate to a nominal form (iii).

- (i) Exceed-1 comparatives: *VIETNAMESE*
 vang qui hon bac
 gold valuable exceed silver
 ‘Gold is worth more than silver.’
- (ii) Exceed-2 comparatives: *SWAHILI*
 mti huu ni merefu ku-shinda ule
 tree this is big INF-exceed that
 ‘This tree is taller than that tree.’
- (iii) Exceed-3 comparatives: *HAUSA*
 doki ya-fi rago girma
 horse it-exceed goat bigness
 ‘A horse is bigger than a goat.’

In Amis, *-ki-*, as an affixal verbal morpheme, does not fit well into any of Stassen’s ‘exceed’ subtypes. In this paper, we still consider *-ki-* as one ‘exceed’ type since the morphological complex *-ki-V* behaves exactly like a ‘transitive-like’ *exceed* as suggested by Stassen.

- (16) mi-ki-lalok \emptyset -ci panay ci-aki-an
 AF-exceed-diligent NOM-NCM PN NCM-PN-LOC
 ‘Panay is more diligent than Aki.’
- (17) mi-palu \emptyset -ci panay takuwanan (Wu 2000: 69)¹¹
 AF-hit NOM-NCM PN 1SG.LOC
 ‘Panay is going to hit me.’

The standard NP can also appear with a nominative case, as long as it co-occurs with a conjunction *a* or *atu*, as shown below.

- (18) mi-ki-lalok \emptyset -ci panay a/atu \emptyset -ci aki
 AF-exceed-diligent NOM-NCM PN CONJ NOM-NCM PN
 ‘Panay is more diligent than Aki.’

The word order is flexible in AF -ki- constructions as expected:

- (19) a. mi-ki-fangcal kaku tisuwanan
 AF-exceed-good 1SG.NOM 2SG.LOC
 ‘I am better than you.’
- b. mi-ki-fangcal tisuwanan kaku
 AF-exceed-good 2SG.LOC 1SG.NOM
 ‘I am better than you.’
- c. u kaka aku, mi-ki-fangcal tisuwanan
 NCM older.sibling 1SG.GEN AF-exceed-good 2SG.LOC
 ‘My brother is better than you.’

3.3.2 PF -ki- Comparatives

Among the four PF affixes (*ma-*, *-en*, *mi...an*, *maka-*), the focus/voice affix in -ki-comparatives is strictly restricted to ‘*ma-*’. Similar to the case-marking of a non-comparative PF construction, the comparee NP as the agent is assigned genitive case while the standard NP as the theme is assigned nominative case, as shown in (20).

- (20) ma-ki-lalok n-i panay \emptyset -ci aki
 PF-exceed-diligent GEN-NCM PN NOM-NCM PN
 ‘Panay is more diligent than Aki.’

Unlike a flexible word order attested in AF -ki- comparatives, PF -ki- type has a rather fixed order, the same observation found in a non-comparative PF construction which allows only a VOS order. Consider (21).

- (21) a. ma-ki-lalok n-i panay \emptyset -ci Aki
 PF-exceed-diligent GEN-NCM PN NOM-NCM PN
- b. *ma-ki-lalok \emptyset -ci aki ni panay
 PF-exceed-diligent NOM-NCM PN GEN-NCM PN

¹¹ The example has been slightly re-glossed according to our coding symbols.

- c. *n-i panay ma-ki-lalok ø-ci aki
 GEN-NCM PN PF-exceed-diligent NOM-NCM PN
 ‘Panay is more diligent than Aki.’

In addition, unlike nominal comparatives and AF -ki- comparatives where the standard NP may appear with the same case marking as the comparee NP by means of a conjunction, PF -ki- comparatives do not allow such an alternation. This is exemplified below:

- (22) a. *ma-ki-lalok n-i panay a n-i aki
 PF-exceed-diligent GEN-NCM PN CONJ GEN-NCM PN
 b. *ma-ki-lalok n-i panay a ø-ci aki
 PF-exceed-diligent GEN-NCM PN CONJ NOM-NCM PN
 ‘Panay is more diligent than Aki.’

3.4 *ikaka/isafa* Comparatives

Just like in Kavalan, Amis also employs a pair of lexical items *ikaka/isafa* ‘be.more/be.less’, derived from the kinship terms *kaka/safa* ‘older.sibling/younger.sibling’, to instantiate a comparison event.¹² Unlike *qaqa/suani* ‘older sibling/younger sibling’ in Kavalan, *kaka/safa* do not involve a morphological Ca-reduplication to denote a comparison. A locative prefix *i-*, instead, is obligatorily attached, deriving *ikaka/isafa* with a meaning similar to *more* in English. Compare (23) with (8a) in Kavalan:

- (23) *ikaka* ø-ci panay takuwan/takuwanan
 be.more.(in.age) NOM-NCM PN 1SG.ACC/1SG.LOC
 ‘Panay is older than me.’
 (repeated)
- (8) a. *qa-qaqa* ti-abas tu pataqsian ‘nay
 RED-older.sibling NCM-PN OBL student that
 ‘Abas is older than that student.’

And *ikaka/isafa* in Amis, compared to *qaqa/suani* in Kavalan, is relatively common in expressing all kinds of comparison in addition to a comparison of age as in (23), and its syntactic structure is rather different. Consider (24).

- (24) a. *ikaka* k-u fana’ n-i mama aku tisuwan/tisuwanan
 be.more NOM-NCM know GEN-NCM father 1SG.GEN 2SG.ACC/2SG.LOC
 ‘My father knows more than you do.’
 b. *isafa* k-u takaraw n-i panay takuwan/takuwanan
 be.less NOM-NCM tall GEN-NCM PN 1SG.ACC/1SG.LOC
 ‘Panay is less taller than me.’

¹² There exists a preference of using *ikaka* to *isafa* among different informants. Such a preference of usage is not uncommon: cross-linguistically the nature of gradable terms often follows a universal principle by which ‘more’ is unmarked and ‘less’ isn’t; ‘tall’ is unmarked and ‘short’ isn’t. Therefore, it is natural as well as intuitive to choose *ikaka* ‘be more’ as the default for a linguistic comparative device.

In (24), serving as the grammatical subject of the main predicate *ikaka/isafa*, the complex NP with a nominative case is composed of a nominalized (non-finite) gradable predicate and a genitive comparee NP.¹³ Thus (24a) should be literally interpreted as *My father's knowing (knowledge) is more than yours*, and (24b) should be *Panay's tallness is less than mine*.

With respect to the standard NP, accusative (*takuwan*) locative (*takuwanan*) cases are possible choices shown in (24). The option of accusative or locative pronouns results from an ellipsis of the standard NP *tu takaraw* 'my tallness', as exemplified in (25).

- (25)

	ikaka	k-u	takaraw	n-i	panay
	be.more	NOM-NCM	tall	GEN-NCM	PN
	t-u	takaraw	aku		
	OBL-NCM	tall	SG.GEN		

'Panay is taller than me.'

As for the word order, in addition to the VSO order given in (24a), other possible variations as in (26) are attested in this construction.

- (26) a.

ikaka	tisuwan/tisuwanan	[NP k-u	fana'	n-i	mama aku]
	be.more 2SG.ACC/2SG.LOC	NOM-NCM	know	GEN-NCM	father 1SG.GEN

b.

[NP k-u	fana'	n-i	mama aku]
NOM-NCM	know	GEN-NCM	father 1SG.GEN

ikaka kisuwan/kisuwanan
 be.more 2SG.ACC/2SG.LOC
 'My father knows more than you do.'

4. Concluding Remarks

Table 1 summarizes our discussion so far with respect to the morphosyntactic characteristics of each type of comparative attested in Kavalan and Amis. As the table shows, two types of comparative constructions are identified in Kavalan while four are found in Amis. None of the constructions examined in both languages is like the English-type comparative, which makes use of a degree modifier *more/-er* and a special comparative particle *than*. Instead, either parallel clauses marking a contrast, the most straightforward strategy, or special lexical items/morphemes such as *qaqa/suani*, *ikaka/isafa*, or *-ki-* are employed to indicate a comparison. It is also interesting to find that constructions such as 'addity' type in Kavalan or *-ki-* in Amis function syntactically similar to non-comparative transitive-like structures.

¹³ The finite form of *fana'* is *ma-fana'* as shown in (i):
 (i)

ma-fana'	kaku	tuni	a	demak	(Wu 2000: 72)
AF-know	I	this	LNK	matter	

 'I know this.'

Table 1: *Comparative constructions in Kavalan and Amis*

Kavalan					
		Components of a Comparison			Word order
		Gradable predicate	Comparee NP	Standard NP	
Juxtaposed type		As a main predicate	NOM	NOM	V1 NP1; V2 NP2 V1 NP1; not V1 NP2
Addity type		As a main predicate	NOM	OBL/LOC	AF: flexible order NAF: fixed order
Amis					
Juxtaposed type		As a main predicate	NOM	NOM	V1 NP1; V2 NP2 V1 NP1; not V1 NP2
Nominal type		Occur inside an equational construction	caseless	ACC/LOC/ NOM	NP NP
-ki- type	AF mi-	Intergrated with -ki-	NOM	ACC/LOC/NOM	flexible order
	PF ma-		GEN	NOM	fixed order
ikaka/isafa type		Infitive form occurring with a genitive comparee NP	GEN	ACC/LOC	flexible order

References

- Chang, Yung-li. 1997. *Voice, case and agreement in Seediq and Kavalan*. Taiwan: National Tsing Hua University dissertation.
- Chang, Yung-li. 2000. *A Reference grammar of Kavalan*. Taipei: Yuanliu Publishing Co. (in Chinese)
- Huang, Lillian M. 1995. The case markers and pronominal system in Amis. *The Journal of National Chengchi University* 70: 217-258. Taipei: National Chengchi University.
- Lee, Amy Pei-jung. 1997. *The case-marking and focus systems in Kavalan*. Taiwan: National Tsing Hua University MA thesis.
- Lehmann, Winfred P. 1972. Comparative constructions in Germanic of the OV type. Evelyn Scherabon Firchow et al. (eds.). *Studies for Einar Haugen. Janua linguarum*, series major 59. The Hague: Mouton. 323-330
- Li, Paul Jen-kuei. 1996. *The Formosan Tribes and Languages in I-Lan*. I-Lan: I-Lan County Government, Taiwan. (in Chinese)
- Li, Paul Jen-kuei. 1997. A syntactic typology of Formosan languages – case markers on nouns and pronouns. *Proceedings of the Fourth International Symposium on Chinese Languages and Linguistics (IsCLL-IV)*, ed. by Ho and Tseng, 270-289. Taipei: Academia Sinica.
- Lin, Michelle H. 1995. Two Amis suffixes: -ay and -an. *Studies in English Literature and Linguistics* 21: 159-173.

- Liu, Dorinda Tsai-hsiu. 1999. *Cleft constructions in Amis*. Taiwan: National Taiwan University MA thesis.
- Liu, Emma En-hsing. 2003. *Conjunction and modification in Amis*. Taiwan: National Tsing Hua University MA thesis.
- Stassen, L. 1985. *Comparison and Universal Grammar*. Oxford: Blackwell.
- Wu, Joy. 1995. *Complex sentences in Amis*. Taiwan: National Taiwan Normal University MA thesis.
- Wu, Joy. 2000. *A Reference grammar of Amis*. Taipei: Yuanliu Publishing Co. (in Chinese)
- Wu, Joy. 2003. Clausal modifiers in Amis. *Concentric: Studies in English Literature and Linguistics* 29.2: 59-81.

GRAMMATICALIZATION PATHS OF THE THAI VERB *dây*: A CORPUS-BASED STUDY

Kiyoko Takahashi

Kanda University of International Studies

<kiyoko@kanda.kuis.ac.jp>

1. The semantics of *dây*

The Thai verb *dây* is usually translated into an English agentive verb such as ‘get’, ‘obtain’, ‘gain’, ‘attain’, and so on. For instance, (1) below is usually translated as ‘(he) got money’, as in (1a). But I do not consider *dây* as an agentive verb, since it cannot co-occur with a manner verb indicating an agent’s volition such as *phayaayaam* ‘make an effort’, as exemplified in (2). It follows that the topic person of (1), which may or may not be overtly expressed, is not an agent (i.e. a conscious, willful and responsible actor with the ability to control the course of the event) but rather an experiencer (i.e. an undergoer of the event). On this basis, I regard *dây* as a non-agentive, non-volitional achievement verb describing a momentaneous event of ‘emergence’, as interpreted in (1c). It is worth noticing that an emergence verb in Thai, such as *kàət* ‘take place’ and *praako t* ‘appear’, is normally followed by a noun phrase naming an emerging entity, as in (3). Like *dây*, *kàət* ‘take place’ is incompatible with *phayaayaam* ‘make an effort’, as in (4).

- (1) (*kháw*) *dây* *ηən*
(PRONOUN) DAY money
a. (He) got money.
b. (He) came to have money.
c. Money emerged (for him, and he got it).
- (2) * (*kháw*) *phayaayaam* *dây* *ηən*
(PRONOUN) make an effort DAY money
(He) tried to get money. (intended meaning)
- (3) *kàət* *panhǎa*
occur problem
A problem took place.
- (4) * (*kháw*) *phayaayaam* *kàət* *panhǎa*
(PRONOUN) make an effort occur problem
(He) tried to bring about a problem. (intended meaning)

Since I do not take it for granted that the semantics of *dây* entails the presence of a human being, I hesitate to interpret *dây* as ‘come to have’ (Enfield 2003), as in (1b). The lexical meaning of *dây* does not encompass the state of possession of something by a particular person. The sense of ‘possession’ is an implication deriving from the given pragmatic context. That is to say, when the topic person had wanted an entity denoted by the post-

verbal nominal, we naturally infer that after the entity came into existence at the locus of the person, he would take it and become its possessor. In the case of (1), if the topic person wanted money, we would readily understand that he got the money as a result of its emergence. Thus, I would rather prefer to translate (1) as ‘money emerged’. This is, I think, the default verbal meaning of *dây* (Takahashi & Methapisit 2004).

In addition to the verbal meaning, *dây* has two main functional meanings which in this study I will call ‘realization’ and ‘possibility’ for the sake of convenience. Examples of expressions including *dây* as a functional morpheme are given in (5) and (6).

(5) (*kháw*) *dây* *pay*
 (PRONOUN) DAY go
 The event of (his) going is realized.

(6) (*kháw*) *pay* *dây*
 (PRONOUN) go DAY
 The event of (his) going is possible.

The pre-verbal *dây*, as in (5), is the marker for ‘realization’ (Takahashi & Methapisit 2004) or ‘participant-external actuality’ (van der Auwera & Plungian 1998). The linguistic concept ‘realization’ is a hybrid comprising an inchoative aspect and a realis-assertion. It specifies the speaker’s evaluation or understanding that a situation in question has actually occurred as a result of some prior non-specific event (Enfield 2003). The basis of a realis-assertion characterized by ‘realization’ is the speaker’s belief (or subjective certainty) that the prior event which is backgrounded but can be pragmatically retrieved from the given context should be connected with the realization of the situation in question. Take (5) for example. The realized event of the topic person’s going to a certain place may result from such a prior event that the person was invited to visit the place he had wanted or expected to go to.

On the other hand, the post-verbal *dây*, as in (6), functions as the marker for ‘possibility’. The linguistic concept ‘possibility’ subsumes a variety of subcategories such as circumstantial possibility, non-human capacity, agentive possibility (ability), probability, and permissibility (Takahashi & Methapisit 2004). The post-verbal *dây* basically denotes the most inclusive sense of possibility, which is differently called, e.g., ‘possibility in a world independent of the speaker’ (Traugott 1989), ‘externally conditioned possibility’ (Shibuya 1993), ‘participant-external possibility’ (van der Auwera & Plungian 1998), and ‘circumstantial possibility’ (Narrog 2005). Rather specific interpretations of the possibility meaning of *dây*, such as ability, probability, and the like, are achieved through our inferences in a particular pragmatic context.

Previous studies on the semantics of *dây* in Thai and Lao (Bisang 1996, Diller 2001, Enfield 2003, Matisoff 1991, Meesat 1997, Sindhvananda 1970, inter alia) assume that the original, core meaning of *dây* is ‘be able’ or ‘get’ or ‘come to have’, all of which presuppose the presence of a human being as an agent (actor) or an experiencer (undergoer). However, few, if any, studies have seriously tried to provide historical evidence for this assumption. This study, therefore, aims at offering a hypothesis on *dây*’s original meaning and grammaticalization paths based on empirical research dealing with

historical corpus data. Thus, this is a case study of grammaticalization using a data-driven approach.

2. Hypotheses on the mechanisms of grammaticalization

In my previous studies (Takahashi 2005, 2006a, 2006b; Takahashi & Shinzato 2005), I analyzed actual discourses in Thai inscriptions from the 13th century through the present time and reconstructed most plausible grammaticalization paths of *dây* into the two functional morphemes. In the study focusing on the ‘realization’ marker (Takahashi 2006b), I found that in the process that *dây* has developed into the ‘realization’ marker, (a) ‘changes into a less specific meaning’ (cf. ‘semantic generalization’ Bybee et al. 1994; ‘depletion’ Givón 1975; ‘desemanticization’ Heine & Kuteva 2002; ‘schematization’ Langacker 1991; ‘semantic bleaching’ Sweetser 1988) and (b) ‘changes into a more specific meaning’ (cf. ‘specification’ Kuteva 1999; ‘pragmatic strengthening’ Traugott 1988; ‘subjectification’ Traugott 1989, 1995) are both involved. (7a) and (7b) below describe the two directions of semantic change which I attested to be involved in *dây*’s evolution into the ‘realization’ marker.

- (7) a. Changes into a less specific meaning or ‘generalization’: Generalization of the referential, content meaning of *dây* and its argument nominal (i.e. bleaching of the meaning of the described emergence event with an emerging entity)
- b. Changes into a more specific meaning or ‘specification’: Specification of the constructional, ‘frame’-like meaning of *dây* constructions and a certain modal meaning associated (i.e. strengthening of the meaning of the speaker’s subjective construal)

Furthermore, I hypothesized that these two types of semantic change, namely generalization and specification, interact differently at different stages of the development of grammaticalization, as explicated in (8).

- (8) a. In early stages of the development of grammaticalization, a certain specification always precedes a certain generalization. In other words, the former triggers the latter. For example, forming a specific construction triggers bleaching of the meaning of an argument nominal.
- b. In late stages, on the other hand, both of the two types of semantic change occur at the same time, as if they were the two sides of the same coin. For example, bleaching of the meaning of *dây* and fixing of the pre-verbal *dây* construction occur simultaneously.

The main purpose of this study is to show that my hypotheses on the mechanisms of grammaticalization based on the examination of *dây*’s grammaticalization into the ‘realization’ marker, which are summarized in (7) and (8) above, are indeed applicable to *dây*’s grammaticalization into the ‘possibility’ marker.

3. Grammaticalization into the ‘realization’ marker

First, I will review *dây*’s grammaticalization into the ‘realization’ marker. Using the diachronic corpus data, I reconstructed a grammaticalization path along which *dây* has

changed from a verb for ‘emergence’ to the marker for ‘realization’, as shown in Table 1 below.

Table 1: Grammaticalization path from ‘emergence’ to ‘realization’

CONSTRUCTION	MEANING
1. [<i>dây</i> quantity-NP]	A quantity emerges. (9)
2. [VP] [<i>dây</i> quantity-NP]	A quantity emerges as a result of a prior situation. (10)
3. [VP] [<i>dây</i> NP]	Something emerges as a result of a prior situation. (11)
4. [VP] [<i>dây</i> NP DATIVE human-NP]	Something emerges for the person as a result of a prior situation. (12)
5. topic-NP [OPTATIVE CAUSATIVE <i>dây</i> DATIVE human-NP]	Hope this thing will (will not) come to emerge for the person. (13)
6. preceding discourse [OPTATIVE <i>dây</i> (as I wish)]	May this story occur as I wish! (14)
7. [OPTATIVE <i>dây</i> VP]	Hope a situation will (will not) occur. (15)
8. [<i>dây</i> VP]	A situation is realized. (16)

All the eight constructions listed in Table 1 were used in the Sukhothai period (the 13th century through the 15th century). I tried to see which constructions disappeared in earlier ages, or on the contrary, which constructions got to be frequently used in later ages. I have found that by the end of the 18th century, Constructions 1 to 5 became less used, and they, except for 2 and 3 which have diverged into other *dây* constructions, disappeared before the 20th century. Constructions 6 and 7 were used even into the 20th century for some time, but now are no longer seen. In contrast, Construction 8 became very common in the period of the present dynasty (since 1782). These observations lead to the following hypothesis.

Originally *dây* was an achievement verb designating an emergence of a quantity like the number and the period of something. I believe so owing to the fact that it was frequently followed by a noun phrase expressing a certain quantity ([*dây* quantity-NP]). This is Construction 1 meaning that ‘a quantity emerges’, as in (9).

- (9) *sàkkaràat* *dây* *cèt rójy sām sīp pèet*
 era DAY 738
 As for the era, 738 years emerged. (The period of the era amounted to 738 years.) (1376)

Used as the second verb phrase, then, Construction 1 changed into Construction 2 ([VP] [*dây* quantity-NP]) meaning that ‘a quantity emerges as a result of a prior situation’, as in (10).

- (10) *phim rûup phrá?* *dûay hìak dûay din*
 mold Buddhist image with tin with clay
dây mùun phan rójy pèet ?an
 DAY 11,108 CLASSIFIER
 (They) molded Buddhist images with tin and clay and the number of the images amounted to 11,108. (1339)

Shifting to Construction 3 ([VP] [*dây* NP]), not only a quantity noun phrase but also a noun phrase in general became able to take place after *dây*. Construction 3 denotes that ‘something emerges as a result of a prior situation’, as in (11).

- (11) *kuu pay thỏ bán thỏ mua*
 PRONOUN go attack country attack country
dây chấn *dây nua* *dây pà*
 DAY elephant DAY elephant DAY man
dây naa *dây nua* *dây thỏ*
 DAY woman DAY silver DAY gold
 I attacked countries, and elephants, people and treasures emerged (and I got them).
 (1292)

The important point is that *dây* in this bipartite construction does not simply express an emergence event but rather encodes a change and the resultant state.

Construction 4 ([VP] [*dây* NP DATIVE human-NP]) came from the addition of a dative prepositional phrase to the end of Construction 3. The dative prepositional phrase refers to an experiencer who is more or less affected by an emerging entity. Construction 4 represents that ‘something emerges for the person as a result of a prior situation’, as in (12).

- (12) *phî kuu taay cù*
 elder brother PRONOUN die CONJUNCTION
dây mua kèe kuu thánklom
 DAY kingdom DATIVE PRONOUN whole
 My elder brother died and the whole kingdom emerged for me (and I got it). (1292)

This result-oriented meaning of Construction 4 appears to indirectly trigger an interpretation of a *dây* construction with a topic person, like (1) above, that ‘the topic person eventually gets a thing emerged’.

Construction 5 (topic-NP [OPTATIVE CAUSATIVE *dây* DATIVE human-NP]) arose when the emerging entity became topicalized and *dây* with the dative prepositional phrase (*dây kèe* NP) became preceded by the combination of the optative marker (the marker for wishing) (e.g. *cù*, *co*, *yàa*) and the causative marker (i.e. *hây*, *hũu*). Construction 5 characteristically encodes the writer’s wishing that ‘the topic entity will (or will not) come to appear for a particular person’.

- (13) a. *bàp ... cù hây dây*
 sin OPTATIVE CAUSATIVE DAY
kèe phũu bở sũu
 DATIVE person NEGATIVE honest
 As for that sin, (I) hope it will come to emerge for the non-honest person. (1392)
- b. *thỏt ?an sảabaan ní*
 punishment RELATIVE PRONOUN swear this
yàa hây dây
 NEGATIVE OPTATIVE CAUSATIVE DAY

strengthening of the language users' inferences is behind the establishment of Construction 8.

Note that Construction 8 has been used since the Sukhothai period, though the frequency of its use in early ages was much less than that in the present time. The radical increase of the use of the pre-verbal *dây* in the present dynasty, however, shows that it must be in recent times that *dây* in Construction 8 came to be fully recognized as the 'realization' marker.

Table 2 below summarizes the semantic changes involved in the grammaticalization into the 'realization' marker. Instances of (a) 'generalization of the semantics of the verb *dây* and its argument nominal' are listed in the left column, and instances of specification, including (b) 'specification of the semantics of constructions (entrenchment of specific event structures)' and (c) 'specification of modal meanings (entrenchment of a specific mood as well as subjectification)', are listed in the right column.

Table 2: *Semantic changes involved in grammaticalization into the 'realization' marker*

1→2		(b) becoming used as the second verb phrase
2→3	(a) semantic generalization of the emerging entity (extension from quantity to thing)	
3→4		(b) co-occurrence with a dative prepositional phrase indicating the locus of emergence
4→5		(b) co-occurrence with the optative and the causative (b) topicalization of the noun phrase naming an emerging thing (c) entrenchment of the optative mood
5→6	(a) semantic generalization of the emerging entity (extension from thing to event)	(b) establishment of formulaic expression for wishing for the emergence of a desirable situation (disappearance of the causative and the topic noun phrase)
6→7	(a) semantic generalization of the emerging entity (extension from event to propositional concept)	(b) taking a verbal complement
7→8	(a) bleaching of the emergence sense (extension from emergence to realization)	(b) establishment of the construction for 'realization' through syntactic reanalysis rendering <i>dây</i> a functional morpheme (disappearance of the optative) (c) subjectification (becoming the 'realization' marker)

From Table 2 we can see the following two points, which I regard as empirical evidence in support of my hypotheses stated in (7) and (8) above. First, in early stages of the development of grammaticalization, a specification precedes a generalization: (i) In the shift from Stage 1 to Stage 2, *dây* followed by a quantity noun phrase ([*dây* quantity-NP]) came to be commonly used as the second verb phrase to express a change and the resultant state. With this specification, the meaning of the emerging entity extended from quantity to thing in general, which gave rise to Stage 3. (ii) In the shift from Stage 4 to Stage 5, the optative mood wishing for a certain emergence (or non-emergence) became entrenched. With this specification, the meaning of the emerging entity extended from thing to event, which gave

rise to Stage 6. And, (iii) in the shift from Stage 5 to Stage 6, the causative and the topic noun phrase disappeared and formed a formulaic expression for wishing for the emergence of a desirable event. With this specification, the meaning of the emerging entity extended from event to propositional concept, which gave rise to Stage 7.

Second, in late stages of the development of grammaticalization, a generalization and a specification occur at the same time: (i) The shift from Stage 5 to Stage 6 involved an abstraction of the emerging entity (from thing to event), and concurrently, the formulaic expression for wishing for a desirable event became fixed. (ii) The shift from Stage 6 to Stage 7 involved a further abstraction of the emerging entity (from event to propositional concept), and concurrently, *dây* came to take a verbal complement. And, (iii) the shift from Stage 7 to Stage 8 involved a bleaching of the verbal meaning of *dây* (from emergence to realization), and concurrently, the construction underwent ‘syntactic reanalysis’ (Langacker 1977) rendering *dây* the ‘realization’ marker that is always followed by a verb phrase.

4. Grammaticalization into the ‘possibility’ marker

We now turn to an examination of *dây*’s grammaticalization path into the marker for ‘possibility’ in order to further verify my hypotheses. Table 3 below shows my corpus-based reconstruction of *dây*’s grammaticalization into the ‘possibility’ marker.

Table 3: Grammaticalization path from ‘emergence’ to ‘possibility’

CONSTRUCTION	MEANING
1. [<i>dây</i> quantity-NP]	A quantity emerges. (9)
2. [seek for (NP)] (CONJ.) [NEGATIVE <i>dây</i> quantity-NP]	After seeking for something, a certain quantity of it does not emerge. (17)
3. [seek for (NP)] (CONJ.) [NEGATIVE <i>dây</i>]	After seeking for something, it does not emerge. (18)
4. [VP] (CONJ.) [NEGATIVE <i>dây</i>]	After doing something, it is not achieved. (19)
5. [VP NEGATIVE <i>dây</i>]	It is not possible to do/be something. (20)
6. [VP <i>dây</i>]	It is possible to do/be something. (21)

Like the constructions listed in Table 1, the constructions listed in Table 3 were all seen in the Sukhothai inscriptions. I tried to identify which ones got to be less and less used while which ones got to be more and more used in later ages. I have found that Constructions 2 and 3, which have diverged into other *dây* constructions, disappeared by the end of the 14th century. Constructions 4 to 6 remain until now. Especially in the present dynasty (since 1782), Construction 6, which does not have the negative, is quite frequently used. These observations lead to the following hypothesis.

Shifting from Construction 1 ([*dây* quantity-NP]) to Construction 2 ([seek for (NP)] (CONJ.) [NEGATIVE *dây* quantity-NP]), the *dây* verb phrase came to be negated and follow another verb phrase including the verb *hăa* ‘seek for’. Construction 2 means that ‘a certain quantity of a thing does not emerge after seeking for the thing’, as in (17).

- (17) *fũuŋ khon* *?an*
 people RELATIVE PRONOUN
càk *rúu bun tham*
 MODAL know virtue right principles

The semantic changes involved in grammaticalization into the ‘possibility’ marker are summarized in Table 4 below.

Table 4: *Semantic changes involved in grammaticalization into the ‘possibility’ marker*

1→2		(b) becoming used as the latter verb phrase following the former verb phrase meaning ‘to seek for a thing’
2→3	(a) semantic generalization of the emerging entity (extension from non-emergence of quantity to non-emergence of thing)	(b) disappearance of the numeral after <i>dây</i>
3→4	(a) semantic generalization of the former verb phrase (extension from ‘seeking for a thing’ to ‘doing something’) (a) semantic generalization of the emerging entity (extension from non-emergence of thing to non-emergence of event)	(b) entrenchment of the construction ‘[VP] (CONJ.) [NEGATIVE <i>dây</i>]’ (implying ‘non-achievement’ sense)
4→5	(a) semantic generalization of the emerging entity (extension from non-emergence of event to non-emergence of propositional concept)	(b) complete disappearance of the conjunction (undergoing syntactic reanalysis which results in an interpretation of the construction as a single clause) (c) subjectification (implying ‘impossibility’ sense)
5→6	(a) bleaching of the emergence sense (extension from non-emergence of propositional concept, i.e. impossibility, to possibility in general)	(c) subjectification (becoming the marker for ‘possibility’)

From Table 4 we can see the following, which can be considered as another piece of evidence to support my hypotheses regarding the mechanisms of grammaticalization. First, a specification precedes a generalization in early stages of the development of grammaticalization: (i) In the shift from Stage 1 to Stage 2, the negated *dây* verb phrase ([NEGATIVE *dây* quantity-NP]) came to follow the *hãa* verb phrase ([seek for (NP)]) to encode a change and the resultant state. With this specification, a particular quantity became unmentioned and the negated *dây* alone remained in the second verb phrase, which gave rise to Stage 3. (ii) In the shift from Stage 2 to Stage 3, the quantity noun phrase after *dây* disappeared and the form [NEGATIVE *dây*] was fixed. With this specification, the meaning of the first verb phrase extended from ‘seeking for a thing’ to ‘doing something’ and the meaning of the emerging entity also extended from thing to event, which gave rise to Stage 4. (iii) In the shift from Stage 3 to Stage 4, the negated *dây* verb phrase preceded by another verb phrase ([VP] (CONJ.) [NEGATIVE *dây*]) came to be conventionally used to denote the sense of ‘non-achievement’ of a particular event (i.e. after doing something, it is not achieved). With this specification, the meaning of the construction as a whole extended from non-emergence of event (‘non-achievement’) to non-emergence of propositional concept (‘impossibility’), which gave rise to Stage 5. And, (iv) in the shift from Stage 4 to Stage 5, the conjunction *lee* ‘and’ totally disappeared and the construction underwent syntactic reanalysis resulting in an interpretation of the construction as a single clause. With this specification, the meaning of the construction as a whole extended from non-emergence of propositional concept

(‘impossibility’) to ‘possibility’ in general, which gave rise to Stage 6. Second, a generalization and a specification occur at the same time in late stages of the development of grammaticalization: (i) The shift from Stage 3 to Stage 4 involved an abstraction of the emerging entity (from thing to event), and concurrently, the form [NEGATIVE *dây*] came to imply ‘non-achievement’. (ii) The shift from Stage 4 to Stage 5 involved a further abstraction of the emerging entity (from event to propositional concept), and concurrently, the form [NEGATIVE *dây*] came to imply ‘impossibility’. And, (iii) the shift from Stage 5 to Stage 6 involved a bleaching of the verbal meaning of *dây* (from propositional concept to modal concept), and concurrently, *dây* became the ‘possibility’ marker.

5. Conclusion

In concluding, I would like to claim the following three points.

First, in order to concretely understand the mechanisms of grammaticalization as a whole, it is necessary to pay attention to the two different directions of semantic change, namely generalization and specification. It is not enough for the purpose only to consider exactly what kinds of inference such as metaphor and metonymy would motivate a semantic change from a lexical meaning to a functional meaning.

Second, having done empirical research on grammaticalization, I claim that we have to investigate grammaticalization paths by analyzing corpus data of actual discourses in the past; otherwise, we will fail to see how the two types of semantic change interact at each particular stage of the evolution of grammaticalization. That is, in early stages a specification precedes a generalization, while in late stages both occur simultaneously.

Third, Diewald (2002: 117) is right in saying that “the decisive factors for the triggering and continuation of a grammaticalization process are not to be found exclusively in the grammaticalizing items themselves, but also in changes in related linguistic categories and subsystems” and that “the split between the older, more lexical meaning and the newly grammaticalizing meaning [...] is reinforced not only by changes concerning the new meaning and function, but also by the further development of the older, lexical reading”. I concur with his view, and therefore I will continue investigating changes in related linguistic categories and subsystems such as historical changes of the negative and the causative systems in Thai and also splits into other *dây* constructions (e.g. human-NP [*dây* NP], [VP] [*dây* evaluation-VP], [VP *hây* *dây*], [*hăa* NP/VP *mây*], etc.), so that we can have a better understanding of grammaticalization paths of *dây*.

References

- Bisang, Walter. 1996. Areal typology and grammaticalization: Process of grammaticalization based on nouns and verbs in East and mainland South East Asian languages. *Studies in Language* 20:3.519-597.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. *New reflections on grammaticalization*, ed. by I. Wischer and G. Diewald, 103-120. Amsterdam: John Benjamins.

- Diller, Anthony. 2001. Grammaticalization and Tai syntactic change. *Essays in Tai linguistics*, ed. by M.R. K. Tingsabadh and A. S. Abramson, 139-175. Bangkok: Chulalongkorn University Press.
- Enfield, N. J. 2003. *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia*. London: Routledge Curzon.
- Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. *Word order and word order change*, ed. by C. Li, 47-112. Austin: University of Texas Press.
- Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Kuteva, Tania A. 1999. Specification in grammar. *Cultural, psychological and typological issues in cognitive linguistics*, ed. by M. K. Hiraga, C. Sinha and S. Wilcom, 269-284. Amsterdam: John Benjamins.
- Langacker, Ronald W. 1977. Syntactic reanalysis. *Mechanisms of syntactic change*, ed. by C. Li, 57-139. Austin: University of Texas Press.
- Langacker, Ronald W. 1991. *Foundation of cognitive grammar Vol.2: Descriptive application*. Stanford: Stanford University Press.
- Matisoff, James A. 1991. Areal and universal dimensions of grammaticalization in Lahu. *Approaches to grammaticalization*, ed. by E. C. Traugott and B. Heine, 383-453. Amsterdam: John Benjamins.
- Meesat, Paitaya. 1997. *A study of auxiliary verbs developed from verbs in Thai*. Master's thesis, Chulalongkorn University.
- Narrog, Heiko. 2005. Modality, mood and change of modal meanings: A new perspective. *Cognitive Linguistics* 16:4.677-731.
- Shibuya, Kastumi. 1993. Aspects of Japanese potential expressions and their historical development. *Memoirs of the faculty of letters*, Osaka University 33:1.
- Sindhvananda, Kanchana. 1970. *The verb in modern Thai*. Ph.D. dissertation, Georgetown University.
- Sweetser, Eve E. 1988. Grammaticalization and semantic bleaching. *BLS*14.389-405.
- Takahashi, Kiyoko. 2005. Uses of *dây* in Thai inscriptions. *Memoirs of Kanda University of International Studies* 17.295-353.
- Takahashi, Kiyoko. 2006a. Uses of *dây* in Thai inscriptions (2). *Memoirs of Kanda University of International Studies* 18.427-448.
- Takahashi, Kiyoko. 2006b. From 'emergence' to 'realization': A case study of grammaticalization involving both semantic generalization and specification. *Proceedings of the 6th annual meeting of the Japanese cognitive linguistics association*.193-203.
- Takahashi, Kiyoko and Tasanee Methapisit. 2004. Observations on form and meaning of DAY. *Papers from the 11th annual meeting of the Southeast Asian linguistics society 2001*, ed. by S. Burusphat, 701-719.
- Takahashi, Kiyoko and Rumiko Shinzato. 2005. On grammaticalization of Japanese and Thai emergence verbs. *Proceedings of the 5th annual meeting of the Japanese cognitive linguistics association*.197-206.
- Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. *BLS*14.406-416.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65:1.31-55.

- Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalization. *Subjectivity and subjectivisation: Linguistic perspectives*, ed. by D. Stein and S. Wright, 31-54. Cambridge: Cambridge University Press.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.
- van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2. 79-124.

A COMPARISON OF THE EXPRESSION OF SIMULTANEITY IN THAI AND ENGLISH ADULTS AND CHILDREN USING SHORT ANIMATIONS

Heather Winskel
University of Western Sydney
<h.winskel@uws.edu.au>

Sudaporn Luksaneeyanawin
Chulalongkorn University
<Sudaporn.L@chula.ac.th>

Aims of the study

The aim of the present study was to investigate the linguistic devices that Thai and English children aged 4 years, 5 years, 6 years, 7 years and adults use to express simultaneous or overlapping temporal relations depicted in short animations. Thai has imperfective aspectual morphemes but they are not morphologically grammaticalised on the verb and usage in general is optional and not obligatory. These particular aspectual characteristics of Thai form an interesting comparison with English, which has obligatory grammaticalised aspectual marking on the verb. An additional aim is to investigate if these grammatical differences and degrees of obligatoriness of aspectual marking in English and Thai affect what particular aspects of the external simultaneous events depicted in the animations are expressed, and if so, do children and adults show the same language-specific patterns?

Background literature

Recently, there has been a resurgence of interest in linguistic relativity, the influence that language has on cognition, and if language can influence the way that we think or perceive the world (e.g. Bowerman & Levinson 2001; Gentner and Goldin-Meadow 2003; Gumperz and Levinson 1996; Levinson 2003). A more subtle, less deterministic version of linguistic relativity has been formulated by Slobin, which he has termed *thinking-for-speaking*. This is a type of thinking “that is carried out on-line in the process of speaking” as the grammatical categories of the language that one speaks shapes or filters the way that aspects of the world are expressed (Slobin 1996: 75). So it is reasoned that speakers from languages that *do* have these obligatory, grammatical categories and speakers from languages that *do not* have these categories have different mental space for the particular semantic domain being encoded.

In the temporal domain two distinct language typologies have been identified, based on whether languages have obligatory grammatical marking of the durative-non-durative aspectual contrast or not (Berman & Slobin 1994). In languages, such as Spanish, Turkish and English which have obligatory, grammaticalised aspectual marking, speakers tend to express the overlapping temporal relations depicted in the *frog story* picture book using the aspectual forms in the language, whereas in German and Hebrew, which do not have obligatory grammaticalised aspect, speakers tend not to express the temporal distinctions depicted, even though there are alternative lexical means of expression (Berman & Slobin, 1994), e.g. *Der Hund rennt rennt rennt*.

The dog runs runs runs
The dog is running.

These characteristic patterns are also evident in the youngest children studied, the 3-year-olds. Furthermore, German and Hebrew narrators do not tend to “compensate” by using temporal connectives, whereas Turkish and English speakers use temporal connectives as well as aspectual morphemes to mark the temporal distinctions depicted, and the acquisition of temporal connectives is relatively early in comparison to German and Hebrew. Mandarin Chinese similar to Thai, has non-obligatory aspectual morphemes, and a study by Hickmann (2003) has revealed that speakers compensate for the optional nature of aspectual morphemes by using temporal connectives (temporal conjunctions and adverbials) to mark situational overlaps in discourse to a greater extent than in the other languages studied - French, English and German. Hence, it appears that obligatory, grammaticalised categories within a language have a channeling effect on the attention of the speaker towards particular aspects of temporal events or actions in the external world (Strömquist & Verhoeven 2004).

The main linguistic devices used to express simultaneity in Thai and English

Simultaneity or the overlap of two events or actions has the general shared meaning of “during the time that”, i.e. *event x overlaps with event y* (Bennett & Partee, 1972). However, the actual expression of simultaneity is more complex than conveyed by that definition as the degree of overlap between events can vary, and simultaneity is not often expressed unequivocally through explicit linguistic devices, but is inferred from a range of explicit and implicit linguistic devices as well as contextual cues.

The main linguistic devices, which can be used to express simultaneity in English and Thai are listed in Table 1. Imperfective aspect plays an important role in the expression of simultaneity or the overlap of events or actions, as it plays an important function in backgrounding events and gives an unbounded or durative meaning to the event (Chan 1996; Hopper 1979). In English, imperfective aspect is grammaticalised on the verb and is realized by the morpheme *-ing* (see Table 1). In Thai there are two imperfective markers *kamOlaŋ0¹* and *ju:1*. The preverbal imperfective marker *kamOlaŋ0* signifies ‘the process of doing something’, and postverbal *ju:1* signifies ‘the continuation of an event’ and translates as ‘stay/be alive’ (Burusphat, 1991). While *kamOlaŋ0* is more limited to the progressive meaning and restricts its usage to dynamic verbs, *ju:1* has evolved into a more general imperfective aspect marker which includes stative meaning (Meepoe, 1998), hence the two imperfective markers have distinctive though overlapping, aspectual meanings.

In addition, imperfective aspect can also be implied in both languages by lexical means such as *jaŋ0* ‘still’, reduplication of the adverb, e.g. *ruaŋ2 ruaŋ2* ‘continually’, *khɔŋ2 khɔŋ2* ‘gradually’, which can all give an ongoing interpretation to the clause. As Thai has imperfective aspect markers that are not obligatory similar to Mandarin Chinese,

¹ Tones are marked in the Thai examples cited in this paper as follows; 0=mid, 1=low, 2=falling, 3=high, 4=rising. This system is based on the system that was developed at the Linguistics Research Unit (LRU) of Chulalongkorn University (Luksaneeyanawin, 1993). IPA transcription is used for the transcription of all other Thai text.

then situation aspect, the intrinsic aspect of the verb and its predicate, can play an important additional role in signalling simultaneity (Smith & Erbaugh, 2005).

Additional principal means of encoding simultaneous temporal relations are through the use of simultaneous connectives, e.g. *when, while*, which explicitly relate two events or actions either in the foreground, background or across background and foreground. The interpretation of *when* is somewhat ambiguous and open to interpretation, as it can give either a simultaneous or successive interpretation dependent on the aspect of the verb and its predicate in conjunction with world knowledge (Moens & Steedman, 1986). Even though the interpretation of devices which have a ‘when’ type meaning are somewhat ambiguous, they are classified as simultaneous connectives in the current study, as events they relate temporally either overlap or at least are closely aligned. A common strategy in young children is to juxtapose two independent clauses without a temporal connective, e.g. *We eat. We go play*. Children later add a temporal connective to explicitly relate the two clauses e.g. *While I walk to school, I eat my breakfast*. The multifunctional, coordinating conjunction *and* in English or *lɛ3* ‘and’ in Thai is somewhat ambiguous as it can theoretically express a simultaneous or successive relationship between its two clauses.

On the basis that Thai does not have obligatory aspectual marking similar in this respect to Hebrew and German, we can predict that durative aspects of the events and the explicit simultaneous relationship between the two events depicted, will be expressed to a lesser degree in Thai than English speakers, and there will be a relatively late acquisition of simultaneous temporal connectives in Thai in comparison with English. Alternatively, on the basis that Thai has grammatical aspect as does English, Turkish and Spanish, we can predict that durative aspects of the events and the explicit simultaneous relationship between the two events depicted will be expressed in Thai to a similar degree as in English, and there will be a corresponding relatively early acquisition of temporal connectives.

Table 1: *The main linguistic devices used to express simultaneity in Thai and English.*

	English	Thai
Imperfective aspect	Grammaticised on the verb -ing The child is sleeping Obligatory	Not grammaticised on the verb – separate morphemes kam0laŋ0 ju:1 dek1 kam0laŋ0 nɔ:n0 ju:1
Lexical means	e.g. still, continues, keeps	Not obligatory, optional in usage <i>e.g. jaŋ0</i> ‘still’, <i>ruaŋ2 ruaŋ2</i> , <i>tɔ:1</i> <i>paj0</i> ‘continually’
Simultaneous connectives	e.g. while, when, as, during, meanwhile, whilst, at the same time.	e.g. <i>kha1na1 thi:2</i> ‘while’ <i>phɔ:0</i> , <i>mua2</i> ‘when’ <i>ra1wa:ŋ0</i> ‘during’ <i>phrɔm3kap1</i> ‘concurrently/at the same time’
Coordinating conjunction “and”	and	<i>lɛ3</i> ‘and’

Method

Participants: The participants consisted of 10 participants per age group; 4-year-olds, 5-year-olds, 6-year-olds, 7-year-olds and adults for Thai and English. Monolingual children were recruited from preschools and schools in Bangkok and Sydney. Participants were from predominantly middle SES. Adults were mainly university students studying at the University of Western Sydney or Chulalongkorn University.

Procedure: 18 short animations depicting different degrees of overlap between two actions were presented to participants individually on a laptop. Instructions given were: “Can you tell me what happens here/in this picture?” In the current study only the co-occurring simultaneous animation was selected for analysis (3 trials per participant and 30 trials per age group) (see Figure 1). The co-occurring animations consisted of ‘a person painting while a monkey was drawing on a piece of paper’, or ‘a girl telephoning while a baby was rolling a ball’. Responses were audio-recorded, later transcribed and analysed.

Analysis: Transcripts were analysed for the linguistic devices used to express the different aspects of the co-occurring or simultaneous actions depicted. As can be seen from Figure 1 the co-occurring animation consists of two co-occurring ongoing actions, e.g. *The person is painting while the monkey is drawing*. In order to ascertain what aspects of the co-occurring events Thai and English speakers are attending to the linguistic expressions, both imperfective aspect and lexical devices used to give duration to either one of the actions or both actions depicted in the animations were recorded e.g., one action: *The person is painting*, *the monkey draws*, or two actions: *The person is painting*, *the monkey is drawing* (see Table 2). The usage of simultaneous connectives to explicitly relate the co-occurrence of the two actions in the animations was also recorded, e.g., *The person is painting while the monkey is drawing*. Specifically imperfective and lexical expressions used to give duration to the two actions, and the connective used to relate the two actions were recorded – refer to Table 1.

Speakers have various choices in what aspect of the situation they encode linguistically. In the co-occurring animation speakers have several choices, they can express: 1. the ongoing action of one of the events only, 2. the ongoing action of both events, or 3. not explicitly express either of the ongoing actions. For the simultaneous relationship between the two actions speakers can: 1. explicitly express the co-occurring relation through the use of simultaneous temporal connectives, or 2. use a more ambiguous or less explicit connective, such as “and”, or 3. not express the relationship at all, i.e. juxtapose the two events without a connective.

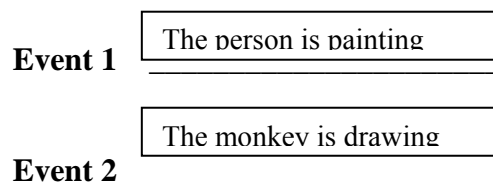


Figure 1: *The Co-occurring animation*

Results

In Table 2 the expression of the ongoingness of one of the events or both the events, and the simultaneous relationship between the two events are recorded. Analysis of variance was conducted to compare the expression of ongoingness for one clause, two clauses and the explicit expression of the simultaneous temporal relationship between the two actions in Thai and English participants to express the simultaneous relations depicted in the animations. It was found that expressing duration of one clause was not significantly different between Thai and English participants, however the expression of duration of both clauses was significantly different, as English participants more frequently expressed durative properties of the two actions depicted than Thai participants, (over all age groups $F(1,4)=44.2$ $p<.001$, for 4-year-olds $F(1,4)=30.27$ $p<.001$, 6-year-olds $F(1,4)=28.6$ $p<.001$, 7-year-olds $F(1,4)=19.6$ $p<.001$ and adults $F(1,4)=7.6$ $p<.01$). Furthermore, in relation to explicitly expressing the simultaneous relationship it can be seen from Table 2 that the Thai children did not explicitly express the simultaneous relationship at all and only the adults did. In English the 6-, 7-year-olds and adults expressed the simultaneous relation using simultaneous connectives. However, the usage of simultaneous connectives was not significantly different between Thai and English adults.

Table 2: *The expression of ongoingness of events and of the simultaneous relationship between the two events depicted in the simultaneous animations in Thai and English speakers.*

	Ongoingness of one event expressed		Ongoingness of both events expressed		Simultaneous temporal relationship expressed	
	Thai N=30	English N=30	Thai N=30	English N=30	Thai N=30	English N=30
4-year-olds	4	5	1**	13**	0	0
5-year-olds	12	10	7	7	0	0
6-year-olds	8	9	4**	21**	0	5
7-year-olds	11	6	9**	24**	0	10
Adults	12	6	13*	23*	19	14
All ages N=180	47	36	35**	89**	19	29

** $p<.001$ * $p<.01$

Table 3 shows other common strategies used by Thai and English participants to express the overlapping temporal relations depicted. It can be seen that juxtaposing clauses without a connective is a common strategy used by the younger children, 4 and 5 year olds in both Thai and English. However, 6- and 7-year-old English children do not use this strategy as much as the Thai children (for 4-year-olds $F(1,4)=18.6$ $p<.001$, 6-year-olds $F(1,4)=14.4$ $p<.001$, 7-year-olds $F(1,4)=25.4$ $p<.001$). This strategy declines in Thai adults and is not significantly different from English adults. English participants commonly used the

coordinating conjunction *and* to link clauses or actions in comparison with Thai participants, although to a lesser extent in English adults (for 4-year-olds $F(1,4)=40.8$ $p<.001$, 5-year-olds $F(1,4)=14.9$ $p<.001$, 6-year-olds $F(1,4)=12.7$, 7-year olds $F(1,4)=11.3$ $p<.001$ but not significant in adults). Figure 1 gives an overview of the main linguistic devices used to express the simultaneous temporal relation in Thai and English speakers. It can be seen that the imperfective is used more frequently by English than Thai participants to express the simultaneous relationship.

Table 3: *Other common strategies used by Thai and English participants to express the simultaneous animations*

	Clauses juxtaposed without connective		“and” used	
	Thai N=30	English N=30	Thai N=30	English N=30
4 year olds	24**	14**	0**	15**
5 year olds	19	16	0**	14**
6 year olds	18**	5**	4**	16**
7 year olds	14**	0**	7**	20**
Adults	4	6	4	10

** $p<.001$ * $p<.01$

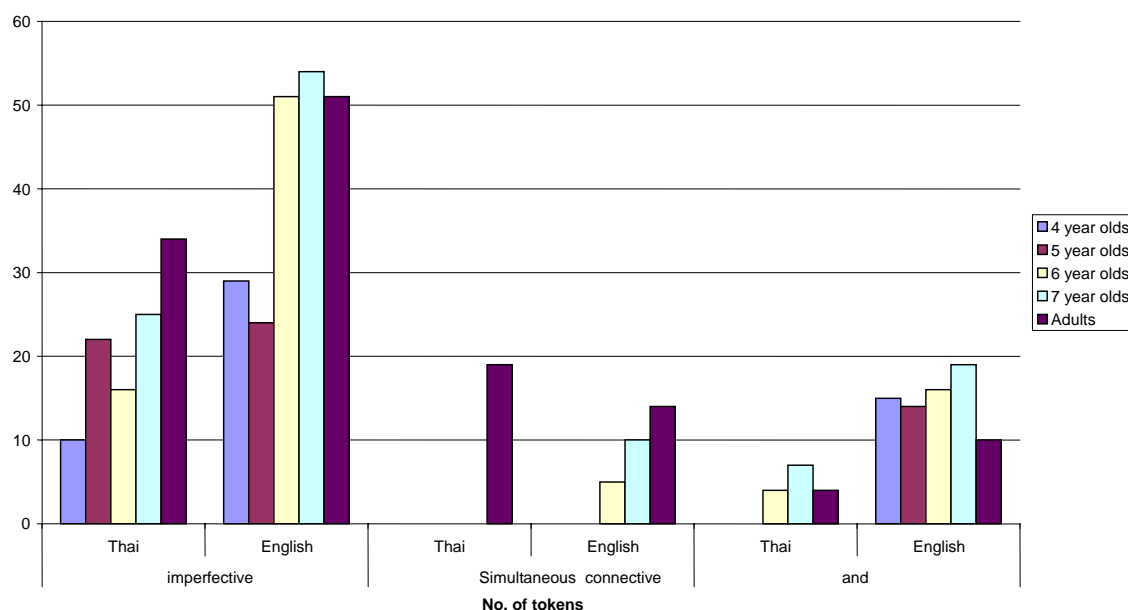


Figure 1: A summary of the main linguistic devices used in Thai and English to express the simultaneous animations.

Discussion

The current study has found that English speakers tend to express ongoingness of the events depicted more than Thai speakers. The simultaneous temporal relation between the two events is also expressed more by English than Thai children. Thai adults, however, tend to express this relation using simultaneous connectives to a similar extent as English adults. Temporal connectives were used by English children but not by Thai children, which indicates that English children produce these connectives at an earlier stage than Thai children at least for this particular task.

Language-specific preferences or differences were found; “and” was used more frequently by English speakers, whereas juxtaposition of two clauses without connective was used to a greater extent by Thai speakers. Young Thai and English children both show a predilection to juxtapose two clauses without a connective and in the older Thai children it is still a common strategy. However, this strategy declines in Thai adults.

In conclusion, it appears that the obligatoriness and ease of access of grammatical aspect is a factor in shaping Thai and English children’s responses to the overlapping temporal relations depicted. Thai children appear to be influenced by the non-obligatoriness of the aspectual system and consequently do not express the temporal relations as much as English children. However, Thai adults are not influenced to the same extent as the children and are able to draw on the full range of rhetorical options available in their language to express the temporal relations depicted.

References

- Bennett, M. & Partee, B. 1972. *Toward the logic of tense and aspect in English*. Bloomington: Indiana University Linguistics Club, 1978.
- Berman, R.A. and Slobin, D.I. (eds.) 1994. *Different ways of relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bowerman, M. and Levinson, S.C. 2001. *Language acquisition and conceptual development*, Cambridge: Cambridge University Press.
- Burusphat, S. 1991. *The structure of Thai narrative*. Dallas, Texas: The Summer Institute of Linguistics.
- Chan, Y-C. 1996. A semantic and cross-linguistic study of imperfective aspect in narrative discourse. *Research Papers in Linguistics and Literature* 5, 49-68.
- Gumperz, J.S. and Levinson, S.C. 1996. *Rethinking Linguistic Relativity*, Cambridge: Cambridge University Press.
- Hickmann, M. 2003. *Children's discourse: Person, space and time across languages*. Cambridge: Cambridge University Press.
- Hopper, P. J. 1979. Aspect and foregrounding in discourse. In T. Givon (ed.), *Syntax and Semantics 12: Discourse and Syntax*, (pp. 213-241). NY: Academic Press.
- Levinson, S.C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Luksaneeyanawin, S. 1993. Speech computing and speech technology in Thailand. *Proceedings of the Symposium on Natural Language Processing in Thailand*. Chulalongkorn University, 17-21 March 1993, 276-321.
- Meepoe, T.A. 1998. The interaction between lexical aspect and progressive imperfective in Thai: a discourse analysis of *kamlaj* and *yùu*. *Journal of Language and Linguistics*, 16(2), 56-66.
- Moens, M. & Steedman, M.J. 1986. *Temporal information and natural language processing*. Edinburgh Research Papers in Cognitive Science: Research Paper EUCCS/RP-2, Centre for Cognitive Science, Edinburgh University.
- Slobin, D.I. 1996. "From "thought and language" to "thinking for speaking"". In *Rethinking linguistic relativity*, J.J. Gumperz and S.C. Levinson (eds.), 70-96, Cambridge: Cambridge University Press.
- Smith, C.S. & Erbaugh, M. 2005. Temporal Interpretation in Mandarin Chinese. *Linguistics* 43(4), 713-756.
- Strömquist, S. and Verhoeven, L. 2004. "Typological and contextual perspectives on narrative development". In *Relating events in narrative: Typological and contextual perspectives*. S. Strömquist. and L. Verhoeven. (eds.), Mahwah, NJ: Lawrence Erlbaum.

REDUCED RELATIVE CLAUSES IN THAI AND VIETNAMESE¹

Natchanan Yaowapat
Chulalongkorn University
<Natchanan.Y@student.chula.ac.th>

Amara Prasithrathsint
Chulalongkorn University
<Amara.Pr@chula.ac.th>

Abstract

The present study focuses on a string of noun–modifier in two Southeast Asian languages, that is, Thai and Vietnamese. This string of noun–modifier, such as *nǎj sǎi phɛɛŋ* (book-expensive) ‘expensive book’ in Thai, often brings about a controversy as to what linguistic category the modifier belongs to. Some linguists label such a modifier an adjective. Others consider it to be a relative clause. In this study, we argue that it is a type of relative clauses; namely, reduced relative clause because it has important universal characteristics of relative clauses, that is, it functions as a noun modifier, contains a main verb and a gap that is co-referential with the head noun. We also aim to uncover syntactic and semantic constraints that govern its occurrence. Syntactically, it is found that the clause normally begins with a verb without a relativizer at the beginning of the clause. It is also found that the reduced relative clause is obligatorily adjacent to the head noun. Concerning relativization, this type of relative clause cannot occur with relativized indirect object. Semantically, we found that reduced relative clauses modify only indefinite and non-specific head nouns.

1. Introduction

According to Greenberg (1966:73-113), in most SVO languages nouns are normally followed by modifiers, as in the following examples from Thai.

Thai

- (1) *bān chǎn*
house I
‘my house’

¹ An earlier version of this paper was presented at The 16th Annual Meeting of Southeast Asian Linguistics Society (SEALS XVI), at Atma Jaya University, Jakarta, Indonesia. We would like to express our deep appreciation to the participants who made comments on our paper. They really helped us improve this paper. Also, we would like to acknowledge the financial support from the Thailand Research Fund (TRF) granted to us through the Royal Golden Jubilee Ph.D. Program (Grant No.PHD/0008/2547).

Paul Sidwell & Uri Tadmor, eds. *SEALS XVI: papers from the 16th meeting of the Southeast Asian Linguistics Society* (2006). Canberra, Pacific Linguistics, 2008, pp.143-159.

© Natchanan Yaowapat and Amara Prasithrathsint

- (2) *khon dii*
 person good
 ‘good person’
- (3) *nǎŋsǎi bon tó*
 book on table
 ‘a book on the table’
- (4) *nǎŋsǎi thîi yùu bon tó*
 book which be on table
 ‘a book which is on the table’

In this type of languages, there is a certain string of noun–modifier, such as *khon dii* (person-good) ‘a good person’, that often brings about a controversy in labeling the modifier. Several syntacticians, such as Upakitsilapasarn (1937), Noss (1964), Panupong (1989), Sookgasem (1996), label *dii* ‘good’ in (2) an adjective in the same way as they label other words like *kàw* ‘old’, *mày* ‘new’, *sǔuŋ* ‘tall’, *yày* ‘big’, *dam* ‘black’ etc. They consider those words to be adjectives according to functional and semantic criterion. Functionally, adjectives are lexical words that modify nouns. Semantically, adjectives are words that describe the quality or state of the nouns they modify.

On the other hand, other linguists, such as Kuno and Wongkhomthong (1981), Savetamalya (1996), Prasithrathsint (2000), Kullavanijaya (2006), label *dii* ‘good’ in (2) a relative clause and consider it to be syntactically equivalent to the relative clause with the relativizer *thîi*, as in (5).

- (5) *khon thîi dii*
 person that good
 ‘a person that is good’

According to them, the relative clause *dii* in (2) and *thîi dii* in (5) modify the head noun *khon*, and both consist of the main verb *dii*. Kuno and Wongkhomthong (1981) calls this type of relative clauses “thii-less relative clauses”. Savetamalya (1996) calls it “verbal relative clauses” in order to contrast with “nominal relative clauses” or relative clauses with *thîi*. Prasithrathsint (2000:260) labels it “non-finite relative clauses” in order to contrast with “finite relative clauses”.

In this study, we will argue that this type of noun modifier is a relative clause without a relativizer, henceforth, a “reduced relative clause”.

2. Previous analyses of reduced relative clauses

2.1 Previous analyses of reduced relative clauses in Western languages

According to Siloni (1997) and Doherty (2000), there is a type of relative clause without relative pronoun in English and French, and it functions as a noun modifier. They gave some examples as follow.

English

- (6) A man *reading a newspaper in the street* is a spy. (Siloni 1997:110)
 (7) He has found the key *you lost yesterday*. (Doherty 2000:57)
 (8) There is a girl *wants to see you*. (Doherty 2000:58)

French (Siloni 1997:132-133)

- (9) La fille *arrivant à Genève* est sud-américain.
 ‘The girl arriving in Geneva is South-American.’
 (10) La fille *arrivé à Genève* est sud-américain.
 ‘The girl arrived in Geneva is South-American.’

The clauses in italic in the examples (6) – (10) are labeled differently by different linguists, such as “participial relatives” and “reduced relatives” by Siloni (1997), “that-less relatives” by Doherty (2000), “contact clauses” by Jespersen (1969: 151-152). Regarding participial relatives, there are two subtypes, that is, “present participial relatives” as in (9), and “past participial relatives” as in (10).

These clauses are claimed to be “relative clauses” by these linguists because they function the same as what normal relative clauses do. That is, they restrict the class of objects denoted by the nouns they modify. For example, *arrivant à Genève* in (9) restricts the class of the girls.

In addition, Doherty (2000:59) mentioned a certain syntactic constraint of reduced relative clauses, namely, adjacency constraint. This constraint states that a reduced relative clause must appear immediately adjacent to a modified noun. Therefore, if it is extraposed from the modified noun, the sentence will be ungrammatical, as in Example (12) below.

English

- (11) John gave a book *he wrote* to Mary.
 (12) *John gave a book to Mary *he wrote*.
 (13) John gave a book to Mary *that he wrote*.

According to Siloni (1997:131), reduced relative clauses in French permit relativization of grammatical subjects, as in (9) and (10).). Relativization of others such as objects results in ungrammaticality in French. On the other hand, English allows relativization of both subjects and objects, as in (7) and (8).

Doherty (2000) observed that reduced relative clauses with relativized subjects in English normally occur in existential sentences and it-cleft sentences. For example,

English (Doherty 2000:72)

- (14) I have an idea *might work*. (existential)
 (15) There is something *keeps upsetting him*. (existential)
 (16) It was Bill *did it*. (it-cleft)

In addition, she states that a noun modified by a reduced relative clause is non-referential. In other words, the modified noun fails to denote an individual in the real world, as can be seen that reduced relative clauses can modify *someone* in (17) and modify *everybody* in (18).

English (Doherty 2000:91-92)

(17) We want someone *knows John*.

(18) Everybody *lives in the mountains* has an accent.

To summarize, reduced relative clauses are recognized by people who studied relative clauses in English and French.

2.2 Previous analyses of reduced relative clauses in Thai

A number of works mention reduced relative clauses in Thai, such as Kuno and Wongkhomthong (1981), Savetamalya (1996), Prasithrathsint (2000), Kullavanijaya (2006). These studies provide knowledge of syntactic and semantic characteristics of reduced relative clauses.

Savetamalya (1989, 1996) who labels reduced relative clauses in Thai as “verbal relative clauses” recognizes two syntactic characteristics of reduced relative clauses. Firstly, a reduced relative clause immediately follows the head noun. Secondly, there is one missing noun phrase in the relative clause and that missing noun phrase is co-referential with the head noun, for example.

(19) *khon rák rôt* (Savetamalya 1996:634)
 person love car
 ‘person who loves cars’

In the example (19), there is a missing subject noun phrase in the relative clause *rák rôt* which modifies the preceding noun *khon*. And that missing subject is co-referential with the head noun *khon*.

In addition, it is observed by Savetamalya (1989:75) that reduced relative clauses cannot modify deitic nouns like pronouns or proper nouns. In the following examples, reduced relative clauses are put in square brackets.

(20) *chǎn chōp phūyǐŋ [sǔay]*
 I like woman beautiful
 ‘I like a beautiful woman.’

(21) **chǎn chōp thəə [sǔay]*
 I like you beautiful

(22) **chǎn chōp púk [sǔay]*
 I like Puk beautiful

Kuno and Wongkhomthong (1981:198-204) who call reduced relative clauses in Thai as “thii-less relative clauses” also describe some syntactic characteristics of reduced relative clauses in order to contrast with compounds.

Firstly, a constituent or an element in the clause can take a modifier. If it is a compound, any constituent in it cannot have a modifier.

- (23) (Kuno and Wongkhomthong 1981:198)
phǒm rúucàk khon [sǐ nǎŋsǐ ní]
 I know person buy book this
 ‘I know a person who bought this book.’
- (24) (Kuno and Wongkhomthong 1981:201)
Khon tii hǔa khun mǎawaanní pen khray
 person hit head you yesterday be who
 ‘Who is the person who hit your head yesterday?’

Secondly, a constituent in the relative clause can be coordinated or conjoined, for example.

- (25) (Kuno and Wongkhomthong 1981:203)
thíi roonriian ní mii ʔaacaan [hǔa-dii tɛɛ cay-khêɛp] lǎay khon
 at school this have teacher head-good but heart-narrow several CLF
 ‘In this school, there are some teachers who are bright but narrow-minded.’

The characteristics of modifying and coordinating a constituent in the clause differentiate this type of relative clauses from compound nouns like *khon-khàp-rót* (person-drive-car) ‘chauffeur’, or *sǎaphâa-kàw* (cloth-old) ‘used cloth’. A constituent in compound nouns, such as *khàp rót* or *kàw* cannot be modified or coordinated.

Concerning the semantics of reduced relative clauses in Thai, four observations can be summarized from the previous studies. Firstly, Kuno and Wongkhomthong (1981), Savetamalya (1989, 1996) state that reduced relative clauses are semantically compositional. The meaning of reduced relative clauses can be predicted from their elements or parts, whereas the meaning of compound nouns is non-compositional. For example, the meaning of the compound noun *khon-khàp-rót* is not ‘a person who drives a car’ but it is ‘driver’ or ‘chauffeur’.

Besides, Kuno and Wongkhomthong (1981:207) describe that a reduced relative clause gives general description, general characterization, or general evaluation to its head noun, whereas a relative clause with relativizer gives personal judgment or evaluation to its head noun.

- (26) (Kuno and Wongkhomthong 1981:215)
mǎawaanní phǒm cəə khon [thíi phúut phaasǎa ʔaŋkrít kèn]
 yesterday I meet person REL speak language English well
 ‘Yesterday I met a person who (I think) speaks English well.’
- (27) (Kuno and Wongkhomthong 1981:215)
khon [phúut phaasǎa ʔaŋkrít kèn] mák cà pen khon hǔa-dii
 person speak language English well frequently be person head-good
 ‘A person who (people in general think) speaks English well is frequently clever.’

In addition to the two semantic characteristics of reduced relative clauses mentioned earlier, Prasithrathsint (2000:260) states that a reduced relative clause is used when the concepts expressed by the head noun and the relative clause are very close, whereas the one with relativizer is used when the two concepts are distant. She gave an example as follow.

- (28) *phrá? [dii] mii mâak tɛ coon [thîi dii] mii nɔɔy*
 monk good have many but thief REL good have few
 ‘There are many good monks but few good thieves.’

In (28), the concept of the head noun *phrá?* ‘monk’ is typically good and its concept is very close to the meaning of the verb *dii* ‘good’. So it is compatible with the unmarked relative clause *dii*. On the other hand, the concept of the head noun *coon* ‘thief’ is typically bad and very distant from the meaning of *dii*. So it is compatible with the marked relative clause *thîi dii*.

The last observation on semantic characteristics of reduced relative clauses in the previous studies is the one described by Kullavanijaya (2006:48). She states that a head noun modified by a reduced relative clause is indefinite, whereas a head noun modified by a relative clause with relativizer is definite. She gave the same examples as in Kuno and Wonkhonthong (1981:215) as follow.

- (29) *mâwaannîi phǒm cəə khon [thîi phûut phaasǎa ʔaŋkrít kèŋ]*
 yesterday I meet person REL speak language English well
 ‘Yesterday I met a person who (I think) speaks English well.’

- (30) *khon [phûut phaasǎa ʔaŋkrít kèŋ] mák cà pen khon hǔa-dii*
 person speak language English well frequently be person head-good
 ‘A person who (people in general think) speaks English well is frequently clever.’

In (29), the head noun *khon* ‘person’ is definite in a way that the speaker can identify its referent. In other words, the referent of the head noun is marked in the speaker’s thought so the speaker uses the marked relative clause to modify the head noun. On the other hand, in (30), the speaker has no particular referent in his thought so he uses the unmarked relative clause to modify the head noun.

2.3 Previous analyses of reduced relative clauses in Vietnamese

According to Emeneau (1951), Miller (1976), Thompson (1991), Nguyễn (1997), and Nguyen (2004), a relative clause in Vietnamese is marked by the invariant particle or the relativizer *mà* at the beginning of the clause and the relative clause follows a modified noun, for example.

- (31) Thompson (1991:263)
tôi đã tìm thấy quyển sách [mà anh nói hôm nọ]
 I PST find see CLF book REL you talk the other day
 ‘I found the book that you were talking about the other day.’

So far, there is no in-depth study on reduced relative clauses in Vietnamese. Certain scholars who have dealt with relative clauses in Vietnamese, such as Miller (1976), Nguyễn (1997), and Nguyen (2004), only mention that the relativizer *mà* at the beginning of relative clauses can be absent or optional, for example.

- (32) *cái người thợ may* [(*mà*) *đến sáng ngày*] (Nguyễn 1997:183)
 CLF person tailor REL come morning day
 ‘the tailor who came this morning’
- (33) *cuốn từ điển* [(*mà*) *tôi thích*] (Nguyen 2004:59)
 CLF dictionary REL I like
 ‘the dictionary that I like’
- (34) *khi* [(*mà*) *mẹ tôi mất*] (Nguyễn 1997:183)
 time REL mother I die
 ‘the time when my mother died’
- (35) *nơi* [(*mà*) *tôi ra đời*] (Nguyễn 1997:183)
 place REL I be born
 ‘the place where I was born’

In addition, there is no explanation by those scholars for the syntactic and semantic constraints on the presence and the absence of the *mà* marker.

As can be seen from the review of previous studies, reduced relative clauses are mentioned in various languages but are not analyzed thoroughly in Southeast Asian languages. Such an analysis is needed so as to solve the continual problem of labeling such a grammatical category. Even though some syntacticians recognize reduced relative clauses in Thai and Vietnamese and describe, to some extent, their syntactic and semantic characteristics, there are still questions: If they are true relative clauses, how similar they are to universal relative clauses? Are there any features in their behavior that distinguish them from normal relative clauses? This study aims to find out the answers to these questions.

Therefore, the purpose of this study is to justify that reduced relative clauses in Thai and Vietnamese are a type of relative clause and to uncover syntactic and semantic constraints that govern their occurrence.

It is hypothesized here that there is a type of relative clause without relativizer, namely, reduced relative clauses, in the two languages, along with another typical type of relative clause with relativizer and that reduced relative clauses have some syntactic and semantic characteristics which are different from those of relative clauses with relativizer.

The data that the analysis is based on was taken from written texts, particularly from newspapers, elicitation, and observations. Approximately, there are fifty sentences for analysis in each language.

The data were analyzed in the framework of typology. It is hoped that the result will provide insight into the syntactic typology of Southeast Asian languages and the universality of human languages.

3. Characteristics of reduced relative clauses in Thai and Vietnamese

This section shows the result of the analysis of reduced relative clauses in Thai and Vietnamese. In 3.1, we will strongly argue that there is a type of relative clauses called *reduced relative clauses* in the two languages with a number of reasons. In 3.2, we will present syntactic differences between reduced relative clauses and relative clauses with relativizer. In 3.3, we will present semantic differences between reduced relative clauses and relative clauses with relativizer.

3.1 Syntactic similarities between reduced relative clauses and relative clauses with relativizer

In this study, we argue that the noun modifier *dii* ‘good’ in the string *khon dii* (person-good) is a relative clause, namely, reduced relative clause which begins with a verb and there is no relativizer at the beginning of the clause. Such the noun modifier like *dii* has three universal characteristics of relative clauses. That is, 1) it follows the modified noun, 2) it contains a co-referential gap with the modified noun, and 3) the co-referential gap functions as subject, direct object, oblique or possessor.

Firstly, a reduced relative clause follows the head noun as relative clauses with relativizer do. Some examples from Thai and Vietnamese are illustrated below. In this study, reduced relative clauses are surrounded by square brackets and modified nouns are typed in bold.

Thai

- (36a) *ráan níi khǎay tɛɛ **ʔaahǎan** [∅ ʔarɔ̀y]*
 shop this sell only food delicious
 ‘This shop sells only delicious food.’
- (36b) *ráan níi khǎay tɛɛ **ʔaahǎan** [thîi ∅ ʔarɔ̀y]*
 shop this sell only food REL delicious
 ‘This shop sells only delicious food.’
- (37a) *chán mây chɔ̀ɔp **khon** [∅ mây mii khwaamrúu]*
 I not like person not have knowledge
 ‘I do not like persons who have no knowledge/who are not knowledgeable.’
- (37b) *chán mây chɔ̀ɔp **khon** [thîi ∅ mây mii khwaamrúu]*
 I not like person REL not have knowledge
 ‘I do not like persons who have no knowledge/who are not knowledgeable.’

Vietnamese

(38a) *Tôi thích những người [∅ tối]*
 I like PLU person kind
 ‘I like people who are kind.’

(38b) *Tôi thích những người [mà ∅ tối]*
 I like PLU person REL kind
 ‘I like people who are kind.’

(39a) *Anh ta có ý tưởng [∅ không sáng tạo]*
 he have idea not creative
 ‘He has ideas which are not creative.’

(39b) *Anh ta có ý tưởng [mà ∅ không sáng tạo]*
 he have idea REL not creative
 ‘He has ideas which are not creative.’

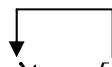
In (36a), *thơm* ‘delicious’ is considered a relative clause solely consisting of the verb and it follows the head noun *thức ăn* ‘food’ as the relative clause *thức ăn thơm* in (36b) does. In (37a), *mây mii khwaamrúu* ‘who have no knowledge’ is a relative clause following the head noun *khon* ‘person’ as the relative clause *thức ăn mây mii khwaamrúu* in (37b) does. In (38a), *tối* ‘kind’ is a relative clause following the head noun *những người* ‘people’ as the relative clause *mà tối* in (38b) does. In (39a), *không sáng tạo* ‘which are not creative’ is a relative clause following the head noun *ý tưởng* ‘idea’ as the relative clause *mà không sáng tạo* in (39b) does.

Secondly, reduced relative clauses in Thai and Vietnamese contain a gap that is co-referential with the head noun as relative clauses with relativizer do. Taken (36a) – (39b) again as the examples of the case, there is a gap or a missing noun phrase in the subject position (preceding the main verb in the relative clause). The gaps in (36a) – (39b) are represented by the symbol \emptyset and the coreference between the gaps and the head nouns is represented by the arrows. In (36a) and (36b), the missing subject is co-referential with the head noun *thức ăn* ‘food’. In (37a) and (37b), the missing subject is co-referential with the head noun *khon* ‘person’. In (38a) and (38b), the missing subject is co-referential with the head noun *những người* ‘people’. In (39a) and (39b), the missing subject is co-referential with the head noun *ý tưởng* ‘idea’.

Thirdly, a head noun modified by a reduced relative clause can have a co-referential gap which functions as subject, direct object, oblique or possessor as a head noun modified by relative clause with relativizer. For instance:

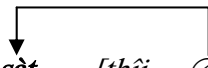
Thai

(40a) Subject



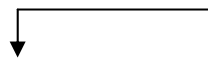
ηuu pen sàt [∅ mii samǎɔŋ khanàat lék]
 Snake be animal SUBJ have brain size small
 ‘Snakes are animals that have small brains.’

(40b) Subject



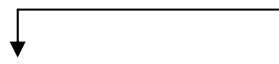
ηuu pen sàt [thîi ∅ mii samǎɔŋ khanàat lék]
 Snake be animal REL SUBJ have brain size small
 ‘Snakes are animals that have small brains.’

(41a) Direct object



chǎn chǎɔp ʔàan nǎŋsǎi [yîim ∅ maa càak hǎŋsamùt]
 I like read book borrowDO come from library
 ‘I like reading books that are borrowed from a library.’

(41b) Direct object

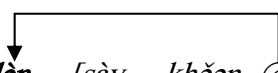


chǎn chǎɔp ʔàan nǎŋsǎi [thîi yîim ∅ maa càak hǎŋsamùt]
 I like read book REL borrow DO come from library
 ‘I like reading books that are borrowed from a library.’

(42a) Oblique

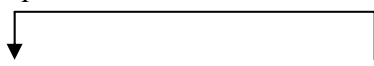
sày khǎɔŋ nay klǎŋ yǎʔ kǎən pay
 put thing in box many over go
 ‘Put too many things in a box.’

(42b) Oblique



klǎŋ [sày khǎɔŋ ∅ yǎʔ kǎən pay] mǎk phaŋ rew
 box put thing OBL many over go often break fast
 ‘A box in which (you) put too many things often gets broken soon.’

(42c) Oblique



klǎŋ [thîi sày khǎɔŋ ∅ yǎʔ kǎən pay] mǎk phaŋ rew
 box REL put thing OBL many over go often break fast
 ‘A box in which (you) put too many things often gets broken soon.’

(43a) Possessor

chǎn mây chǎp plùuk tǒnmáay [râak Ø yaaw]
 I not like grow tree root POSS long
 ‘I do not like growing trees the root of which is long.’

(43b) Possessor

chǎn mây chǎp plùuk tǒnmáay [thîi râak Ø yaaw]
 I not like grow tree REL root POSS long
 ‘I do not like growing trees the root of which is long.’

Vietnamese

(44a) Subject

Cách đó vài giường là một bệnh nhân [Ø vừa qua đời vì bệnh AIDS]
 Beside several bed be one patient SUBJ just pass away because AIDS
 ‘Beside several beds is a patient who has just passed away because of AIDS.’

(44b) Subject

Cách đó vài giường là một bệnh nhân [mà Ø vừa qua đời vì bệnh AIDS]
 Beside several bed be one patient REL SUBJ just pass away because AIDS
 ‘Beside several beds is a patient who has just passed away because of AIDS.’

(45a) Direct object

Thức ăn [bạn nấu Ø] ngon
 food you cook DO delicious
 ‘Food that you cook is delicious.’

(45b) Direct object

Thức ăn [mà bạn nấu Ø] ngon
 food REL you cook DO delicious
 ‘Food that you cook is delicious.’

(46a) Oblique

tôi nghĩ đến khi [mẹ tôi mất ∅]
 I think of time mother I die OBL
 ‘the time when my mother died’

(46b) Oblique

tôi nghĩ đến khi [mà mẹ tôi mất ∅]
 I think of time REL mother I die OBL
 ‘the time when my mother died’

(47a) Possessor

tôi thích phụ nữ [tóc ∅ dài]
 I like woman hair POSS long
 ‘I like a woman whose hair is long.’

(47b) Possessor

tôi thích phụ nữ [mà tóc ∅ dài]
 I like woman REL hair POSS long
 ‘I like a woman whose hair is long.’

Because of the three universal characteristics previously presented, we strongly maintain that in Thai and Vietnamese, there are reduced relative clauses which follow head nouns and contain a co-referential gap with various functions.

3.2 Syntactic differences between reduced relative clauses and relative clauses with relativizer

Syntactically, there are three differences between reduced relative clauses and relative clauses with relativizer in Thai and Vietnamese. That is, 1) a reduced relative clause begins with a verb; 2) a reduced relative clause is obligatorily adjacent to the modified noun; and 3) a reduced relative clause cannot have a co-referential gap that functions as an indirect object.

Firstly, a reduced relative clause normally begins with a verb, whereas a relative clause with relativizer begins with a relativizer. For reduced relative clauses in Thai, there is no *thìi* at the beginning of the clause. Example (48a) below illustrates a reduced relative clause, whereas example (48b) illustrates relative clauses with *thìi*.

Thai

(48a) *ηuu pen sàt [mii samǎŋ khanàat lék]*
 snake be animal have brain size small
 ‘Snakes are animals that have small brains.’

- (48b) *ḡuu pen sàt [thī mīi samǎḡḡ khanàat lék]*
 snake be animal REL have brain size small
 ‘Snakes are animals that have small brains.’

In Vietnamese, there is no *mà* at the beginning of reduced relative clauses. Example (49a) below illustrates a relative clause without *mà*, whereas example (49b) illustrates a relative clause with *mà*.

- (49a) *Chúng tôi không thích những chính trị gia [lấy tiền ừ các doanh nghiệp]*
 we not like PLU politician take money from
 PLU businessman
 ‘We do not like politicians who take money from businessmen.’

- (49b) *Chúng tôi không thích những chính trị gia [mà lấy tiền từ các doanh nghiệp]*
 we not like PLU politician REL take money from
 PLU businessman
 ‘We do not like politicians who take money from businessmen.’

The fact that reduced relative clauses normally begin with a verb supports what Savetamalya (1989, 1996) claimed about “verbal relative clauses”. However, there is an exception for this. When a modified noun has a co-referential gap functioning as a possessor, the reduced relative clause begins with a noun, not a verb. For instance,

Thai

- (50) *chǎn chǎp phūuyǎ [phǎm ∅ yaaw]*
 I like woman hair POSS long
 ‘I like women with long hair.’
-

Vietnamese

- (51) *Tôi thích phụ nữ [tóc ∅ dài]*
 I like woman hair POSS long
 ‘I like women with long hair.’
-

Concerning the second syntactic constraint, it is found in this study that a reduced relative clause, unlike relative clauses with relativizer, must be adjacent to a head noun. It cannot be extraposed from the head noun. For example,

Thai

- (52a) *ʔaacaan* [sǎɔn dii] mii kii khon
 teacher teach good have how many CLF
 ‘How many teachers who teach well?’
- (52b) **ʔaacaan* mii kii khon [sǎɔn dii] (Extrapolation)
 teacher have how many CLF teach good
- (53a) *ʔaacaan* [thii sǎɔn dii] mii kii khon
 teacher REL teach good have how many CLF
 ‘How many teachers who teach well?’
- (53b) *ʔaacaan* mii kii khon [thii sǎɔn dii] (Extrapolation)
 teacher have how many CLF REL teach good

The example (53b) shows that if a relative clause is marked by *thii*, it is not necessarily adjacent to the modified noun *ʔaacaan* ‘teacher’.

The last syntactic difference between reduced relative clauses and relative clauses with relativizer is that a reduced relative clause cannot have a co-referential gap functioning as an indirect object whereas a relative clause with relativizer can. In other words, the occurrence of reduced relative clauses with co-referential indirect objects results in ungrammaticality, for instance.

Thai

- (54a) **dək* [chǎn sǐ nǎŋsǐ hây ∅] nâasǎɔnsǎan
 Child I buy book give IO pitiful
 ‘A child whom I gave a book is pitiful.’
- (54b) *dək* [thii chǎn sǐ nǎŋsǐ hây ∅] nâasǎɔnsǎan
 child REL I buy book give IO pitiful
 ‘A child whom I gave a book is pitiful.’

Vietnamese

- (55a) **Tôi gặp cậu bé* [bạn cho ∅ tiền hômqua]
 I meet boy you give IO money yesterday
 ‘I met the boy whom you gave money yesterday.’

- (55b) Tôi gặp **cậubé** [mà bạn cho ∅ tiền hômqua]
 I meet boy REL you give IO money yesterday
 ‘I met the boy whom you gave money yesterday.’

So far, we have proposed three syntactic differences between reduced relative clauses and relative clauses with relativizer in Thai and Vietnamese. In the next section, we will show the semantic differences between the two types of relative clauses.

3.3 Semantic differences between reduced relative clauses and relative clauses with relativizer

In addition to syntactic constraints, there are some semantic constraints on the occurrence of reduced relative clauses in Thai and Vietnamese. It is found that reduced relative clauses modify only indefinite or non-specific head nouns. For definite or specific head nouns, relative clauses with relativizer are required, as can be seen in the following examples.

Thai

- (56a) Definite (Relative clause with *thîi*)

dèk sǎam khon nán [thîi khamooy khǎw] *thùuk càp léew*
 Child three CLF that REL steal thing PASS arrest already
 ‘Those three children who stole the thing were arrested.’

- (56b) Definite (Reduced relative clause)

**dèk sǎam khon nán [khamooy khǎw]* *thùuk càp léew*
 Child three CLF that steal thing PASS arrest already
 ‘Those three children who stole the thing were arrested.’

- (57a) Specific

paacaan thîi sǎon phâakwíchaa phaasǎasàat [thîi sǎon dii] mii
 Teacher REL teach department linguistics REL teach good have
kii khon
 how many CLF
 ‘How many teachers from the linguistics department who teach well?’

- (57b) Specific (Reduced relative clause)

**paacaan thîi sǎon phâakwíchaa phaasǎasàat [sǎon dii] mii*
 Teacher REL teach department linguistics teach good have
kii khon
 how many CLF
 ‘How many teachers from the linguistics department who teach well?’

Vietnamese(58a) Definite (Relative clause with *mà*)*10 sinhviên Thái [mà học tiếng Việt Nam] đó sẽ sang*

10 student Thai REL study language Vietnam that will go

Việt Nam tháng sau

Vietnam week next

‘Those 10 Thai students who studied Vietnamese will go to Vietnam next week.’

(58b) Definite (Reduced relative clause)

**10 sinhviên Thái [học tiếng Việt Nam] đó sẽ sang*

10 student Thai study language Vietnam that will go

Việt Nam tháng sau

Vietnam week next

Việt Nam tháng sau

Vietnam week next

‘Those 10 Thai students who studied Vietnamese will go to Vietnam next week.’

Concerning the last syntactic constraint that we propose in Section 3.2 which states that reduced relative clauses are impossible when relativizing indirect objects, it is observed that definiteness of head nouns is relevant. Taken (54a) and (54b) as the examples again, we normally know for whom we will buy a book, therefore the indirect object or the recipient in this case is definite. We can identify who that person is. Therefore, that is why reduced relative clauses are not compatible with relativized indirect objects.

It should be noticed that what we found here support what has been observed in certain previous studies. Syntactically, the facts that a reduced relative clause follows a head noun and that there is a co-referential gap in the clause support the claim by Savetamalya (1989, 1996). Semantically, that fact that a modified noun is indefinite supports the claim by Kullavanijaya (2006).

4. Conclusion

The present study argues that there is a type of relative clause without relativizer in two Southeast Asian languages, Thai and Vietnamese. That type of relative clauses is called here “reduced relative clauses”. The classification of this type of construction as a type of relative clauses is justified by their similarities to relative clauses with relativizer. However, it is found that there are some differences between reduced relative clauses and the other type, such that it is obligatorily adjacent to the modified noun phrase, that it cannot have an indirect object that is coreferential with the modified noun phrase, and that the modified noun phrase must be indefinite or non-specific.

References

- Doherty, Cathal. 2000. Clauses without “That”: The Case for Bare Sentential *Complementation in English*. New York & London: Garland Publishing.
- Emeneau, M. B. 1951. *Studies in Vietnamese (Annamese) Grammar*. Berkeley: University of California Press.

- Greenberg, Joseph. H. 1966. *Universal of Language*. Second Edition. Cambridge: M.I.T Press.
- Jespersen, Otto. 1969. *Analytic Syntax*. New York: Holt, Rinehart and Winston Inc.
- Kullavanijaya, Pranee. 2006. Noun-modifying clauses. *Controversial Constructions in Thai Grammar: Relative Clause Constructions, Complement Clause Constructions, Serial Verb Constructions, and Passive Constructions*, edited by Amara Prasithrathsint. Research Monograph. Bangkok: Chulalongkorn University Press.
- Kuno, Susumu and Preeya Wongkhomthong. 1981. Relative clauses in Thai. *Studies in Language* 5. Amsterdam: John Benjamins B.V. Publisher.
- Miller, Carolyn. 1976. Structural Ambiguity in the Vietnamese Relative Clauses. *Mon-Khmer Studies* 5: 233-67.
- Nguyễn, Đình-Hoà. 1997. *Vietnamese*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nguyen, Tuong Hung. 2004. The Structure of the Vietnamese Noun Phrase. Ph.D.Dissertation. Boston University.
- Noss, Richard B. 1964. *Thai: Reference grammar*. Washington: Foreign Service Institute.
- Panupong, Vichin. 1989. The Structure of Thai: Grammatical System. Bangkok: Ramkhamhaeng University Press.
- Prasithrathsint, Amara. 2000. Adjectives as verbs in Thai. *Linguistic Typology* 4:251-271.
- Savetamalya, Saranya. 1989. *Thai Nouns and Noun Phrases: A Lexicase Analysis*. Ph.D.Dissertation. Honolulu: University of Hawaii.
- _____. 1996. Verbal relative clauses as adnominal modifiers in Thai. *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, January 8-10, 1996, Volume 2:627-646.
- Siloni, Tal. 1997. *Noun Phrases and Nominalizations: The Syntax of DPs*. Dordrecht: Kluwer Academic Publishers.
- Sookasem, Prapa. 1996. The predicative–adjective construction in Thai. *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, January 8-10, 1996, Volume 2:579-607.
- Thompson, Laurence C. 1965. *A Vietnamese Reference Grammar*. Honolulu: University of Hawaii Press.
- Upakitsilapasarn. 1937. *Principles of Thai: Syntax*. Bangkok: Thaiwatanapanich Press.

Symbols used in this study

Thai

v = mid tone

∨̂ = low tone

∨̂ = falling tone

∨̂ = high tone

∨̂ = rising tone

CLF = classifier

DO = direct object

IO = indirect object

OBL = oblique

PASS = passive

PLU = plural

POSS = possessive

REL = relativizer

SUBJ = subject