TEMPORAL AND CO-VARYING CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

Niklas Jonsson
Temporal and co-varying clause combining in Austronesian languages
Semantics, morpho-syntax and distributional patterns

Niklas Jonsson
Abstract

Various semantic relations are represented by combined clause constructions in the languages of the world. This study investigates combined clause constructions for ten distinct semantic relations in a cross-section of Austronesian languages. The semantic relations selected for the study are of a temporal or co-varying nature, the former commonly expressed in English by such markers as when, then, until, etc. and the latter by if, so, because, etc. The research falls into three main domains: semantics, morpho-syntax and distributional patterns.

First, the study provides an overview of the semantics of temporal and co-varying relations in the Austronesian languages. Several subdistinctions are found to be made within the semantic relations investigated, some cross-linguistically rare, such as general vs. fulfilled purpose, and others more common, such as distinguishing counterfactuality and concessivity for conditionals. The study also explores polysemic relation markers, and several patterns are identified. The single most common pattern is the overlap between (open) conditional and (non-past) co-occurrence relations, for which many Austronesian languages employ the same relation marker.

Second, the study develops a morpho-syntactic typology of Austronesian clause combining based on three parameters related to features common to clause combining constructions. The typology divides the constructions into five different types that are ranked with regard to structural tightness based on notions of compactness and dependency. Some constructions, cutting across several types, are also discussed; in particular, asymmetric coordination, which involves the use of a coordinator to connect a fronted topicalized adverbial clause to the rest of the sentence. It is argued that the construction facilitates cognitive processing as the coordinator renders both clauses as independent, and the hearer does not have to try and fit in a cumbersome initial constituent into the sentence structure.

Finally, the study explores the distributional patterns of the morphosyntactic types across the semantic relations, as well as across three geographical areas in the Austronesian region. In the former case, a clear correlation is found between posteriority and result relations on the one hand and looser structural types on the other. The other relations show more complex patterns and cannot easily be ordered into a linear hierarchy. One of the analyses made, however, shows that several pairs of individual relations display clear differences in average tightness. The distribution of types across three Austronesian macro-regions revealed few differences between the areas, although two tendencies could be detected: the Oceanic languages generally employ slightly looser morpho-syntax for clause combining constructions, while the Formosan and Philippine languages employ slightly tighter morpho-syntax.
# Table of contents

*Abbreviations* .................................................................................................................................................. xi

*Acknowledgements* ........................................................................................................................................ xvii

1 Introduction ...................................................................................................................................................... 1

1.1 The domain of study .................................................................................................................................. 2

1.2 Aim and scope of the study ...................................................................................................................... 7

1.3 Why Austronesian? Why temporality and co-variation? ......................................................................... 9

1.4 Terminological notes ................................................................................................................................ 11

1.4.1 Clauses, combined clauses and sentences ......................................................................................... 11

1.4.2 Relation marker .................................................................................................................................... 12

1.4.3 Construction ......................................................................................................................................... 12

1.5 Organization of the thesis ......................................................................................................................... 14

2 Earlier research on clause combining ....................................................................................................... 15

2.1 Coordination and subordination .............................................................................................................. 16

2.1.1 The notions of coordination and subordination ............................................................................... 16

2.1.2 Defining criteria for subordination .................................................................................................... 19

2.2 Parallel continua, iconicity and implicational hierarchies ..................................................................... 29

2.2.1 Parallel continua: Lehmann's typology ............................................................................................. 29

2.2.2 Iconicity in clause combining: the RRG approach ........................................................................... 33

2.2.3 Typological hierarchies and functional motivations: Cristofaro (2005) on subordination ............... 40

2.3 Clause combining in Austronesian languages ........................................................................................ 42

3 Austronesian languages and sources of data ............................................................................................. 48

3.1 The Austronesian language family ......................................................................................................... 48

3.1.1 Genealogical relationships .................................................................................................................. 49

3.1.2 Typological characteristics ................................................................................................................. 59

3.2 Sources, sample and database ............................................................................................................... 65

3.2.1 Sources of data ................................................................................................................................... 65

3.2.2 Sample of languages ............................................................................................................................ 66

3.2.3 The database ...................................................................................................................................... 72

3.2.4 Factors possibly distorting the results ............................................................................................... 72
4 Semantic relations ........................................................................................................ 75
  4.1 General discussion ................................................................................................. 75
  4.1.1 Background and the semantic relations used in the present study .................. 75
  4.1.2 The universality of semantic relations ............................................................. 77
  4.2 Practical considerations in the semantic classification ........................................... 81
  4.2.1 Principles and procedures ................................................................................ 81
  4.2.2 Discourse levels of interpretation ..................................................................... 85
  4.2.3 Some terminological notes with regard to temporal ordering ............................ 86
  4.2.4 Semantic relations reference map ..................................................................... 88
  4.3 Relational categories in the Austronesian languages ............................................. 90
     4.3.1 Co-occurrence ............................................................................................... 91
     4.3.2 Posteriority .................................................................................................. 95
     4.3.3 Anteriority .................................................................................................... 96
     4.3.4 Terminal boundary ....................................................................................... 99
     4.3.5 Initial boundary ............................................................................................ 103
     4.3.6 Condition ...................................................................................................... 104
     4.3.7 Concession ................................................................................................... 113
     4.3.8 Purpose ......................................................................................................... 116
     4.3.9 Reason .......................................................................................................... 121
     4.3.10 Result .......................................................................................................... 123
  5 Patterns of polysemy ............................................................................................... 126
     5.1 Three paths of development well-attested cross-linguistically ......................... 126
     5.2 Austronesian polysemic relation markers ............................................................ 134
        5.2.1 Co-occurrence-condition polysemy ............................................................... 135
        5.2.2 Posteriority-result polysemy ....................................................................... 137
        5.2.3 Condition-concession polysemy .................................................................. 139
        5.2.4 Purpose-reason and purpose-result polysemy ............................................ 141
        5.2.5 Posteriority-anteriority polysemy ............................................................... 144
  6 Morpho-syntactic properties of Austronesian clause combining ................................. 146
     6.1 Classifying parameters ...................................................................................... 146
        6.1.1 Deviating morpho-syntax ............................................................................ 147
        6.1.2 Syntactic relation between clauses ............................................................... 155
        6.1.3 Explicitness of semantic relation ................................................................. 164
     6.2 Morpho-syntactic types ...................................................................................... 171
        6.2.1 The tightness scale ....................................................................................... 171
        6.2.2 Asymmetric coordination ............................................................................ 186
7 Distributions ........................................................................................................... 199
  7.1 Distribution of semantic relations across structural types ....................... 199
    7.1.1 Paradigmatic iconicity ............................................................................. 199
    7.1.2 Distributional patterns .......................................................................... 202
    7.1.3 Range of variation .................................................................................. 213
    7.1.4 Discussion ............................................................................................... 219
  7.2 Distribution across areas .............................................................................. 222

8 Summary ............................................................................................................. 227
  8.1 Semantics of Austronesian clause combining ............................................. 228
    8.1.1 Relation internal nuances of meaning ..................................................... 228
    8.1.2 Polysemy .................................................................................................. 233
  8.2 Morpho-syntax of Austronesian clause combining ....................................... 234
    8.2.1 Structural typology ................................................................................ 234
    8.1.2 Constructions of special interest ............................................................. 236
  8.3 Distributional patterns .................................................................................. 239

References .............................................................................................................. 242

Appendix 1: Native speaker’s questionnaire ...................................................... 263
Appendix 2: Linguist’s questionnaire .................................................................... 263
Appendix 3: Principal source material ................................................................. 268
Appendix 4: Language samples .......................................................................... 272
Abbreviations

Below, three sets of abbreviations are listed that are used throughout the thesis, mainly in connection with language examples (in the glossing of grammatical terminology and when referring to a native speaker consultant as the source of an example), but also in tables to save space and in text when the full form is somewhat cumbersome (see Other abbreviations below). At times, I altered the glosses and/or translations of original examples; however, in such cases I was always careful to not distort the relevant original meanings. The reason for altering glosses is that the originals sometimes use certain glosses for specific theoretical reasons that may not be relevant for our present purposes. Free translations of examples may be slightly changed for better illuminating some point under discussion.

Glossing of grammatical terminology:

1 first person
2 second person
3 third person
ABIL abilitative
ABS absolutive
ACC accusative
ACT action
ACTR actor
ACTV active voice
AF actor focus voice
ANA anaphora
ANT anteriority
APPL applicative
ART article
ASS assertion
AUG augmentative
BEN benefactive
CASE case marker
CAUS causative
CF counterfactual
CIT citation form
CL classifier
CMPL  completer
CNT    continuative
COMP   complementizer
CONC   concessive
CONN   connective
CONST  construct state
COOC   co-ocurrence
DAT    dative
DEF    definite
DEIC   deictic
DEM    demonstrative
DER    derivational
DET    determiner
DETRANS detransitivizer
DIR    directional
DISTPST distant past tense
DRAM   dramatic
DS     different subject
DU     dual
DUB    dubitative
DX     indexer
EMPH   emphatic
ERG    ergative
ES     echo subject
EX     exclusive
EXT    extended
FEM    feminine
FOC    focus marker
FUT    future
FX     affix
GEN    genitive
GIV    given information
GNR    generic tense
HES    hesitation marker
HI     verb prefix in Muna (multifunctional)
HOD    hodiernal (recent) past tense
HON    honorific
HYP    hypothetical mood
IF     instrument focus voice
IMM    immediacy
IMMFUT immediate future tense
IMP    imperative
IN     inclusive
INCH   inchoative aspect
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCOOC</td>
<td>initial stage co-occurrence</td>
</tr>
<tr>
<td>IND</td>
<td>indicative</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INSTR</td>
<td>instrument</td>
</tr>
<tr>
<td>INT</td>
<td>intention</td>
</tr>
<tr>
<td>INTEN</td>
<td>intensifier</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>IPFV</td>
<td>imperfective aspect</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>LIG</td>
<td>ligature</td>
</tr>
<tr>
<td>LIM</td>
<td>limiter</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>MASC</td>
<td>masculine</td>
</tr>
<tr>
<td>MOD</td>
<td>mood marker</td>
</tr>
<tr>
<td>MOM</td>
<td>momentary</td>
</tr>
<tr>
<td>NEC</td>
<td>necessity</td>
</tr>
<tr>
<td>NEG</td>
<td>negative</td>
</tr>
<tr>
<td>NF</td>
<td>nonfactuality</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NPST</td>
<td>non-past tense</td>
</tr>
<tr>
<td>NSG</td>
<td>non-singular</td>
</tr>
<tr>
<td>NTOP</td>
<td>non-topic</td>
</tr>
<tr>
<td>NZR</td>
<td>nominalizer</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>OBLG</td>
<td>obligation</td>
</tr>
<tr>
<td>PART</td>
<td>particle</td>
</tr>
<tr>
<td>PASS</td>
<td>passive voice</td>
</tr>
<tr>
<td>PERI</td>
<td>peripatetic (moving person marker in Coastal Konjo)</td>
</tr>
<tr>
<td>PF</td>
<td>patient focus voice</td>
</tr>
<tr>
<td>PFT</td>
<td>perfect tense</td>
</tr>
<tr>
<td>PFTV</td>
<td>perfective aspect</td>
</tr>
<tr>
<td>PHO</td>
<td>phoric</td>
</tr>
<tr>
<td>PHR</td>
<td>general phrase introducer</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>POT</td>
<td>potential</td>
</tr>
<tr>
<td>PRES</td>
<td>present tense</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>PROP</td>
<td>proprieteive</td>
</tr>
<tr>
<td>PRS</td>
<td>presentative</td>
</tr>
</tbody>
</table>
Native speaker consultants (referred to in the study):

AJ     Aiga Jonsson     Samoan
MJ     Marijane Jonsson  Tagalog

Other abbreviations:

ant.     anteriority
C. Cagayan Agta  Central Cagayan Agta
CEMP     Central-East Malayo-Polynesian
CMP      Central Malayo-Polynesian
conc.    concession
cond.    condition
coo      coordinate
coo-oc.  co-occurrence
dev      deviating
EMP      East Malayo-Polynesian
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>expl</td>
<td>explicit</td>
</tr>
<tr>
<td>Form</td>
<td>Formosan</td>
</tr>
<tr>
<td>freq.</td>
<td>frequency</td>
</tr>
<tr>
<td>impl</td>
<td>implicit</td>
</tr>
<tr>
<td>Indon.</td>
<td>Indonesia</td>
</tr>
<tr>
<td>init b</td>
<td>initial boundary</td>
</tr>
<tr>
<td>initial b.</td>
<td>initial boundary</td>
</tr>
<tr>
<td>initial bound.</td>
<td>initial boundary</td>
</tr>
<tr>
<td>MI</td>
<td>The Malaysia-Indonesia area</td>
</tr>
<tr>
<td>non-d</td>
<td>non-deviating</td>
</tr>
<tr>
<td>OC</td>
<td>The Pacific Ocean area</td>
</tr>
<tr>
<td>p.c.</td>
<td>personal communication</td>
</tr>
<tr>
<td>PNG</td>
<td>Papua New Guinea</td>
</tr>
<tr>
<td>post.</td>
<td>posteriority</td>
</tr>
<tr>
<td>PT</td>
<td>The Philippines-Taiwan area</td>
</tr>
<tr>
<td>purp.</td>
<td>purpose</td>
</tr>
<tr>
<td>reas.</td>
<td>reason</td>
</tr>
<tr>
<td>SHWNG</td>
<td>South Halmahera West New Guinea</td>
</tr>
<tr>
<td>sub</td>
<td>subordinate</td>
</tr>
<tr>
<td>TC</td>
<td>temporal and co-varying</td>
</tr>
<tr>
<td>term. b.</td>
<td>terminal boundary</td>
</tr>
<tr>
<td>term. bound.</td>
<td>terminal boundary</td>
</tr>
<tr>
<td>term. boundary</td>
<td>terminal boundary</td>
</tr>
<tr>
<td>WMP</td>
<td>West Malayo-Polynesian</td>
</tr>
</tbody>
</table>
Many people have helped and supported me in many different ways during the writing of this dissertation. Firstly, I would like to express my gratitude to my supervisor, Masja Koptjevskaja-Tamm, for guiding me through this process. Even when my work for various reasons was less than productive, she always showed interest in my progress. She has been a constant support since I first came to the Department of Linguistics at Stockholm University, sharing my interest in typology as well as in Polynesian languages. Thanks also, Masja, for inviting my family and me to visit your family at your lovely summer house in Norrtälje.

I would also like to thank my assistant supervisor, Östen Dahl, for so meticulously reading and commenting on my dissertation. Although you were my supervisor only during the final stages of my writing, your comments and advice have definitely helped to make this a better dissertation than it otherwise would have been.

My very warm thanks also goes to Ulrike Mosel for opening up her home to me when I was in Kiel to find Austronesian texts and grammars at the Christian-Albrechts-Universität at the beginning of my dissertation work. Thank you, Ulrike, for many interesting discussions about Samoan, and for taking time to show me around some nice places in the Kiel area.

Collecting data for the dissertation was a very time-consuming task. I would not have been able to gain access to all the grammars, texts and articles from various faraway places without the help of several people. My sincere thanks goes especially to Henry Yungli for providing me with various materials on Formosan languages, to Marian Klamer for sending me her papers on Kambera, to Joel Bradshaw for sending me his work on Yabem, Iwal and Numbami (including much as yet unpublished material), to John Bowden for sending me material on Taba and Ambai, to Ronald Whistler for sending me his Sawai discourse notes and to Alexandre François for sending me his Araki grammar. All of these people also most graciously answered all the questions that I kept asking. I am indebted to you all.

Other people who did not mind taking their time to discuss details of various Austronesian languages in lengthy e-mail conversations were Geoff Woollams on Karo Batak, Alan Jones on Mekeo, Frantisek Lichtenberk on Manam, Hope Hurlbut on Labuk Kadazan, Arthur Holmer on Seediq, and Barbara Friberg on Coastal Konjo. Thanks for sharing your knowledge and wisdom with me.
A very special thanks goes to Timothy Friberg who spent hours, days and even weeks filling in my linguistic questionnaire on clause combining constructions in the most detailed way for Coastal Konjo. You made so much information accessible to me in this language, and I only hope that you yourself will be able to make some use of the data that you collected for my questionnaire. I would also like to thank the following people who most kindly filled in my questionnaire as well: Jan Allen for Kankanaey, Sherri Brainard for Karao and John A. Walsh for Ulithian.

I also had the opportunity to interview a number of native speakers of Austronesian languages about clause combining constructions in their respective languages. Although this method of collecting data proved to be too time consuming to be carried out for a large number of languages, I am very grateful to all who participated for the time they spent and for the information they willingly shared: Ruben Maury on Kurudu, Teresita Luistro on Hiligaynon, Jeanina Santiago on Tagalog, Bessie Rave on Ilokano, Josepbine Eriksson on Cebuano, Veronique Forsberg on Malagasy and Dewi Carina Widowaty on Indonesian.

At the department in Stockholm, I would like to thank my office mates through the years, especially Harald Berthelsen, Josefin Larsson, and Ljuba Veselinova, with whom I shared rooms for extended periods of time. Thanks for sharing ideas, teaching experiences and thoughts in general. I would also like to specially thank my next-door office neighbor for several years, Mikael Parkvall, whose linguistic interests as well as taste in music coincide with mine to such a large extent.

For much moral support as well as interesting discussions, I wish to warmly thank Andrea Kiso and Christine Ericsdotter, who were always ready to share a smile and a kind word. I have also received support and good advice over the years from Päivi Juvonen, Eva Lindström, Henrik Liljegren and Emil Perder. Thank you so much. And I will not forget, Emil, how you willingly spent the better part of a couple of hours helping me retrieve my data when my computer crashed towards the end of the writing process.

I am very grateful as well to Elizabeth Spenader for checking the English example sentences in my questionnaire, to Kjell Lindström for all the interesting discussions about languages and linguistics (and everything else) and for commenting on the entire thesis, and to Leif Asplund for providing valuable comments especially on Austronesian language examples and genealogical classification. Thank you all.

I also wish to extend my gratitude to Christian Nilsson and Karin Dellby at Scania for allowing me to adjust my working hours there between 2009 and 2012 in order to have time to finish the dissertation.

Finally and most importantly, I would like to thank my family: my parents, Ruth and Claes, for instilling in me the patience that has been so much needed in the process of completing a dissertation; my brother, Markus, and
his wife, Marijane, for all the encouragement, as well as for picking up my books in Manila when they were there on vacation. Marijane – a native speaker of Tagalog – also kindly answered all my questions about her language. Most of all, I thank my wife, Aiga, for putting up with countless questions about Samoan, her native language, and for putting up with long working hours for so many years; and I thank my beautiful children, Nancy and Noel, for waiting so patiently for dad to be done with his book. At last, it is done.

Niklas Jonsson
Stockholm, February 2012
This thesis is an investigation into the semantics and morpho-syntax of Austronesian clause combining. As such, it explores the form and meaning, as well as the relation between form and meaning, of a set of clause combining constructions in a cross-section of Austronesian languages. These constructions involve those representing temporal relations (p before/after/until q, etc.), as in example (1) below, and those representing co-varying relations (p if/because/so that q, etc.), as in example (2).

1. **Palauan**

```
a  sèchèl-ik  a  mirrael  ěr  a  uche  ěr
PHR  friend-1.SG  PHR  leave.PST  in  PHR  front  of
```

```
a  k-bo  k-mèrek  ěr  a  urer-èk
PHR  1.SG.HYP-be.PST  1.SG.HYP-finish.PST  to  PHR  work-1.SG
```

'My friend left before I finished my work.' (Josephs 1975, p 447)

2. **Tagalog**

```
hindi  ako  tu-tugtog  ng  piyano
be.not  1.SG.TOP  RD:AF.IRR-play  GEN  piano
```

```
pag  na-tu-tulog  ang  tatay
if  AF-IPFV-sleep  TOP  father
```

'I won't play the piano, if father is sleeping.'
(Schachter & Otanes 1972, p 468)

One of the most important questions that the thesis attempts to answer is whether or not there is a systematic correlation between the structural "tightness" of these constructions and the semantic relations that they represent. I will refer to constructions showing a greater number of interclausal dependencies and more compact morpho-syntax as structurally tighter than those that show fewer and less. Interclausal dependencies prevail, for instance, when one clause has a structure that cannot occur on its own as a simplex clause, or when one clause depends on the other for the interpretation of
grammatical features and/or the identity of actants. Compact morpho-syntax in clause combining is manifested by fewer morphemes/lexemes and/or more bound morphology than in corresponding simplex clauses. Consider these English examples, which express a purpose and a result, respectively.

3. She goes out to look for him.
4. He is quite tired so he goes back home.

As is evident, the purpose construction in example (3) is morpho-syntactically more compact and displays stronger interclausal dependence than the result construction in example (4). For instance, the clause denoting the purpose lacks an explicit subject and has a verb in the infinitive form that lack any TMA expression. Both of these features distinguish it from the result clause in the second example, which displays patterns identical to that of a simplex clause in English. The purpose clause is also subordinate to the preceding matrix clause, while the result clause is coordinate to the preceding clause. Some scholars (see further section 1.2) claim that there is a regular cross-linguistic correspondence between the structural tightness of clause combining constructions and their relational meaning, such that some semantic relations are consistently expressed by tighter morpho-syntax, while other relations are expressed by looser morpho-syntax. What is even more interesting is that there are also scholars who claim the opposite to be true: that the structural variation displayed by clause combining constructions in the languages of the world cuts across the semantic relations expressed, without clear correspondences. The present study will investigate the connection between form and meaning for clause combining constructions in the Austronesian language family. In the course of the investigation, several semantic and grammatical aspects of Austronesian clause combining will be explored.

1.1 The domain of study

The research reported here is conducted within the framework of linguistic typology. The dominant method of conducting cross-linguistic typological research is to delimit a domain of study in semantic/functional terms and then to investigate the different forms and structures used to represent this semantic/functional domain in human languages. The assumption is that while all languages may essentially express the same ideas, the forms they employ to do so are language specific. Therefore, formal properties are not always ideal in delimiting a domain of study when it comes to comparison between languages, as discussed by Stassen (1985) and Croft (2003), among others. In reality, most typological research simultaneously employs both
formal and semantic properties to varying degrees. Even in semantically based typological research, there are often implicit structural conditions involved. Comrie (1985, p 9), for instance, in his book on tense, defines his domain of study as "grammaticalized expression[s] of location in time" but leaves it implicit that he is only concerned with verbs, disregarding markers such as -kaman, 'until', in the Imbabura Quechua word luniskamanka, 'until Monday' (Cole 1985, p 127), that also express time. Other authors are more explicit about the significance of both semantic and formal/structural criteria in defining their domains of study (see for example Haspelmath 1997a, 1997b, Wälchli 2003).

The present work will involve both semantic and morpho-syntactic criteria in the characterization of its domain of study. At the most general level, the study is fenced off in semantic terms, as it is concerned with various linguistic manifestations of the temporal and co-varying (TC) relations between states of affairs.¹ Temporal relations concern the orientation in time of one state of affairs in relation to another, while co-varying relations concern different ways in which one state of affairs co-varies with another in some sense. Specific temporal relations include expressions of co-occurrence, anteriority, posteriority, terminal boundary and initial boundary. Specific co-varying relations include expressions of conditionality, concessivity, purpose, reason and result. These will be described in detail in chapter 4. Some of the relations are illustrated in examples (5)-(9) below.

5. *We took care of their garden while they were on vacation.*
   \[\text{CO-OCCURRENCE}\]

6. *She sat on her bed half asleep until the alarm rang.*
   \[\text{TERMINAL BOUNDARY}\]

7. *He can't spell because he never went to school.*
   \[\text{REASON}\]

8. *We grant you unlimited access, provided you comply with our terms.*
   \[\text{CONDITION}\]

9. *She can already read although she's only three.*
   \[\text{CONCESSION}\]

Although TC relations may of course obtain between more than two states of affairs, scholarly attention has only been directed to relations between two

¹ An explanatory note should be made about the term 'state of affairs'. It is used here in the same sense as in much of the literature on Functional Grammar (e.g. Dik 1997; Cristofaro 2005) to mean the "conception of something which can be the case in some world" (Dik 1997, p 107). In Functional Grammar, it is taken to be a cover term for conceptual entities commonly referred to as processes, actions, states, etc. Other terms often used with similar connotations as 'state of affairs' include 'situation' and 'event' (see e.g. Comrie 1985; Bybee et al 1994; Haspelmath 1997b). Occasionally in this study, I have used the term 'event' synonymously with 'state of affairs', since the latter is sometimes a bit cumbersome in text.
states of affairs. The semantic criteria thus constitute two components: relevant constructions (i) code two principal states of affairs, and (ii) express a TC relation between those states of affairs.

A word about the term 'co-varying relations' is also necessary here. These relations are quite heterogeneous semantically, and their differences lie on several planes. They have sometimes been called fundamentally causal and sometimes fundamentally conditional.\(^2\) Kortmann (1997) refers to them simply as CCC-relations (for condition, concession and cause). None of these terms seems to capture what they have in common. It seems to me that the term 'co-varying relations' does, however, whether the co-variation of the states of affairs involved is hypothetical (as in conditional relations), expected but altered (as in concessive relations), intended (as in purpose relations) or directly/indirectly caused (as in reason and result relations).

The general semantic criteria are complemented by morpho-syntactic ones. First of all, since states of affairs are typically coded as clauses, we are mainly concerned with temporality and co-variation as manifested in constructions of clause combining. What constitutes a clause may be difficult to specify universally, but the standard of measure has been simplex clauses in the languages of concern. In the unmarked case, there is a correlation between states of affairs and clauses, but some of the morpho-syntactic strategies that languages have developed to code relations between states of affairs involve processes by which structures are deprived of some of their clausal properties (see section 6.1.1). Such constructions may include nominalized clauses and so-called converb forms (structures based on verb forms used adverbially), as well as other derived constructions. Some English equivalents to these constructions are exemplified below.

10. *The study was designed before the opening of the new medical center.* NOMINALIZATION
11. *He stood up to see better.* INFINITIVE
12. *I’m always careful (when) driving at night.* CONVERB

As long as constructions such as these constitute the conventional way of representing specific TC relations in a language, they are clearly relevant for the present study. However, the study is not concerned with TC relations as represented between a clause and an underived simple noun phrase. The clause-hood criterion thus directs our attention to the type of constructions in example (13) rather than the type in example (14).

\(^2\) Erben (1972 [cited in Kortmann 1997, p 83]), for example, on German grammar, labels cause, condition and concession as causal relations, while Eisenberg (1989 [cited in Kortmann 1997, p 83], also on German, considers the same semantic relations to be conditional.
13. *He stayed in his room because [he didn’t feel well].*  
   (Bracketed text is a clause.)

14. *He had to close his eyes because of [the light].*  
   (Bracketed text is NOT a clause.)

Furthermore, we are concerned with relational expressions represented as single sentences (though multi-clausal) rather than as individual clauses that are merely adjacent in discourse. That is, the example in (15) is relevant while that in (16) is not.

15. *We felt cold because he opened the window.*  
16. *He opened the window. Therefore we felt cold.*

Identifying equivalents to the latter example as two individual clauses rather than as a single clause combining construction could of course be difficult with respect to languages one does not speak. In this matter, I simply followed the writing conventions of my sources. These conventions are sometimes – though far from always – explicitly explained in linguistic terms (mainly prosody and syntax).

Relational meaning can be emphasized or represented in various other stylistically marked ways in most languages. What we are interested in here, however, is the unmarked, that is, most conventional and stylistically neutral way of rendering the relation between two states of affairs. Thus, while the construction in (17) below is not the target of our attention, the construction in (18) is.

17. *The very reason he returned was that Clara had moved back into town.*  
18. *He returned because Clara had moved back into town.*

This means that overarching both the semantic and morpho-syntactic criteria is the pragmatic criterion, by which constructions selected for study should be unmarked and stylistically neutral.

The criteria delimiting our focus of interest may thus be summarized and stated in the following terms:
Table 1. Criteria delimiting the domain of study

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic</td>
<td>Conventional and unmarked</td>
</tr>
<tr>
<td></td>
<td>The constructions studied constitute the conventional and unmarked ways of expressing the relevant semantic relations in the respective languages.</td>
</tr>
<tr>
<td>Semantic</td>
<td>Two states of affairs</td>
</tr>
<tr>
<td></td>
<td>the constructions code two principal states of affairs (not intimately linked as subcomponents of a single overall happening).</td>
</tr>
<tr>
<td>TC</td>
<td>The relation between the states of affairs is of a temporal or co-varying character.</td>
</tr>
<tr>
<td>Morpho-syntactic</td>
<td>Clause combining</td>
</tr>
<tr>
<td></td>
<td>Each of the two states of affairs is coded by a clause, using simplex clauses in the respective languages as the standard of measure, or a structure with clause-like features, if this is the unmarked way of expressing a certain relation.</td>
</tr>
<tr>
<td>Single sentence</td>
<td>The construction is possible to interpret as a single sentence – prosodically, syntactically and/or according to normal writing conventions.</td>
</tr>
</tbody>
</table>

Given that the distinction between clauses and non-clauses is not cross-linguistically (or language internally) absolute, but rather scalar, it has been impossible to delimit the selection of constructions precisely. At the fringes of the selection, there are some gray zone cases, and by necessity, some arbitrary decisions have been made in the process of collecting data. The present criteria, nonetheless, go a long way toward the selection of a representative and comparable set of constructions. Descriptive grammars and other sources of data naturally differ in resolution of detail and prioritized topics, but assuming that grammarians and experts in the languages selected for this study chose relevant constructions to describe, representative of actual language use (and we cannot really assume otherwise, unless other sources clearly indicate this), we can still be reasonably confident that we are studying the most common and conventionalized patterns for each language.
1.2 Aim and scope of the study

The present study has several aims – briefly mentioned initially in this chapter – in relation to TC clause combining in Austronesian languages. These all build on each other and will be outlined in more detail in this section.

Firstly, the study aims to provide an overview of the semantics of TC relations in the Austronesian languages. What relations tend to be present in the sample languages, and what nuances of meaning are expressed? Which Austronesian distinctions are common, and which are rare? An example of a very rare typological distinction found in a few of the Austronesian languages is one that holds between general purpose and realized purpose constructions. In the relevant languages, the equivalent of *Sam went to the store to buy groceries* may be represented by one construction if the speaker wants to communicate that Sam both had the intention and also actually did buy groceries, and by a different construction if the speaker wants to convey Sam's intention without any information about whether he actually bought groceries or not. Within the set of basic relational meanings investigated, several Austronesian sub-distinctions will be reviewed. A more specific aim concerning the semantics of clause combining constructions is to investigate the polysemy of relation markers. To what extent can one relation marker be used in constructions expressing different semantic relations? The thesis also aims to discuss Austronesian polysemic patterns against the background of diachronic shifts of meaning shown to have taken place in relation markers cross-linguistically.

The next general aim is to develop a structural typology of Austronesian clause combining. The typology will be based on three different parameters related to features common to clause combining constructions in general. Some of the Austronesian constructions have not received much attention in the previous literature, neither in Austronesian studies nor in general typological studies on clause combining, which makes them an interesting target of attention. In particular, the Austronesian preference for deriving relation markers for some relations from verbs has implications for various constructions, and the tendency to use coordinators to link initial subordinate clauses to the rest of the sentence is relevant to the discussion of the distinction between coordination and subordination. This means that the Austronesian situation may also contribute to our general understanding of issues pertaining to clause combining and TC relations.

A further aim – and the most central one to the thesis – is to map out the distribution of the construction types across semantic relational categories, such as co-occurrence, condition, and concession, etc. The distribution of types across semantic categories is relevant for an evaluation of the notion of
paradigmatic iconicity in clause combining, i.e. the extent to which the degree of integration between states of affairs for different semantic relations corresponds to the degree of tightness for morpho-syntactic types. Such an iconic correspondence is proposed universally by some (e.g. Foley & Van Valin 1984; Cristofaro 2005) and denied by others (e.g. Lehmann 1988; Harris 1988). The proponents of an iconic correlation draw upon examples such as (3) and (4) above. Examples showing similar patterns can also be found in Austronesian languages. Compare the two Palauan examples below, the first of which semantically encodes a purpose, and the second one a result.

19. Palauan

a. ak ulusbech ěr a droteo
   1.SG need.PST at PHR NAME
   ěl mɐruul ěr a subeš-e-k
   PURP do.PRES at PHR homework-1.SG.POSS

   'I needed Droteo to (help me) do my homework.'
   (Josephs 1975, p 301)

b. ak di mililil
   1.SG just play.PST

   mɐ ak mle otsir ěr a test
   and.so 1.SG go.PST fail at PHR test

   'I just fooled around, so I failed the test.' (Josephs 1975, p 440)

In the (a) example, the subject of the purpose clause is obligatorily left out, and the verb must occur in the present tense form, while the result clause in the (b) example has no such restrictions – it exhibits the same morpho-syntactic pattern as a simplex clause in Palauan. The purpose clause therefore shows a more compact structural pattern than the result clause in this case. It is also dependent on the preceding clause for the interpretation of the would-be subject, while this is not the case for the result clause. However, a semantic relation can be represented by more than one type of construction in most languages. Compare the two purpose constructions from Erromanggan below.
In this case, the purpose clause in (a) has the pattern of a simplex clause, and the argument marker may vary independently of the subject in the preceding clause (although the tense value is restricted to present or future). The purpose clause would also allow a lexical subject had it been different from the subject in the clause preceding it (Crowley 1998). The purpose clause in (b), on the other hand, does not allow a lexical subject, since it obligatory must be the same as that in the preceding clause in this construction. The purpose clause verb, which cannot be used in simplex clauses, carries a purposive prefix, disallowing regular argument/TMA markers. Thus, by showing stronger dependencies, the purpose construction in (b) is considered structurally tighter than that in (a). In light of on examples such as these, scholars rejecting an iconic correlation in clause combining argue that stylistic and pragmatic factors are more important.

Finally, with regard to distribution, another aim is to shed light on the extent to which structural types correlate with geographic areas within the Austronesian region.

1.3 Why Austronesian? Why temporality and co-variation?

The most important reason for choosing the Austronesian languages as target languages for these investigations is that constructions coding TC relations between states of affairs were not studied systematically for the entire Austronesian language family previously, making it a virgin area of research. Comparative research of grammatical phenomena in a specific language family is also a way of evaluating cross-linguistic generalizations, as well as highlighting the dividing line between what is truly universal and what is
common mainly in one family. In this respect, Austronesian was a natural choice, since it is a language family towards which my research interests have been largely directed in the past.

With regard to the linguistic constructions being investigated here, there are a number of reasons why TC relations are an interesting object of study. One thing that distinguishes us as human beings is our ability to understand, interpret and predict the consequences of different states of affairs, including how two or more states of affairs are tied together in a co-varying relationship. This ability is, of course, fundamental both in human social interactions and a prerequisite for any advances in all sciences. In fact, we are constantly looking for reasons and results, and making hypotheses about correlations in trying to understand the world around us. This is a universal human trait that we cannot easily abandon, and this ability is absent or at least far less developed in all other animals, including the other primates (see Gärdenfors 2003). A temporal relationship prevails by necessity between any two factual states of affairs, whether also co-variationally related or not. Thus, as we perceive the world, it is inevitable that we construe reality in terms of temporal and co-varying relationships between states of affairs. Temporality and co-variation is at the core of human thinking.

Differences and similarities in the way these are expressed in the languages of the world may provide clues to a deeper understanding of human conceptualization at large. The fundamental character of these relations is reflected in several ways in the world's languages. (i) Most if not all languages have systematic and regular ways of expressing at least some of the specific relations subsumable under the labels of temporality and co-variation (e.g. reason, condition, co-occurrence, etc.). (ii) There is evidence that these semantic relations receive special treatment compared to other semantic relations in grammar and discourse. In European languages, for instance, adverbial subordinating conjunctions expressing co-variation and temporality between states of affairs are more frequently used, and more often grammaticalized as one word subordinators, than those expressing other comparable relations (Kortmann 1997). (iii) Many co-varying and temporal relations seem to be some of the first semantic relations that children learn to code linguistically (see Bloom et al 1980 for a study of English-speaking children).

Furthermore, relations of temporality and co-variation are linked synchronically in polysemic patterns and diachronically in patterns of semantic change (Kortmann 1997). A case in point is the English adverbial subordinator since, with both a temporal and causal reading, the latter being historically derived from the former. This indicates a clear conceptual link between the two types of relations and provides interesting prospects for investigating both together (see further chapter 5).
1. INTRODUCTION

1.4 Terminological notes

Terminological issues will mostly be taken up in the specific chapters where they are relevant, but some very general concepts and distinctions will be discussed here.

1.4.1 Clauses, combined clauses and sentences

By the term 'clause', I am referring to a syntagm containing at least one predicate, expressing at least one state of affairs. By 'simplex clause', I mean a clause containing only one predicate, expressing just one state of affairs, and being capable of constituting a single grammatical utterance in a language.

By 'clause combining construction', I intend to refer to any combination of clauses interpretable as one grammatical utterance, as well as combinations of a clause and clause-like structures (such as nominalized clauses or converb forms). This means that clause-like structures, forming part of some clause combining constructions, will also normally be referred to as 'clauses' in this study.

The distinction between a clause and a sentence will not be crucial in the present study. I take the term 'sentence' in most cases to be synonymous with 'clause', i.e. a cover term subsuming all constructions referred to by the terms 'simplex clause' and 'clause combining construction'. However, a sentence must always be grammatical on its own, while a clause may be part of a clause combining construction and may not necessarily be grammatical on its own. While I use the term 'clause', for instance, for clause-like structures used in clause combining constructions, the term 'sentence' cannot be used in this way.

The terms 'clause' and 'sentence' may also be differentiated in perspective, such that clauses are building blocks in syntax, while sentences are building blocks within discourse. In some cases it might be difficult to know whether two clauses are combined into a clause combining construction or rather constitute sentences linked in discourse. This is particularly the case with asyndetically linked clauses where both have simplex clause patterns, but also often when the relation marker has an adverbial function within its clause (e.g. then). Some of the relation markers used in clause combining are also used as discourse linkage devices. Many authors refer to prosodic criteria in determining whether two clauses constitute a clause combining construction or a series of sentences in discourse. Thus, Grimes (1991) speaks of juxtaposed clauses in Buru with a non-final (rising) phonological juncture between them as clause combining rather than as sentences in a series, and Woollams (1996) mentions that Karo Batak clauses may be combined into
one construction merely through phonological means without a formal link between them.

Similarly, only prosodic properties (or punctuation in text, if used consistently) may distinguish between the examples below as being a case of clause combining (21) and discourse linking (22).

21. *It's cold and rainy, so I won't go out tonight.*
22. *It's cold and rainy. So, I won't go out tonight.*

Unfortunately, punctuation is not as consistently used as one would wish. In the present thesis, I have only investigated relation markers used in clause combining constructions and excluded markers that may only function as discourse linking devices, based on the information and examples found in the source material.

### 1.4.2 Relation marker

By relation marker (RM), I mean the element(s) indicating the semantic and/or syntactic relation between the clauses in question. In this study, I am only concerned with relation markers consisting of grammatical items or function words, or sequences of words showing signs of grammaticalization. Relation markers may have a purely syntactic function, such as the neutral coordinator *and* or the neutral complementizer *that*, or a purely semantic function such as the coherence adverbs *then* and *so*, or both a semantic and syntactic function, such as the subordinators *when*, *because*, *after*, *if*, etc.

### 1.4.3 Construction

I will restrict the use of the term 'construction' to clause combining constructions, referring to specific morpho-syntactic patterns. Since the distinction between different semantic relations is crucial to this thesis, I will refer to constructions containing different relation markers as different constructions, even if they follow the same morpho-syntactic pattern. Thus, the English examples below constitute different constructions, according to this view:

23. *We will go to the beach if the sun is shining.*  
    CONDITION
24. *We will go to the beach because the sun is shining.*  
    REASON

However, in many languages, the same relation marker can be used for different semantic relations. In these cases, as long as the morpho-syntactic patterns of the clauses involved stay the same, various possible relational interpretations are regarded as constituting the same construction. Compare
the Samoan examples below, both manifesting the same construction, although the semantic relation expressed is different.

25. Samoan

a. 'āfai 'ole'ā e alu e te moe,
   when FUT 2.SG go 2.SG GNR sleep.SG
   ia, funulu ou nifo
   well brush.PL 2.PL.POSS tooth

   'When you go to bed, brush your teeth.' (AJ)

b. 'āfai 'ole'ā e mālō,
   if FUT 2.SG win
   ia, vili mai a'u
   well call DIR 1.SG

   'If you win, call me' (AJ)

The following principles were used for judging constructions to be identical or different:

I. Same or no relation marker (RM) and same grammatical features in the clauses but different relational meanings > same construction

II. Slightly different RMs where one is clearly a contraction or other regular variant of the other, while other features in the clauses are the same > same construction

III. Clearly different RMs, regardless of other features in the clauses > different constructions

IV. Same or no RM but different grammatical features in the clauses and different relational meanings > different constructions

V. Same or no RM, but different grammatical features in the clauses and no difference in relational meaning > see a-c below:
   a. Difference consisting of optional non-RM morphemes (particle, adverb, etc.) > same construction
   b. Difference consisting of paradigmatic grammatical variation (e.g. TMA) > same construction
   c. Difference consisting of other formal features (e.g. nominalization) > different constructions
1.5 Organization of the thesis

The thesis is organized as follows: Following the introduction in chapter 1, chapter 2 provides a review of earlier research relating to clause combining, in general as well as in Austronesian languages. Chapter 3 provides some background information on the Austronesian language family from a genealogical and typological viewpoint, as well as information about the language samples, sources of data and database of constructions used in the thesis. Chapter 4 describes the semantic relations being investigated, as well as various subdivisions of these relations found in the sample languages. Chapter 5 is also semantically oriented and discusses polysemic patterns for Austronesian relation markers and diachronic developments. Chapter 6 classifies the Austronesian constructions identified along three parameters into types, and in addition, the chapter describes an interesting strategy for information structuring that I will refer to as asymmetric coordination, apparently common in Austronesian languages. In chapter 7, the distribution of semantic relation across structural types and geographic areas is mapped out, which for the former will constitute an evaluation of claims made in the linguistic literature of iconic patterns existing between semantic relations and morphosyntax, such that specific relations are consistently represented by more compact morpho-syntax than other relations in the languages of the world. Chapter 8 concludes the thesis by summarizing its findings.
Earlier research on clause combining

The subject of clause combining has been studied extensively in linguistics from a wide variety of viewpoints and theoretical perspectives, ranging from discourse coherence studies, to semantically based studies, to logical investigations, to morpho-syntactic studies on subordination and coordination and other aspects of clause combining. The subject has also been treated in some detail in many individual languages, mostly European and in particular English, but also others, including many Austronesian languages, as for instance in a series of articles in the framework of tagmemics, stemming mostly from the 1960s and 1970s. However, with few exceptions – notably Longacre (1970) and Crowley (2002a) (see section 2.3), neither of which focuses exclusively on combined clauses – there are no volume-sized typological studies that I am aware of on any larger group of Austronesian languages with regards to clause combining, and certainly none for the entire Austronesian language family.

Many introductions to linguistics offer brief definitions of notions such as coordination and subordination of clauses (e.g. Bloomfield 1933; Lyons 1968). In more depth, clause combining is dealt with in works specializing in various sub-areas of linguistics, such as Matthews (1981), Hopper & Traugott (1983), Haiman (1985), Givón (2001 [1990]), and Butler (2003), to mention but a few. Extensive descriptive grammars (especially those on major European languages) also often discuss topics pertaining to clause combining thoroughly and may touch upon issues with a greater theoretical scope (e.g. Quirk et al 1985 for English; Teleman et al 1999 for Swedish). Specifically on the topic of combined clauses, several important theoretically oriented works have seen the light of day in the last 40 years or so on issues such as coordination, subordination, adverbial constructions and clause chaining, both in the form of edited volumes (e.g. Haiman & Munro 1983; Brugman et al 1984; Shopen 1985 [2007]; Haiman & Thompson 1988; Haspelmath & König 1995; Korzen & Herslund 1998; Bybee & Noonan 2002; Vincent 2006; Dixon & Aikhenvald 2009; Bril 2010a), and in the form of monographs and reports (e.g. Dik 1968; Ohori 1992; Kortmann 1997; Diesssel 2004; Cristofaro 2005). Others have focused on specific semantic relations between clauses (a topic that will be dealt with in chapter 4), in particular, conditional relations (e.g. Traugott et al 1986; Athanasiadou & Dirven
1997), but also other relations, such as Couper-Kuhlen & Kortmann (2000), who treat co-varying relations broadly in their edited volume, which is often referred to as the CCCC volume (for cause, condition, concession and contrast). There is also an abundance of journal articles on the subject, some of which will be discussed below.

Of the different syntactic types of clause combining constructions, relative clauses (see e.g. Levi et al 1972; Keenan & Comrie 1977; Keenan 1985; Alexiadou et al 2000) and complement clauses (see e.g. Bolinger 1972; Givón 1980; Noonan 1985) have been studied in most detail, but there are also several studies specifically on coordination (see Dik 1968 for an early example and Haspelmath 2004a for a more recent one) and adverbial clauses (e.g. Haspelmath & König 1995; Kortmann 1997). The references made above are to works mostly of the functional school of linguistics to which the present study also belongs. Both coordination and subordination have of course been covered also in formal linguistics, prominent examples of which are Van Oirsouw (1987) (coordination) and Haumann (1997) (subordination).

The following sections constitute an overview of some of the major strands of interest in the literature concerned with clauses in combination. The first section deals with the distinction between coordination and subordination, the second one reviews some influential works that have served as sources both of facts and of inspiration in the writing of the present thesis, and the third one covers some earlier studies on Austronesian clause combining specifically.

## 2.1 Coordination and subordination

### 2.1.1 The notions of coordination and subordination

Traditionally, in the description of complex sentences, a distinction is made between coordinate and subordinate clauses, such that the former involve a principally symmetric relation between two (or more) clauses of equal syntactic status, no one being dependent on another, while the latter involve a part-whole relationship in which one clause, the subordinate, has a syntactic function in another, i.e., is grammatically dependent on it. Depending on the nature of their syntactic function, subordinate clauses are further divided into complement clauses (usually arguments to verbs), relative clauses (modifiers to nouns), and adverbial clauses (adverbials), of which the last mentioned are commonly – but not solely – associated with the temporal and co-varying relations studied here. The following examples illustrate the different types.
26. [Tom met Reese] and [they went to a restaurant].

27. He said [that he was hungry].

28. Tom was a man [who could eat a lot].

29. But Reese had to pay [because Tom was broke].

However, from a cross-linguistic perspective, or even within languages, it has proven difficult to make a clear-cut distinction between subordination and coordination (see e.g. Foley & Van Valin 1984; Lehmann 1988; Van Valin & LaPolla 1997; Givón 2001; Diessel 2001). Complement clauses and relative clauses may be somewhat easier to identify, the former normally being an obligatory constituent of a matrix clause and the latter an NP-internal modifier. A definite line between coordinate and adverbial clauses, on the other hand, seems especially difficult to draw. The presence of a subordinator (or subordinating conjunction) is often traditionally used as an indication of the subordination of a clause, but as pointed out by Kortmann (1997), the risk of circularity is apparent, since subordinators are often defined as items introducing subordinate clauses. Besides, in the face of cross-linguistic evidence, subordinators can hardly be considered a necessary means for identifying subordinate clauses. Also, in English we have examples such as that in (30).

30. I'm delighted he's coming.

As noted by Palmer (1987), without intonation in speech or punctuation in writing, the above sentence is ambiguous between one interpretation of it as a matrix clause with a complement clause and another as two independent clauses (the state of affairs of the second giving a reason for that in the first). Furthermore, it may not always be obvious that a certain relation marker is to be classified as a subordinator, coordinator, or something else, such as a coherence adverb or a verb particle. Consider for instance the Samoan example below.

31. Samoan

    alu  atu  'oe  suga  vika
go   DIR  you  girl  NAME

e    'aumai  tā  solo  ma  fasimoli
TMA  bring  our  towel  and  soap

'Go now, Vika, to bring us towels and soap.'
(Mosel & Hovdhaugen 1992, p 618)
The clause expressing purpose is introduced with *e*, which is a very general TMA marker used when the temporal or aspectual character of a state of affairs is indeterminate or constant (as with stative predicates), or is of no importance for what the speaker wants to convey. It is used in purpose clauses as purposes normally refer to yet unrealized states of affairs, and it is not really necessary to specify tense and aspect. The use of *e* in purpose clauses has given it the air of a purpose marker, and it is not obvious whether it should be regarded as a purposive subordinator (or coordinator) or whether the construction should be seen as involving the juxtaposition of clauses, and in that case, whether the purpose clause is coordinated to the initial clause (since it has the structure of an independent simplex clause) or subordinated to it (since the choice of TMA marker is restricted).

Also, consider the two following English examples, of which (32) is usually taken to involve adverbial subordination, and (33) coordination:

32. *I couldn’t sleep because it was cold.*  
33. *I couldn’t sleep for it was cold.*

If grammatical dependency is crucial for clausal subordination (as implied, for instance, by Lyons 1968), we can easily make a case for the string *because it was cold* being subordinate, since it is not grammatical on its own—but neither is the string *for it was cold*. In both cases, we need an appropriate context for these strings to be grammatical. It might be argued that *because it was cold* functions as an adverbial, while *for it was cold* does not. But defining an adverbial clause is not an easy task. The general definition is that an adverbial clause modifies a matrix verb or clause. But in what sense is the because-clause a modifier and the for-clause not? They are both equally optional, and they both certainly have the same meaning. A constituency test might show that *because* is in constituency with *it was cold*, indicated by the fact that *because it was cold* may appear initially as well as finally in (32), while *for* is not in constituency with *it was cold* for the opposite reason, and therefore, the latter two cannot share a single syntactic function. However, in light of this finding (as noted by for instance Quirk et al 1985), examples such as (34) below pose a problem.

34. *The traffic delayed us so that we arrived home late.*

Although *so that we arrived home late* is usually considered to be a subordinate clause, it cannot occur before the other clause in the complex construc-

---

3 See section 4.3, however, for Austronesian constructions denoting realized purpose.
tion above, being, in this respect, more like the coordinate for-clause than the
subordinate because-clause above.

Historically, the conjunction so is derived from the adverb so, which has
grammaticalized together with that into a relation marker. Modeled on the
example (34) above, the development may correspond schematically to the
outline in (35) below.

35. a.  *We arrived home late. The traffic delayed us so.*
b.  >  *The traffic delayed us so (that) we arrived home late.*

Traugott (1985) notes that the development from a backward pointing cohe-
rence adverb (as in (35a)) to a conjunction (as in (35b)) is quite common in
Indo-European languages for conditional and other markers. The obviously
gradual nature of grammaticalization over time is thus often reflected for
clause combining synchronically as a gradual difference over constructions
between subordination and coordination. This is, of course, one of the major
reasons why it is so difficult to pinpoint the individual characteristics of the
two.

2.1.2 Defining criteria for subordination

A number of different criteria by which subordination may be separated
from coordination have been proposed in the linguistic literature over the
years. Some have been put forth in isolation, while others have been pre-
sented in groups, explicitly or implicitly suggesting a prototype definition.
Taken by themselves, these criteria are neither necessary nor sufficient for a
definition of subordinate and coordinate clauses. In spite of their shortcom-
ings, however, taken together they provide useful guidelines when determin-
ing what is and what is not subordination. They are also useful for identify-
ing problematic cases.

In this section, some of the most frequently used subordination criteria
will be described, and some problems with each will be discussed. The crite-
ria are of various kinds and pertain to morpho-syntactic structure, syntactic
distribution within and between clauses, syntactic function, prosodic fea-
tures, and discourse-pragmatic function. The criteria have been applied to
the data in the present study – to the extent possible, given that the data
sources differ in detail (see further section 6.1.2).

Regarding morpho-syntax, one of the most widely used tests of subordi-
nation is possibility of pronominal cataphoric reference (e.g. Palmer 1987;
Haspelmath 1995). Pronouns normally have anaphoric reference, but in ini-
tial subordinate clauses, they may also refer cataphorically to a noun phrase
in the following main clause, as illustrated in the examples below.
36. a. Susan\textsubscript{i} stood up before she\textsubscript{i} began to sing. \textsc{anaphoric}
   b. *She\textsubscript{i} stood up before Susan\textsubscript{i} began to sing. \textsc{cataphoric}
   c. Before Susan\textsubscript{i} began to sing, she\textsubscript{i} stood up. \textsc{anaphoric}
   d. Before she\textsubscript{i} began to sing, Susan\textsubscript{i} stood up. \textsc{cataphoric}

The test works regardless of the syntactic function of the pronoun and the co-referent noun phrase, as may be verified by replacing (36a) by "I knew Susan\textsubscript{i} before she\textsubscript{i} was famous" or "I knew Susan\textsubscript{i} before Tim married her\textsubscript{i}" and similarly for the rest of the examples in (36). This test is often very useful as a criterion of subordination. But it is not flawless. Harris & Bates (2002) note that for at least some types of subordinate clauses, it is possible to achieve cataphoric reference of a pronoun in an initial main clause. Specifically, this is the case when an initial main clause is backgrounded in discourse (cf. Hopper & Thompson 1980), for instance, by using progressive aspect, as in (37c) below.

37. a. *He\textsubscript{i} threatened to leave when Billy\textsubscript{i} noticed that the computer had died.
   b. When he\textsubscript{i} threatened to leave, Billy\textsubscript{i} noticed that the computer had died.
   c. He\textsubscript{i} was threatening to leave when Billy\textsubscript{i} noticed that the computer had died.

In an enlightening paper on pronominalization in discourse, Bolinger (1979) also observes that a certain familiarity with the referent on the part of the interlocutors of a speech act may account for some cases of cataphoric pronouns, regardless of whether the pronoun is in a subordinate clause or not, as exemplified in (38).

38. a. Hey, Gus! Tell him to come in if you see Tom out there, will you?
   b. Hey, Gus! If you see him out there tell Tom to come in, will you?

Furthermore, cataphoric reference may sometimes be allowed between clearly coordinate clauses, as noted by, for instance, Bolinger (1979), Quirk et al (1985), and Matthiesen & Thompson (1988). Example (39) illustrates:

39. I found him hilarious, but Antoine had a serious and gifted side to him too.

Examples such as (39) would normally occur in a context in which Antoine had already been identified, and the pronoun of the first clause could thus be seen as anaphorically co-referring to an antecedent in a previous clause;
however, it is not entirely ruled out that (39) could also occur without Antoine having been mentioned first. Thus, it seems that discourse-pragmatic factors may sometimes override the morpho-syntactic restriction that cataphoric pronouns are only allowed in subordinate structures.

Another morpho-syntactic criterion that is often proposed is that subordinate clauses are reduced in their lexical or grammatical morphology in comparison with simplex clauses. (40) below is an example.

40. *She walked along the road cheerfully singing that old song.*

Lehmann (1988) speaks of desententialization, Bisang (1998) of minus asymmetry, and Van Valin (2005) of operator dependency in these cases, although none of them would define subordination by this criterion alone. Similarly, Cristofaro (2005) speaks of omission of verb morphology and arguments, although she defines subordination in a different way (see end of this section and section 2.2.3). Additionally, Haiman & Thompson (2005) discuss the formal reduction of clauses in connection with subordination, but they see it merely as one of a disparate set of phenomena that are traditionally subsumed under the label of subordination. They point out that many of these phenomena may also affect other types of clauses in discourse – not only those traditionally known as subordinate. For instance, a Mandarin simplex clause does not necessarily specify an agent if its identity is irrelevant or known from context, as the example below shows, and the same applies to many languages in the Austronesian family.

41. **Mandarin (Sino-Tibetan)**

\[
\text{zhè ge difang kěyì huábīng}
\]

\[
\text{this CL place can skate}
\]

'One can skate here.' (Haiman & Thompson 1984, p 513)

For these reasons, Haiman & Thompson (1984) suggest that the term subordination be abandoned entirely, and that the underlying phenomena be investigated separately.\(^4\)

In contrast to the criterion of reduction is the criterion Bisang (1998) calls plus asymmetry. This is the use of additional morphology on the verb of a linked clause that cannot be used in simplex clauses. These cases in-

---

\(^4\) Haiman & Thompson (1984) are often incorrectly cited as representing a prototype approach to the concept of subordination (see e.g. Diessel 2001; Cristofaro 2005). While they do present a number of properties that are often involved in what is traditionally known as subordination, they argue that these properties affect all clauses in a coherent discourse in various ways and should not be assumed to be characteristic of "subordinate" clauses alone.
clude, for instance, subjunctive moods, participle forms, and case markers or other types of nominalizing morphemes. In some cases, nominalizing morphemes may also be attached to the clause as a whole, normally to the last item of the clause. One problem with this criterion, as with the reduction criterion, is that, sometimes, additional morphemes commonly associated with subordinate clauses can be used in simplex clauses as well, as with the use of subjunctives for wishes, etc. (see also section 2.2.2). Another problem with additional morphemes is constructions with affixed coordinating conjunctions, as one can encounter in Latin, Vedic Sanskrit and Chipewyan. An example from the latter language illustrates.

42. Chipewyan (Athabaskan)

\[
\begin{array}{lll}
ts'enidher & ni-u & nathesti & k'i \\
1.SG.wake.up & PST-and & 1.SG.dream & EMPH
\end{array}
\]

'I woke up and it was only a dream.' or 'As I woke up it was only a dream.' (Cook 1992, p 468)

In such cases, it is not always clear how the enclitic conjunction affects the verb or clause to which it is attached. Is it coordinating the clauses or subordinating one to the other? Translating \(-u\) as 'and' may make the answer seem obvious: the clauses are coordinate. But enclitic conjunctions can be a first step in a grammaticalization chain leading to subordination, which is indicated by the second translation of the Chipewyan example. To illustrate further complications, we may look at the non-Austronesian languages of New Guinea. Many of these have different-subject (DS) markers on so called medial verbs that are tightly connected to the final verb of a complex sentence. Such DS markers often seem to originate from grammaticalized enclitic conjunctions (Haiman 1983). Another source from which DS markers on medial verbs can be derived is nominalizing markers, as in Daga (Murane 1974). Thus, there is a functional overlap between coordinating conjunctions and nominalizing morphemes, and since nominalization commonly entails subordination, some constructions are actually indeterminate between coordination or subordination by the criterion of plus asymmetry.

Turning now to criteria related to the syntactic distribution of clauses or the elements within clauses, a famous test for coordination of constituents in general is the so-called "coordinate structure constraint" (CSC) suggested by

---

5 In this respect, cf. Bril's (2010b) informational hierarchy (where coordinators commonly establish a contrast between topic and focus) and referential hierarchy (where deictic markers commonly establish a contrast between reference and assertion). Both hierarchies may lead to subordination. See further section 2.3; see also the concept of asymmetric coordination, discussed in section 6.2.8.
Ross (1986 [1967]). This constraint prohibits the extraction of elements out of either of two coordinate clauses, while extraction of elements out of matrix clauses with subordinate clauses, and even out of some types of subordinate clauses (notably complement clauses), is allowed. The coordinate structure constraint establishes (43) as involving coordination and (44) and (45) as involving subordination.

43. a. *I am reading the paper and she is talking to her sister.
   b. *What are you reading and she is talking to her sister?
   c. *Who are you reading the paper and she is talking to?

44. a. Tim dropped his bags in surprise when he saw me.
   b. What did Tim drop in surprise when he saw me?
   c. (?) Who did Tim drop his bags in surprise when he saw?

45. a. Bern told Sandy he had found a hundred dollar bill.
   b. Who did Bern tell he had found a hundred dollar bill?
   c. What did Bern tell Sandy he had found?

This criterion is mentioned in connection with subordinate clauses, for instance, by Haiman & Thompson (1984) and Haspelmath (1995). However, apart from being difficult to apply in languages one does not speak – since information of extractability is often scarce in descriptive grammars – this criterion also suffers from having a number of obvious counter examples. As has been extensively discussed in the literature (see Goldsmith 1985, Lakoff 1986, Na & Huck 1992, Johannessen 1998), some apparent types of coordinate constructions seem to allow extraction, as in (46).

46. a. Who did you come to the village and look for?
   b. How much can you drink and still stay sober?
   c. What did the committee meet and the chairperson decide?
   d. What do you mention and she screams?

The degree of freedom of clause order in complex sentences is another distributional criterion. Flexibility between final and initial position for a clause vis-à-vis another have often been taken to indicate its subordinate nature in relation to the other clause (see e.g. Haspelmath 1995; Matthiessen & Thompson 1988). However, this criterion is not really a criterion in itself, since it does not identify which of the two clauses is the subordinate. If one clause in a binary complex clause syntagm can occur both initially and finally, so can the other clause. This criterion is only usable when there is some formal asymmetry involved. For instance, it can separate coordinating from
subordinating conjunctions: If the conjunction can move together with one of the clauses without a change in overall meaning of the sentence, then it is a subordinating conjunction, and its clause is a subordinate one. Virtually the same as the criterion of flexibility in clause order is the criterion of tense iconicity. Tense iconicity is said to hold if the temporal interpretation of two clauses depends on their order in discourse. It is often held that subordinate clauses are not tense iconic (Croft 2001). This can of course only be tested in sentences that allow variable clause order (and involve an expression of temporally ordered states of affairs). Lehmann (1988) argues against the flexibility criterion by claiming that if a clause has a fixed position vis-à-vis the other clause in a binary complex clause syntagm, it is a sign of its subordination and integration into the matrix clause, rather than of its independence. However, this seems to be true of a specific set of subordinate and coordinate constructions, e.g. some types of complement clauses and some types of juxtaposed coordinate clauses, respectively (and in the latter case, only those which are not tense-iconic), but not so for other constructions. The criteria of flexible clause order and tense iconicity seem especially useful for identifying most types of adverbial clauses as subordinate and most types of syntactically or asyntactically coordinated clauses as coordinate. Consider the examples below.

47. a. They stopped dead in their tracks when they saw the tiger.
   b. When they saw the tiger, they stopped dead in their tracks.

48. a. They ran inside and he locked the door.
   b. ?And he locked the door, they ran inside.
   c. *He locked the door, they ran inside and.

49. a. They ran inside, he locked the door.
   b. ≠ He locked the door, they ran inside.

Another criterion associated with the linear order of linguistic elements is the criterion of discontinuity. While coordinate clauses are always continuous and non-overlapping, subordinate clauses may occur inside a discontinuous matrix clause (Haspelmath 1995), as illustrated below.

50. Cliff, while he was looking suspiciously at Graeme, sat down in the other chair.

Haiman & Thompson (1984) discuss this criterion in connection with grammatical incorporation, which they take to be one of the component parts of the traditional notion of subordination. However, this criterion of course has restricted applicability, since clauses do not necessarily have to be coordinated to be continuous. There are even the opposite counterexamples, i.e. cases of discontinuous clauses with clearly coordinated clauses inside them, as in (51).
51. *Those people – and of this I am quite certain – watch entirely too much television.*

Even if such examples may be described as meta-comments outside the syntactic structure, they do nevertheless reduce to some extent the applicability of the discontinuity criterion for establishing subordinate clauses.

Other distributional criteria with limited applicability are those pertaining to (i) specific internal word order in subordinate clauses (cf. any general grammar of Swedish or German), (ii) the ungrammaticality of main clauses without their subordinate clause (Palmer 1987), and (iii) the ungrammaticality of subordinate clauses without their main clause (Lyons 1968; Palmer 1987; Van Valin & LaPolla 1997). The latter criterion (iii) is, of course, one of the most widely used criteria for identifying subordinate clauses traditionally, and the problems associated with it have already been discussed in this section. Criteria (i) and (ii) quite obviously apply only to some languages and/or to some types of subordinate clauses.

As for syntactic function, one could say that a clause must function syntactically as a constituent of another clause to be considered subordinate. The term 'embedded' normally entails that the embedded clause has a syntactic function in a matrix clause (e.g. Foley & Van Valin 1984; Lehmann 1988). But as seen earlier in this chapter, the syntactic function of a clause is sometimes no less difficult to pinpoint than its general subordinate status (cf. examples (31)-(33) above).

Prosodic features vary with different kinds of clauses, and some have suggested that these could be used as criteria for subordination (e.g. Bloomfield 1933). Most scholars, however, agree that properties such as intonation and pausation have no direct connection with subordination in syntax, but that syntactic interpretation nevertheless benefits heavily from intonational patterns. Discussion along such lines can be found in Matthews (1981), Bolinger (1984), Chafe (1984), Haiman & Thompson (1984), Palmer (1987), and Lehmann (1988).

Finally, the notions of subordination and coordination have also been characterized in terms of discourse-pragmatic function. Cristofaro (2005), for instance, rejects the scalar relation between subordination and coordination, and suggests that subordination be defined in terms of pragmatic assertion. Cristofaro uses this term in the sense of Lambrecht (1994) to mean the information that the hearer is expected to learn as a result of hearing the sentence uttered, as opposed to the information that the speaker can assume the hearer already knows at the time the sentence is uttered. Cristofaro assumes that subordinate clauses lack assertive power, conveying only (presupposed) non-asserted information. Consider the example below.

52. *He was quite upset when I called.*
In this example, the state of affairs of being upset rather than that of calling is asserted. Since the second clause has no assertive power, it is taken to be subordinate. A similar definition of subordination is proposed by Langacker (1991), though he uses the term 'autonomous profile' instead of 'pragmatic assertion'. Another way of putting it is to say that subordinate clauses do not represent speech acts by themselves, i.e. they lack their own illocutionary force and therefore cannot be assertions. A number of tests may be used to check the assertive power and illocutionary force of clauses. For instance, only assertive clauses are open to challenge by explicit denial or questioning (Givón 1982; Cristofaro 2005), as in example (53).

53. a. *When I opened the door the cat got in.
   b. *It is not the case that when I opened the door the cat got in.
   c. *When I opened the door the cat got in, didn't it? / *didn't I?

Further, only non-asserted clauses can be (i) focused in a cleft construction, (ii) the response to a question, (iii) subject to alternative negation, and (iv) the focus of alternative questioning (Quirk et al. 1985), as illustrated by examples (54) and (55).

54. a. *He likes them because they are always helpful.
   b. *It is because they are always helpful that he likes them. (i)
   c. *Why does he like them? Because they are always helpful. (ii)
   d. *He didn't like them because they are always helpful but because they never complain. (iii)
   e. *Did he like them because they are always helpful or because they never complain? (iv)

55. a. *He likes them and they are always helpful.
   b. *It is and they are always helpful that he likes them. (i)
   c. *Why does he like them? And they are always helpful. (ii)
   d. *He didn't like them and they are always helpful but and they never complain. (iii)
   e. *Did he like them and they are always helpful or and they never complain? (iv)

---

6 Quirk et al. (1985), however, do not discuss these tests in connection with assertiveness.
Delimiting subordinate clauses to those lacking in assertiveness is admittedly an attractive approach. The pragmatic distinction between asserted and non-asserted information, as Cristofaro (2005) uses these terms, does indeed seem to provide a dichotomy that is both universal and absolute. Still, there are (at least) two problems involved.

The first problem is perhaps simply terminological. Although the assertiveness distinction is convenient and useful, the question is whether it is to be equated with the distinction between coordination and subordination in syntax. As Cristofaro acknowledges, equating the two means that clauses such as the bracketed ones in examples (56-58) below are excluded from the set of subordinate clauses by virtue of having their own assertive power (as indicated by the tag questions). Such clauses, however, are traditionally thought to be subordinate, and their structural resemblance (or even identity) to more "genuine" subordinate constructions cannot be denied.

56. a. Ben liked the idea [whereas his wife hated it], didn't she?
   b. You better take your coat, [because it is raining], isn't it?

57. a. I guess [John didn't come in], did he?
   b. I think [the car needs a tune up], doesn't it?

58. a. They are going to Alford, [which is near Skegness], isn't it?
   b. Chris did really well in his exams, [which was a big surprise], wasn't it?

The clauses in (56) are termed disjunct (as opposed to adjunct) adverbial clauses by Quirk et al (1985) (see also Lakoff (1984) for a discussion on assertive adverbial clauses); those in (57) form the object of so-called assertive predicates (Green 1976; Hooper 1976); and those in (58) are non-restrictive relative clauses (Huddleston 1965; Matthiessen & Thompson 1988; Lehmann 1988). Thus, if subordination is defined in terms of pragmatic assertion, it actually entails quite a novel use of the term. What must be kept in mind is that Cristofaro’s (2005) definition of subordination is a pragmatic definition that applies to tokens rather than types. That is, a certain type of construction that is potentially subordinate sometimes cannot be uniquely associated with a lack of illocutionary force. In actual fact, the same structural type may serve as the clausal format for propositions both with and without illocutionary force. This becomes clear when two identical examples can be interpreted either way depending on the context, as shown below (from Haiman & Thompson 1984, p 517).

59. a. They don’t beat us because they love us.
   b. They don’t beat us, because they love us.

In the (a) example, the because-clause has no assertive power (entailing that love is not the reason for the beating, although the beating is still inferred to take place for some other reason), and in the (b) example, both clauses have assertive power (entailing that there was both love and no beating).

The second problem is methodological. Even if the assertiveness distinction can be applied in principle, it is sometimes very difficult in practice to establish what parts of a complex sentence do or do not carry assertive power in a language that the researcher does not speak. Reference grammars seldom supply sufficient data on such issues. Cristofaro’s (2005) solution is to rely on the translations provided in grammars and apply assertiveness test on these. The assumption is that these translations preserve the communicative organization of sentences. This is, of course, not always the case. Consider the Mandarin example below.

60. Mandarin (Sino-Tibetan)

   ta hui  jiā kān qīnqi
   he return home see parents

‘He returned home to see his parents.’ or
‘He returned home and saw his parents.’ (Stassen 1985, p 73)

Without context or intonation cues, the second clause of the Mandarin sentence is ambiguous between a reading with assertive power (second translation) and one without (first translation). Although estimates are accurate in most cases, one cannot with certainty determine the assertive power of a clause in a language one does not speak simply from translations.

As evident from the discussion in this section, and as noted already by Brøndal (1937) (cited in Lehmann 1988, p 220, fn 1) and several times since, none of the various morphological, semantic and logical criteria proposed to determine subordination, and conversely coordination, is by itself sufficient for a definition of these notions. However, if chosen wisely, a number of them are certainly useful in prototype definitions. Indeed, in the present work, subordination vs. coordination is one of the structural parameters along which the Austronesian clause combining constructions are classified (see 6.1.2). In the classification, I have tried to use as many of the subordination criteria as the sources would allow to obtain the most consistent classification possible.
2. EARLIER RESEARCH

2.2 Parallel continua, iconicity and implicational hierarchies

2.2.1 Parallel continua: Lehmann's typology

One way of coming to terms with the problem of defining subordination and coordination, particularly within functional schools of linguistics, is to describe it as being a continuum or gradual cline of some sort. One of the most sophisticated proposals of a typology of clause linkage along these lines is presented by Lehmann (1988). He outlines a model consisting of six scalar parameters, representing semanto-syntactic dimensions along which the lexical and/or grammatical information in combined clauses may be either elaborated or compressed. These parameters are in principle mutually independent, although in practice they are highly interrelated. His model is illustrated below.

Table 2. Parallel continua in clause linkage

| i. hierarchical downgrading       | none: parataxis; strong: embedding |
| ii. syntactic level               | high: sentence; low: word          |
| iii. desententialization          | weak: clause; strong: noun         |
| iv. grammaticalization of main predicate | weak: lexical verb; strong: grammatical affix |
| v. interlacing                    | weak: separate clause properties; strong: overlapping clause properties |
| vi. explicitness of linking       | maximal: syndesis; minimal: asyndesis |

With regards to the scale of hierarchical downgrading, parataxis represents one endpoint, where there is no hierarchical relation between the clauses, while embedding represents the other, where one clause is a well-defined constituent within the other. In Lehmann's terms, parataxis is coordination of clauses, regardless of whether a linking device (syndesis) is present or not (asyndesis). The second parameter concerns the syntactic level at which one clause is integrated with another; it may be at the level of sentence (high) or

---

8 Traditionally, parataxis is often thought of as only including asyndetic coordination of elements (cf. Crystal 1992).
at any constituent level within another clause, down to individual words (low). Examples (61)-(63) show extreme and intermediate values on these two scales.

61.  I was trimming a boomerang, there you came up.

This is a simple juxtaposition of clauses, neither one integrated with the other. Thus, no hierarchical downgrading takes place, and the clauses are related at a very high level of syntax (at the level of text, according to Lehmann).

62.  Hittite

\[
\begin{align*}
&\text{nu kwit LUGALu-s tezzi nu apat iyami} \\
&\text{CONN what king-NOM says CONN that do.1.SG}
\end{align*}
\]

'And what the king says, that I do.' (Lehmann 1988, p 184)

The Hittite example constitutes a correlative construction. The initial relative clause \((\text{nu kwit LUGALu-s tezzi})\) is a dependent structure since it cannot stand on its own, but it is not clearly part of the matrix clause since the demonstrative \(\text{apat}\) occupies its slot there, with which it seems to occur in parallel, as it were. Thus, the relative clause is somewhat downgraded and related to the matrix clause at the level of sentence.

63.  Swedish

\[
\begin{align*}
&\text{han förmodade} \\
&\text{3.SG.MASC suppose.PST}
\end{align*}
\]

\[
\begin{align*}
&\text{att hon hade rätt} \\
&\text{COMP 3.SG.FEM have.PST right}
\end{align*}
\]

'He supposed that she was right.'

Here, the string \(\text{att hon hade rätt}\) is a complement clause and an obligatory constituent of the matrix clause with a well defined syntactic function, i.e. object of the verb \(\text{förmodade}\). Thus, it is tightly embedded at the level of the verb phrase.

Turning now to the scales of desententialization and grammaticalization of the main predicate, both of these concern the reduction of clausal properties: the former of subordinate clauses, and the latter of matrix clauses. However, reduction tends to take place differently in the two cases. With regards to desententialization, Lehmann (1988) enumerates illocutionary force, mood, tense, aspect, actants and circumsttants as common properties of
2. EARLIER RESEARCH

a full fledged clause. Constraints on, and/or loss of, an increasing number of these properties tend to accompany increasing subordination of a clause. Furthermore, these properties tend to be constrained/lost in a fixed order universally with illocutionary force first, followed by modal markers, tense/aspect markers, and arguments, respectively. As clausehood decreases, verbal derivation may take place, and the reduced clause may acquire the ability to combine with prepositions and case affixes. Thus, desententialization goes hand in hand with nominalization, and single verbal nouns constitute the endpoint of the scale. Regarding grammaticalization of the main predicate, this process normally affects the matrix clause in a different way. Reduction of the matrix clause tends to turn lexical verbs into modals, auxiliaries and finally grammatical affixes. Constructions expressing causative and desiderative meanings often develop along these lines. Reduction of clausal properties in the subordinate clause is illustrated in (64), where the complement clause has clear nominal properties (slightly more so in (b) than in (a)), and in (65) to an even greater extent. (65) is also an illustration of reduction in the matrix clause (as well as in the subordinate).

64. a. She objected to [ his constantly reading magazines ].
b. She objected to [ his constant reading of magazines ].

65. Italian

ho fatto [ prendere a mio figlio
have.1.SG made take.INF to my son

un'altra professione ]
another profession

'I had my son choose another profession.' (Lehmann 1988, p 201)

Lehmann's (1988) parameter of interlacing has to do with the sharing of properties between two linked clauses, such as identity in temporal or aspectual reference, or identity in actants or predicate. This may result in the total or partial exclusion from one of the clauses of the relevant property. Examples of this are gapping and non-finite or otherwise reduced verb forms. The sharing of actants may also result in the syntagmatic interweaving of constituents, and in some cases this results in the indeterminacy of whether one is dealing with a coordinate or subordinate construction, as shown by the alternative translation of the Lango example below.9

---

9 As with example (30) above, the presence of intonation or more explicit punctuation would of course resolve the structural ambiguity.
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

66. Lango (Nilo-Saharan)

dákó ódìò icó ókwálò gwènò
woman pressed.3.SG man stole.3.SG chicken

'The woman forced the man to steal the chicken.'
'The woman pressed the man. He stole the chicken.'
(Noonan 1985, p 106)

Finally, the linking device relating clauses (or on a higher level, stretches of text) may vary in explicitness and length, as described by Lehmann's explicitness of linking parameter. He notes, however, that this is often not a question of grammar but of discourse-pragmatics, the linking device being adjusted to the size of the entities linked. Thus, the elaborate initial phrase in 67 links the following sentence syntetically to the preceding context (excluded form the example) and contrasts with the asyndetically linked clauses in 68.

67. These things being done, Caesar had every reason to assume that Gaul was now pacified.
68. You're cold, take may jacket.

Lehmann's (1988) typology is interesting for a number of reasons. It constitutes one of the most systematic attempts to disentangle some of the major factors involved in the linking of clauses. The way he describes the cline multidimensionally between what he calls elaborated and compressed clause linkage with six scalar parameters may be a way of dealing with the fact that various complex constructions are often considered to be midway between coordination and subordination, yet in different ways. They may simply display intermediate values on different scales.

However, Lehmann (1988) is not quite explicit about the relation between subordination and his concept of compression in clause linkage. He seems to allude to the six parameters as relevant for his view of subordination as a prototype – he states: "in the course of the paper, subordination will emerge as a prototypical concept" (Lehmann 1988, p 182) – but it is clear that subordination and compression are not totally equivalent. Prototypicality obviously entails that any single parameter involved in the prototype is neither necessary nor sufficient by itself for identifying the prototype, but parameters should nevertheless be commonly associated with it. However, the last two of Lehmann's parameters are completely indifferent to the notions of coordination and subordination. Regarding interlacing (continuum v), coordinate clauses as well as subordinate clauses may exclude reference to shared elements, as for instance in examples (69) and (70) below.
69. *Bender used to mash the potatoes and eat them with butter.*
70. *Pat stood up in order to see better.*

And as pointed out by Lehmann himself, explicitness of linking (continuum vi) applies equally well to clauses that are clearly subordinate and those that are clearly coordinate. The syndetic and asyndetic coordination displayed in examples (67) and (68) above may be complemented by the syndetic and asyndetic subordination displayed in the examples below.

71. *The student bought a pile of books in order that the professor should consider him intelligent.*
72. *Mary knew Paul wasn't telling the truth.*

Also grammaticalization of main predicates (continuum iv) may be difficult to relate directly to subordination. Complete grammaticalization of the main predicate into a grammatical affix results in the former subordinate clause assuming the role of an independent simplex clause. Thus, while subordination is found between the poles of this continuum, it is absent at its extremes.  

2.2.2 **Iconicity in clause combining: the RRG approach**

Other important studies rejecting the binary opposition between coordination and subordination are those conducted within the framework of the theory of Role and Reference Grammar (RRG) (Foley & Van Valin 1984; Van Valin 1993; Van Valin & LaPolla 1997; Van Valin 2005). Clause combining has a developed position in RRG, and the focus of a number of studies has been exclusively on the linking of clauses (e.g. Van Valin 1984, 2000; Liong 2004; Ohori 1992, 1996, 2001, 2005). The RRG theory of clause combining crucially includes three components: the nexus, which concerns the specific syntactic relation between the combined units; the juncture, which concerns the level of syntax at which these units combine; and the iconic correlation between a syntactic interclausal hierarchy (based on nexus and juncture) and a semantic interclausal hierarchy. All these components will be discussed below.

---

10 It should be noted, however, that the complete grammaticalization of the main predicate does not necessarily entail that the former subordinate clause also assumes all the formal properties of an independent clause, although it is represented as a simplex clause. Retention of subordinate clause patterns in these cases may account for some instances of subordinate clause properties of independent simplex clauses (for examples and some discussion, see section 2.2.2).
Regarding nexus relations, the theory proposes a decomposition of the notion of subordination into two more basic notions: dependency and embedding. An embedded clause is part of a matrix clause, i.e. having a function in it, while a dependent clause cannot constitute a complete sentence, i.e. not being totally interpretable on its own. Coordinate clauses are neither embedded nor dependent, while subordinate clauses are both embedded and dependent. However, there are constructions, most famously the clause chaining constructions of non-Austronesian languages in New Guinea, in which one clause (or more) seems to be dependent but not embedded. Such clauses are termed cosubordinate,\(^{11}\) and an example appear as (73) below.

73. Selepet (Trans-New Guinean)

\[
\text{kawa} \ ari-mu \quad \text{kiap} \quad \text{ya} \quad \text{taka-op}
\]

\text{Name} \ \text{leave-3.SG.DS} \ \text{patrol.officer} \ \text{DEM} \ \text{arrive-3.SG.PST}

'Kawa left and that patrol officer arrived.' (Longacre 1985a, p 238)

Since \textit{*kawa arimu} is neither a grammatical clause nor temporally interpretable on its own, it is clearly dependent on the final clause. However, as is often pointed out in the literature on non-Austronesian New Guinean languages (e.g. Haiman 1980; Olson 1981; Reesink 1983; Roberts 1988), cosubordinate clauses are not embedded and differ from subordinate clauses in a number of ways. For one thing, they do not allow cataphoric pronominal reference (see section 2.1.2). Cf. examples (74a) – subordination – and (74b) – cosubordination – below from Amele.

74. Amele (Trans-New Guinean)

a. \[(uqa)_{i} \quad \text{sa} \quad \text{b} \quad \text{j-igi-an} \quad \text{nu} \quad \text{fred}_{i} \quad \text{ho-i-a} \]

\[\text{he} \ \text{food} \ \text{eat-3.SG.FUT} \ \text{PURP} \ \text{Name} \ \text{come-3.SG.HOD}\]

'Fred came to eat food.' (Roberts 1988, p 56)

b. \[(uqa)_{i} \quad \text{bi-bil-i} \quad \text{fred}_{i} \quad \text{je-i-a} \]

\[\text{he} \ \text{COOC-sit-3.SG.SS} \ \text{Name} \ \text{eat-3.SG.HOD}\]

'While he sat, Fred ate.' (Roberts 1988, p 57)

Since the first clause of example (74a) may include a pronoun cataphorically coreferent with \textit{Fred} in the second clause, it may be considered subordinate to the second. But since a cataphoric coreference interpretation is impossible

\(^{11}\) The notion of cosubordination was originally developed by Olson (1981) for Barai (Papuan).
for the pronoun in the first clause of example (74b), the construction is clearly different from the subordinate one in (74a), and more like a coordinate construction in this respect. Further, as noted by, for instance, MacDonald (1990) for Tauya (Trans-New Guinean), in many chaining languages clause chains may grow fairly long, stretching over whole paragraphs, which further emphasizes the differences between these constructions and ones involving subordinate clauses, at least as they are traditionally known.

A clause can be dependent on another clause (i) for the interpretation of one or several of its features, e.g. tense, and (ii) because it cannot stand on its own as a grammatical simplex clause. In the former case, RRG speaks of operator dependency, and in the latter, of structural dependency (Van Valin & LaPolla 1997). While subordinate clauses display structural but not operator dependence, cosubordinate clauses display both (cf. examples (73) and (74) above). So, in terms of syntactic tightness, RRG takes cosubordination to involve tighter linkage than subordination. The situation may be summarized as in the table below.

Table 3. Types of nexus relations in RRG

<table>
<thead>
<tr>
<th>TYPES OF CLAUSE LINKAGE</th>
<th>Coordination</th>
<th>Subordination</th>
<th>Cosubordination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operator dependency</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Structural dependency</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Embeddedness</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

As for juncture, RRG posits that clauses are universally made up of units in three layers: the nucleus (corresponding to the predicate without its arguments), the core (the predicate with its arguments), and the clause as a whole (the predicate, arguments, and temporal and locative adjuncts). The structure is outlined in example (75).

75. John ate his lunch in the library.

```
  nucleus

    core

  clause
```

Each of the three types of nexus relations in Table 3 (coordination, subordination, and cosubordination) may occur at each layer of the clause so that we

12 In Van Valin (1984), operator dependency and structural dependency are termed grammatical category dependency and distributional dependency, respectively.
end up with nine possible types of complex sentences: clausal coordination, subordination and cosubordination; core coordination, subordination and cosubordination; and nuclear coordination, subordination and cosubordination. These nine juncture-nexus types have long been standard within RRG, but based on examples such as (76) below, Van Valin (2005) suggests that coordination and subordination may also occur at the level of sentence. The detached topic of each clause is taken to be outside of the clause structure and to belong to the sentence structure.

76.  As for Sam, Mary saw him last week, and as for Paul, I saw him yesterday.

Ranked in terms of tightness of linkage, the syntactic hierarchy of interclausal relations presented in Table 4 below emerges.

Table 4. RRG syntactic hierarchy of interclausal relations

<table>
<thead>
<tr>
<th>TIGHTEST SYNTACTIC INTEGRATION BETWEEN UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>nuclear cosubordination</td>
</tr>
<tr>
<td>nuclear subordination</td>
</tr>
<tr>
<td>nuclear coordination</td>
</tr>
<tr>
<td>core cosubordination</td>
</tr>
<tr>
<td>core subordination</td>
</tr>
<tr>
<td>core coordination</td>
</tr>
<tr>
<td>clausal cosubordination</td>
</tr>
<tr>
<td>clausal subordination</td>
</tr>
<tr>
<td>clausal coordination</td>
</tr>
<tr>
<td>sentential subordination</td>
</tr>
<tr>
<td>sentential coordination</td>
</tr>
</tbody>
</table>

| LOOSEST SYNTACTIC INTEGRATION BETWEEN UNITS |

Also semantically, two linked states of affairs may be construed as more or less tightly connected, and as mentioned initially in this section, the syntactic hierarchy of interclausal relations interacts iconically with a semantic hierarchy of interclausal relations. Van Valin (2005) enumerates several points along such a semantic hierarchy. Some of the most relevant ones for the present purposes are presented here (from tight to loose): (1) causative [i]; one state of affairs brings about another directly, the states of affairs being perceived of as one sequence, e.g. Harold pushed open the door; (2) means; one state of affairs constitutes the means by which another is carried out, e.g. Sam opened the box by slicing it with a knife; (3) purposive; one state of
affairs is carried out with the intent to realize another, e.g. *John went to the store to buy milk*; (4) causative [ii]; one state of affairs brings about another, the states of affairs being perceived of as distinct, e.g. *Chris forced Dana to leave the party*; (5) circumstance; one state of affairs specifies the temporal or spatial parameters of another, e.g. *Kim saw Pat after she arrived at the party*; (6) reason; one state of affairs constitutes the cause of another, e.g. *The baby cried because she was hungry*; (7) conditional; one state of affairs is dependent for its realization on another hypothetical state of affairs, e.g. *If it doesn’t rain we will be able to have a picnic*; (8) concessive; one state of affairs is unexpected, given the occurrence of another, e.g. *Bill made it work, even though it was snowing heavily*; (9) simultaneous; one state of affairs co-occurs temporally with another, e.g. *Max danced with Sarah while Susan played the piano*; (10) sequence; one state of affairs takes place after another, with or without temporal overlap, e.g. *Juan told a joke, and then Carlos entered the room*; (11) unordered; the temporal relation between two states of affairs is unexpressed, e.g. *Tyrone likes apples and Don likes oranges*.

Table 5. RRG semantic hierarchy of interclausal relations

<table>
<thead>
<tr>
<th>TIGHTEST SEMANTIC INTEGRATION BETWEEN UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>causative [i]</td>
</tr>
<tr>
<td>means</td>
</tr>
<tr>
<td>purposive</td>
</tr>
<tr>
<td>causative [ii]</td>
</tr>
<tr>
<td>circumstance</td>
</tr>
<tr>
<td>reason</td>
</tr>
<tr>
<td>conditional</td>
</tr>
<tr>
<td>concession</td>
</tr>
<tr>
<td>simultaneous</td>
</tr>
<tr>
<td>sequence</td>
</tr>
<tr>
<td>unordered</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOOSEST SEMANTIC INTEGRATION BETWEEN UNITS</th>
</tr>
</thead>
</table>

The interaction between the syntactic and the semantic hierarchies of interclausal relations is of a complex nature. There is certainly no one-to-one iconic correspondence between the two hierarchies, since a given syntactic type may express more than one semantic relation and a given semantic relation may be instantiated by more than one syntactic type in a certain language. For instance, in English, clausal subordination may be used to express both reason and concession (77), while conversely, purposive relations
may be expressed by both clausal cosubordination and clausal subordination (78).

77. a. Ray laughed because Sandy looked funny. REASON
    b. Ron laughed even though it wasn't funny. CONCESSION

78. a. I wrote it down to remember it. CLAUSAL COSUBORDINATION
    b. I wrote it down so that I would remember it. CLAUSAL SUBORDINATION

However, according to Van Valin & LaPolla (1997), it is always the case that the tightest syntactic linkage realizing a particular semantic relation in a language is higher than or at least as high on the syntactic hierarchy as the tightest syntactic linkage realizing a semantic relation lower on the semantic hierarchy. In this sense, there is iconicity between the two hierarchies. Hence, the tightest linkage type found in a language should always include causative [i] relations (and possibly also others). Likewise, the tightest syntactic linkage realizing, for instance, purpose relations should always be tighter than the tightest syntactic linkage realizing, for instance, conditional relations.

Some aspects of the RRG approach to clause combining are akin to the typology proposed by Lehmann (1988) (section 2.2.1). Embedding, for instance, as understood by RRG, parallels Lehmann's parameter of syntactic downgrading, while the RRG concept of levels of juncture is very similar to Lehmann's parameter of syntactic level. Further, dependency in RRG terms overlaps, at least partially, with both desententialization and interlacing in Lehmann's typology. One difference, however, is that RRG tends to operate with discrete categories in syntax, although hierarchically ordered, whereas Lehmann emphasizes the continuous nature of his parameters. The wide range of constructions found in the world's languages, and the fine-grained differences between them, speak in favor of the continuum approach.13

Another important difference between RRG and Lehmann's typology is their opposing views on the role of syntactic-semantic iconicity in clause linkage. In contrast to RRG on this specific issue, Lehmann denies that semantic relations play a role in determining the syntactic types of linkage. Along the same lines, Harris (1988, 1989) argues that there is no one-to-one correspondence between formal and semantic categories in clause combining, and further that the unmarked construction for a certain function varies diachronically and may be abandoned synchronically for pragmatic, stylistic

---

13 Ohori (1992), within the RRG framework, does in fact acknowledge that adjunct clauses exhibit a wide variety of dependencies that differ in terms of tightness, thus alluding to a continuous dependency scale.
or syntactic reasons. However, the RRG approach does not, as we have seen, deny that different morpho-syntactic types may express the same semantic relation in a language. It is only claimed that semantic relations stand in an iconic relationship to the *tightest* morpho-syntactic type available for each semantic relation. It may well be that examples used to underpin the views of scholars such as Lehmann and Harris on this issue can be incorporated into the kind of iconic relation between syntax and semantics held by RRG to be operating in clause linkage. Other studies affirming the relevance of iconicity in clause combining are Silverstein (1976), Givon (1980, 1985), Palmer (1987), Kortmann (1997), and Cristofaro (2005). We shall have reason to return to this issue from the perspective of the Austronesian languages in section 7.1.

It may also be observed that RRG takes the concept of 'structural dependency' as unproblematic. The dependency of subordinate clauses is often taken for granted in the traditional literature. However, there are problems of both a theoretical and practical nature involved. As Haiman (1985) points out, if structures such as (79) below, consisting of a relation marker and a simplex clause, are considered dependent because they are unacceptable on their own, then structures such as (80), likewise consisting of a relation marker and a structurally simplex clause, must be considered dependent since they are equally unacceptable on their own.

79. *because/if/after/when/although the fish stunk up the house
80. *the fish stunk up the house and

Thus, subordination is sometimes inseparable from coordination in terms of structural dependence. There is also a practical problem in distinguishing between dependent and independent clauses by way of structural properties associated with one or the other, as clauses with typically dependent clause properties may be used as simplex independent clauses in many languages, with varying degrees of conventionalization. This phenomenon has received attention from Evans (2007), who calls it insubordination. The examples from Italian and Diyari below may serve to illustrate.

---

14 Palmer (1987) does not connect the difference he observes between relations of unintended result and relations of intended result (i.e. purpose) explicitly to iconicity. However, he shows that the former are often more coordinate-like than the latter, which are more subordinate-like. And this would indeed seem to be an iconic pattern.
81. Italian

\[ \text{che venga domani} \]
\[ \text{COMP come.3.SG.SUB tomorrow} \]

'(It is possible/likely / I hope/believe) that he will come tomorrow.'
(Evans 2007, p 379)

82. Diyari (Pama-Nyungan)

\[ \text{nhulu-ka kinhthala-li yinanha matha-yathi} \]
\[ 3.SG.FEM.ERG-token dog-ERG you.ACC bite-SUB \]

'This dog might bite you.' (Evans 2007, p 380)

Complementizers and subjunctive verb forms are normally used in subordinate clauses but can also be used in Italian simplex clauses to express a possibility or wish, etc., as in (81). Likewise, inflections commonly used in subordinate clauses in Australian languages to express negative purpose (i.e. 'in order not') can also be used in Diyari simplex clauses to express possibility, as in (82). Thus, it seems that while some structural properties may be typically associated with dependent or independent clauses, the dividing line is neither strict nor stable.

2.2.3 Typological hierarchies and functional motivations: Cristofaro (2005) on subordination

Cristofaro (2005) investigates the typology of subordinate clauses in the languages of the world. She defines her subject of study functionally by the notion of assertive power, equating subordinate clauses with clauses that do not make assertions of their own (cf. section 2.1.2 above). With this definition of subordination, Cristofaro proceeds to examine how a number of morpho-syntactic properties correlate with various kinds of subordinate clauses. The properties she takes into consideration are elimination or alternation of TMA distinctions, elimination or alternation of agreement distinctions on the verb, use of case markers on the verb, and omission or altered coding of verb arguments. Since the set of TMA distinctions, for instance, varies from language to language, Cristofaro uses simplex declarative clauses as the standard of measure for each language. The distribution of the properties is mapped unto classes of complement clauses, relative clauses and adverbial clauses. Of primary concern for the present study are the results she presents for adverbial clauses. Comparing six different semantic relations, she concludes that they may be ordered in two general implicational hierarchies: one
based on the form of the verb (83), and one based on the coding of arguments (84).\textsuperscript{15}

83. purpose &gt; before, after, when &gt; reason, condition \hspace{2cm} \textsc{verb form}
84. purpose &gt; before, after, when, reason, condition \hspace{2cm} \textsc{arguments}

Thus, if the subordinate clause involved in the expression of a certain relation in the hierarchy deviates from a declarative simplex clause with respect to the properties considered, then subordinate clauses involved in the expression of all relations to the left tend to do so as well.

Cristofaro continues on to discuss how these hierarchies, established by morpho-syntactic criteria, correlate with semantic factors. She finds that the relations to the left on the hierarchies tend to share semantic features between the states of affairs to a larger extent than the relations on the right. Thereby she confirms the role of iconicity in clause combining proposed by others (e.g. Foley & Van Valin 1984; Van Valin & LaPolla 1997; Van Valin 2005; Givon 1980, 1985). In addition, she incorporates iconicity into a more encompassing model of functional motivations shaping the morpho-syntax of clause linkage. She identifies four functional motivations underlying the syntax-semantics correlation: syntagmatic economy, iconicity of independence, iconicity of distance, and the cognitive distinction between processes and things. Subordinate clauses in relations further to the left on the hierarchies tend to include semantic components which are recoverable from context (such as reference to participants or temporal setting). The fact that such features do not tend to be marked morpho-syntactically is accounted for by syntagmatic economy. As for iconicity, Cristofaro follows Newmeyer (1992) and separates iconicity of independence, i.e. the correspondence between formal dependency (degree of grammatical integrity) and conceptual dependency (degree of semantic integrity), from iconicity of distance, i.e. the correspondence between formal distance (in terms of number and type of morphemes) and conceptual distance (in terms of shared semantic features) (see further section 7.1.1). Subordinate clauses expressing relations further to the left on the hierarchies tend both to have reduced semantic integrity and share more semantic features with the main clause. For instance, Cristofaro finds that purpose clauses cross-linguistically are often formally reduced compared to simplex clauses (less independent) and normally have the same actors as the main clause, often absent in the purpose clause (less distant). Finally, the greater tendency for clauses expressing relations to the left on the hierarchies to be construed as unitary wholes, just like things, accounts

\textsuperscript{15} Cristofaro (2005) also presents separate hierarchies for each of the morpho-syntactic properties involved. The hierarchies vary only slightly.
for their increased ability to attract nominal features, such as case marking on the verb or coding of arguments as possessors.

### 2.3 Clause combining in Austronesian languages

Most of the studies on clause combining in Austronesian languages have been conducted on individual languages. There are, however, two fields of study that have produced works on the combining of clauses in Austronesian languages: the study of serial verbs, and the study of sentence and discourse structure within the theory of tagmemics. These will each be presented in this section, as well as some other studies.

Let us begin with tagmemics and clause combining. Tagmemics was intended by Kenneth Pike to be a unified theory of human behavior, with linguistics featuring as its most prominent part (Pike 1954-60, 1967). In the 1960s, Robert Longacre ran a project entitled "Discourse, paragraph and sentence structure in selected Philippine languages", which was based on the tagmemic theory. The project involved 25 Philippine languages, and resulted in a three-volume report to the U.S. Office of Education (Longacre 1968). Volume one and two were also later republished as Longacre (1970). In volume two, which is of primary concern for the present study, a wide range of complex clause constructions in Philippine languages are described and exemplified from the viewpoint of tagmemic theory. As an interesting result, Longacre concludes that there is no simple one-to-one match between types of complex constructions (use of conjunction, juxtaposition, merging of clauses, etc.) and semantic relations (result, condition, etc.) across the languages examined. This viewpoint is a precursor to Harris (1988, 1989) and Lehmann's (1988) views and argues against an iconic relation between form and meaning in clause combining (see further section 7.1). However, Longacre nevertheless finds strong ties between groups of categories, if these groups are based on neither entirely semantic nor entirely formal criteria. He presents a so-called neighborhood scheme of sentences reminiscent of a cognitive map. For instance, he groups together the categories of extent,

---


17 Volume three of the original report contained text samples.

18 Longacre (1970) does not himself mention iconicity, nor does he make any explicit remarks about whether his study would support or negate the validity of the notion in clause combining.
sequence, time, simultaneity and circumstance-result. All these categories share the component of temporality. The group has ties to several other groups, two of which also involve semantically based categories: the one consisting of conditional and concessive relations (sharing the component of implicational meaning), and the other consisting of result, warning, remark and question-answer relations (apparently sharing structural traits rather than semantic ones). The examples below are English equivalents given by Longacre (1970, pp 222-5) of sentences featured in some of the categories.

85. a. *First I tried to find a job in London, then I looked for work in Southampton.* SEQUENCE

b. *While they were there, she became pregnant.* SIMULTANEITY

86. a. *If he comes, then I won't go.* CONDITION

b. *Even if he comes still I won't go.* CONCESSION

87. a. *A person is only guessing when it will rain; after all is there any sign of it?* REMARK

b. *Oh, what did she do, she continuously wept.* QUESTION-ANSWER

c. *We shouldn't let our torches out; otherwise we won't be able to see the trail when we go home.* WARNING

Because of the mix of formal and semantic criteria apparently defining the categories and the groups they belong to, it is quite difficult to assess to what extent Longacre's study has any bearing on iconicity in Austronesian clause combining. But it is nonetheless an interesting survey of clause combining constructions in several Philippine languages. A number of articles, as well as a few volume-size works, describing sentence structures found in various Philippine languages from the viewpoint of tagmemics were published in the wake of Longacre's Philippine project (e.g. Reid 1970; Ashley & Ashley 1971; Ballard et al 1971a, 1971b; Elkins 1971; Mayfield 1972; Hall 1973; Walton 1975; Shand 1976; Maryott 1979[19]). The tagmemic model was also used to describe sentence structures in some other Austronesian languages,

[19] Maryott's (1979) article treats sentences in Sangir, which is not a Philippine language proper but belongs to the Sulawesi subgroup of Western Malayo-Polynesian, although it is spoken both in northern Sulawesi and on the Philippine islands of Balut and Sarangani south of Mindanao. Sangir was also included in Longacre's project.
Let us now turn to the subject of serial verb constructions, which is a related area of research that has merited some interest, especially for the Oceanic branch of Austronesian languages. Serial verbs are known to express two subcomponents of an overall state of affairs. Cf. the Paamese example below.

88. Paamese (Southern Oceanic)

\[\text{kaiko ko-muasi-nau nau-vaa netano}\]

2.SG 2.SG.REAL-hit-1.SG 1.SG-go down

'You hit me down.' (Crowley 2002a, p 60)

Of relevance to this study is the fact that the relation between two verbs in serialization may sometimes be of a co-varying character. Syntactically, the relation between the verbs may be seen as a very tight and reduced form of combined clauses. A more literal rendering of the Paamese clause above might be something along the lines of 'You hit me (so that) I went down', which spells out the co-varying relation between the verbs.

However, serial verb constructions are grammaticalized to such a degree that it normally makes no sense to speak of combined clauses; they appear within the single intonation curve of a simplex clause (this property is often included in the definition of the phenomenon), and often they have assumed a grammatical function, e.g. marking case roles or directions/locations, or indicating tense or aspect (cf. Givon 1991). Many researchers also favor a one clause interpretation (e.g. Schachter 1974; Schiller 1990). Johnston (1978) however, for the Nakanai language (Oceanic), argues in favor of a multiple clause analysis of serial verbs, since all Nakanai verbs that are capable of entering into serial constructions have maintained their capacity to appear as main verbs. This situation does of course not hold for all languages, because serial verb constructions are simply not a homogeneous group. There are several attempts at typologizing serial verb constructions. For Oceanic languages, the most obvious one is Crowley (2002a) in a very readable volume on serial verbs in Oceanic languages in general. Following Foley & Olson (1985), Crowley distinguishes between core serialization and nuclear serialization, the former applying to cases in which two serialized verbs within the core may be said to constitute two separate nuclei, the latter to cases in which two serialized verbs may be said to constitute one single nucleus.20 Core and nuclear serialization may be diagnosed by various mor-

---

20 The terms 'core' and 'nucleus' are used in the RRG sense (see section 2.2.2 above).
pho-syntactic criteria (though not universal ones). Consider the examples below from Barai, an SOV language of Papua New Guinea.

89. Barai (Trans-New Guinean)

   a. fu fi fase isoe
      he  sit  letter  write
   'He sat down and wrote a letter.' (Foley & Olson 1985, p 38)

   b. fu fase fi isoe
      he letter  sit  write
   'He sat writing a letter.' (Foley & Olson 1985, p 38)

The first construction is an example of core serialization because of the word order SVOV, where the object is clearly an object of the second verb only. Thus, the verbs constitute separate nuclei. The second construction is an example of nuclear serialization, since the object seems to be an object of the complex predicate 'sit writing', manifested by both verbs together (SOVV). The intransitive fi, 'sit', could hardly take an object without the help of isoe, 'write'. Thus, the verbs have merged into one nucleus. Example (89a) approaches a two clause construction, while example (89b) approaches a one clause construction. This is also supported by Barai prosodic patterns, as (89a) has a two-peak intonation curve, whereas (89b) has a one-peak curve (Foley & Olson 1985). 21 Others who have utilized the distinction between nuclear and core serialization for Oceanic languages include Crowley (1987) for Paamese, Earley (1993) for Lewo, and Sperlich (1993) for Namakir, all closely related languages of the North-Central Vanuatu subgroup. Earley (1993) also makes an interesting comparison with Lehmann's (1988) clause linkage typology (see section 2.2.1) and predicts that since serial verbs express subcomponents of an overall state of affairs, there is pressure towards syntactic compression. Thus, core layer serialization would be expected to move in the direction of nuclear layer serialization.

Another typological parameter explored by some is the nature of the relationships that hold between the actants associated with each of the verbs in serial verb constructions, e.g. whether they are identical or differ between the verbs (see e.g. Crowley (2002a) for Oceanic in general, Crowley (1987) for Paamese, Sperlich (1993) for Namakir, and Bradshaw (1993) for Num-bami and Yabem). Other works on Oceanic serial verb constructions include

21 Core serialization would thus fall outside the widely used definition of serial verbs by which they must occur within one continuous intonation curve (see e.g. Lehmann 1988 and Crowley 2002a, although the latter has some concerns about the practicality of using intonation as a defining criterion for verb serialization).
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

Durie (1988) and Hamel (1993), both of which focus on aspects of serial verbs assuming the character of prepositions.

Apart from tagmemic studies and studies on serial verbs, there are very few studies in clause combining for a general group of Austronesian languages, and those that do exist are of only slight relevance to the topic of temporal and co-varying relations. There are some works partly or wholly on general discourse patterns in groups of Western Malayo-Polynesian languages (Longacre 1970, 1971; Levinsohn 1991). Aldrige (2004) investigates relative clauses, but though the name of her paper may imply otherwise ("Internally headed relative clauses in Austronesian languages"), her investigation only includes data from two languages: Seediq (Formosan) and Tagalog (Meso Philippine). Reid & Liao's (2004) paper presenting a typological outline of Philippine languages also has a section on relative clauses. Bril (2004) and Moyse-Faurie & Lynch (2004), in a collected volume on coordination edited by Haspelmath (2004a), provide interesting discussions on coordination in Oceanic, although only parts of their papers present information on the coordination of clauses. Finally, Bril (2010b), in a recent volume on clause linking and clause hierarchy (edited by her), analyzes clause linking strategies in Oceanic languages. She discusses the structuring of clauses as subordinate by means of (i) the distinction between presupposed and asserted information (informational hierarchy) and (ii) the distinction between referential and asserted information (referential hierarchy). In this context, Bril presents and exemplifies the use of several markers in the Oceanic languages that function as indicators of topical (presupposed), focal (asserted), and referential information. It is especially interesting to note her examples (from both Oceanic and Western Malayo-Polynesian languages) of coordinators used to identify the dividing line between topical and focal information in sentences. It may seem peculiar for coordinators to have this function since topical information is often coded as subordinate constituents. The construction is also noted in the present study – here called asymmetric coordination – and will be discussed in more detail in section 6.2.2.

For individual Austronesian languages, there are several works on clause combining, most of a descriptive nature. Of prime interest for this study are some works on adverbial clauses (e.g. Clausen 1996 and Good 1989 for Ilokano (Philippine) and Kusaiean (Micronesian), respectively), and works on reference tracking across successive clauses (e.g. Lynch 1983 and Bradshaw 1999 for Lenakel (South Vanuatu) and Yaben and Numbami (North New Guinean) respectively). The articles by Ross (1987, 1993) on the influences of non-Austronesian neighboring languages on combined clause patterns in Takia (North New Guinean) are also very interesting. Of some, but less direct, importance are studies of relative clauses in Austronesian languages (e.g. Keenan 1972 on Malagasy (Borneo); Yeoh 1977 on Malaysian (Malayic); Bauer 1982 on Maori (Polynesian); Carpenter 1994 on Luang (Central Malayo-Polynesian); and Donohue 1996 on Tukangbesi (Sulawesi)), as
well as on complement clauses (e.g. Coorman 1984 on Chamorro (Western Malayo-Polynesian); Muller-Gotama 1995 on Malaysian (Malayic)). Of some significance as well are the historically oriented papers on the development of verbs into coordinators or complementizers found in many Austro-Polynesian languages (see Klamer 1999, 2000 on Kambera, Tukangbesi and Buru (of Indonesia), and Klinken 2000 on Tetun (of East Timor)).
3 Austronesian languages and sources of data

In this chapter, some background information on the Austronesian language family will be given, as well as a description of the language sample and the database used. The Austronesian languages will be described from a genealogical and typological viewpoint, for orientation and in order to facilitate understanding of some of the discussion to follow.

3.1 The Austronesian language family

The Austronesian language family is one of the largest language families in the world, whether measured by geographical distribution, number of languages or number of speakers. It is spoken by an estimated 374 million people\(^{22}\) distributed from Taiwan and Hawai‘i in the north to New Zealand in the south, and from Madagascar in the west to Easter Island in the east. The last two locations are separated by a distance corresponding to almost 60 percent of the circumference of the world. Before the great European exodus in modern history that spread the Indo-European languages to almost every part of the globe, Austronesian was the world's most widely distributed language family. It consists of around 1,100 languages (plus/minus about 200 depending on one's criteria for defining a distinct language), which is roughly 20 percent of the number of languages spoken in the world today. Nearly half (40 percent) of the Austronesian languages are spoken in the Pacific Islands, while the other half are spoken in Southeast Asia and Madagascar. However, the overwhelming majority of the Austronesian

---

\(^{22}\) This figure is based on population statistics from the CIA World Factbook (2005 estimates) and the Ethnologue database (15th edition) for the relevant regions (CIA 2005, Gordon 2005). The figure is much higher than the figures commonly stated elsewhere in the literature – 270 million in Bellwood et al (1995) and Adelaar & Himmelmann (2005), and approximately 276 million in Tryon (1995). These figures, however, seem to be based on older statistics. An estimate primarily based on population figures from 1990 (Lahmeyer 1999/2006) matches these lower figures, amounting to approximately 272 million people. For the Austronesian speaking population to reach 374 million in 15 years (i.e. my figure for 2005), a yearly population growth of 2.15% would have been needed, which is not at all improbable for the region.
speaking population – 99 percent – lives in Southeast Asia and Madagascar. Some of the Austronesian languages in this region (such as Javanese, Indonesian, Tagalog, Malagasy and others) have millions of speakers, while the Austronesian languages of the Pacific are comparatively small, most of them spoken by less than a few thousand people each.

Since this study is partly diachronic in nature and will be touching upon issues of genealogical and geographic distribution, the first part of this chapter is intended to give an overview of issues in Austronesian internal sub-grouping. And as a background to the more specific discussion of the grammar of TC clause combining in Austronesian languages, the second part of this chapter will outline some significant typological properties that are associated with Austronesian languages.

3.1.1 Genealogical relationships

The relationship between Malay and the Polynesian languages was first recognized by Hadrian Reland in 1708, and the Austronesian languages were established as a family in 1784 by Lorenzo Hervas y Panduro (Lynch et al 2002). Based on geographical distribution, the name initially used for this family was Malayo-Polynesian, a term that was probably first used in print by Franz Bopp (1841) (Ross 1996). As it became clear that the indigenous languages of Taiwan also belonged to this family, the term Austronesian was established, while Malayo-Polynesian was reserved for one of the first order subgroups. However, the term Malayo-Polynesian can still sometimes be found in use to refer to the entire language family.

Various suggestions of internal subgrouping of the Austronesian family have been proposed through the years. The most widely accepted genealogical classification of the major Austronesian subgroups is based on suggestions by Robert Blust (e.g. 1977, 1980, 1999, 2009; the details of which have been slightly modified over the years). According to the most recent version (Blust 2009), the 20 or so known Austronesian languages of indigenous Taiwan (of which 15 are living languages) – referred to as Formosan languages – are sorted into nine primary subgroups of Austronesian, while all other Austronesian languages belong to a tenth primary subgroup: the Malayo-Polynesian (MP) languages. The MP subgroup in turn is divided into the Western Malayo-Polynesian (WMP) group and the Central-Eastern Malayo-Polynesian (CEMP) group. The former consists of all the languages of the Philippines, Malaysia, and Madagascar, and most of the languages of Indonesia. Also included are the languages spoken on Palau and in most of the Mariana islands (including Guam) in Micronesia, as well as the languages of minorities in China (the Tsat on Hainan Island), Vietnam and Cambodia (the Champa), and Thailand and Burma (the Moklen-Moken of the Mergui Archipelago in the Andaman Sea). The CEMP group is divided into the
Central Malayo-Polynesian (CMP) languages and the East Malayo-Polynesian (EMP) languages. The CMP languages are spoken in most of the lesser Sunda Islands and the Moluccas in eastern Indonesia and on the Bomberai peninsula in western New Guinea. The EMP languages are divided into the South Halmahera-West New Guinea (SHWNG) subgroup and the vast Oceanic subgroup. The former consists of languages of southern Halmahera Island in eastern Indonesia and the Cendrawasih Bay area (coastline and islands) in Western New Guinea, while the latter consists of nearly all the languages of Melanesia, Micronesia and Polynesia in the South Pacific (with the exception of New Guinea, where Oceanic languages are spoken only in coastal enclaves and on the off shore islands). Blust's (2009) classification of the higher order Austronesian subgroups is outlined in Figure 1 below. The location of the languages are indicated on the map in Figure 2, with a close-up of the SHWNG languages in the map in Figure 3.

![Figure 1. Higher order subgroups of the Austronesian languages (after Blust 2009)](image-url)
Figure 2. The geographic location of the Austronesian subgroups (Lynch et al. 2002)
Although the representation of languages in distinct subgroups in a family tree is superior for illustrative purposes, it may not always be ideal for describing the actual history of a group of related languages. The family tree model corresponds best to the reality of cases in which two communities speaking the same language become geographically separated at one point after which the two communities, no longer in contact with each other, develop their own distinct languages. In cases like these, the relation between languages may in principle be mapped onto a tree. The family tree, however, is not easily applicable in situations of extensive language contact between unrelated or distantly related languages, which may result in different mixed language varieties or in what is called a Sprachbund (Trubetzkoy 1928 [cited in Toman 1995]) or a linguistic area (Emeneau 1956), where some linguistic features transcend language and language group borders. There are also other situations, relevant for the Austronesian area, in which languages may be said to be genealogically related, but nevertheless pose problems for the family tree model. For instance, if the separation of speakers in a speech community is not complete, or if changes occur within parts of the home territory of the community, chances are that the dividing line between languages will be less than sharp. As speech varieties grow increasingly diverse, the new languages that evolve may be linked in chains of dialects, displaying partially overlapping features rather than distinct boundaries. In such cases, the family tree model has its obvious limitations. Languages that have developed from a dialect chain may be collectively called a linkage, characterized by partially overlapping sets of innovations, while those that have developed by separation may be called a (sub)family, ideally defined
by innovations present in all daughter languages. In the latter case, a distinct proto-language is often reconstructable, while in the former, the depiction of a proto-language is more diffuse. It will be difficult to say the area or the period of time to which various reconstructed proto-forms belong. The picture is further complicated when a subfamily is the sister of a linkage, i.e. when a portion of an earlier dialect chain breaks off and becomes separated from the rest of the chain, eventually forming a distinct subfamily of languages. In such situations, the "left-behind" linkage does not have a common mother language that is separate from the mother language of the subfamily and the linkage together. Features that are present in the subfamily may be partially present in the linkage, which naturally makes subgrouping difficult.

The distinction between linkage and subfamily is relevant in describing the history of the Austronesian language family. Linkages have been identified and described for several Austronesian areas (see e.g. Tryon 1976; Geraghty 1983; Pawley & Green 1984; Ross 1988, 1995; Rehg 1995). Ross (1995) argues that the Formosan, WMP and CMP languages are "left-behind" linkages,23 since it has proven difficult to reconstruct a single proto-language for these three groups. Attempts often yield proto-languages very similar or identical to the proto-language reconstructed for the above node in the family tree (Blust 1999). In light of this, it is not surprising that there is no real consensus among linguists as to the internal relationships among these languages.

Earlier, the Formosan languages were either classified as a single subfamily (primarily following Dyen 1963) or as three principal subgroups immediately under the Austronesian node (following Ferrell 1969). Considering, however, the heterogeneity among the Formosan languages, such a classification cannot be maintained (cf. Blust 1999, 2009; Adelaar 2005). The nine Formosan subgroups suggested by Blust (1999) each consist of between one and five languages.

WMP is usually conveniently retained as a label for the languages of the WMP area, although this group should probably be dissolved into smaller groups, which would thus all be sisters to the CEMP subgroup, in analogy to what is suggested for the Formosan languages. WMP languages seem to cluster together genealogically in some 20 low-level subgroups (see e.g. Ross 1995, Adelaar 2005), the specific relationships between which are uncertain or unattainable. Nevertheless, some of these are often grouped together into a few larger groups defined geographically for the most part, such as the Sulawesi, Borneo, and Sundic (Malay Peninsula, Sumatra, Java, Bali) groups. Malagasy is assigned to the Borneo group since it has been shown (by Dahl 1951) that its closest relative is the Ma'anyan language belonging to one of the Borneo subgroups. The term 'Philippine languages'...

---

23 Ross (1995) uses the term "stay home" languages.
often refers to a typological category of languages rather than a genealogical subgroup, although some authors have suggested that all the languages of the Philippines (and possibly some of the languages of northern Borneo and Sulawesi) belong to a single Philippine family (e.g. Zorc 1986; Blust 1991; Adelaar 2005). More differentiated outlines of the interrelationships between the Philippine and Formosan languages have been attempted as well, for instance by Reid (1982), who suggests that Amis (Formosan) forms a distinct subgroup together with what he calls the Extra-Formosan Austronesian languages, which in turn split into two groups: the Northern Philippine languages and the MP languages.

Also within the much smaller CMP area, languages cluster together genealogically into a number of small families: at least seven subdivisions are commonly recognized (Ross 1995). The partly overlapping features of these indicate that they stem from a linkage, and perhaps the seven CMP subdivisions could each be seen as sisters to the EMP branch. Blust (1993), however, argues for a distinct genealogically based CMP family.

The MP, CEMP and EMP languages, on the other hand, are usually taken as distinct families rather than linkages (e.g. Ross et al 1998), although there is no real firm evidence for neither CEMP (Ross 1995; Adelaar 2005) nor EMP (Adelaar 2005). There is greater consensus concerning the status of the two daughter groups of EMP, i.e. SHWNG and Oceanic, as families. Of these two, the latter is the linguistically better attested, and indeed is the most firmly established subgroup of all the larger Austronesian subgroups. SHWNG also stands on very solid linguistic ground, although the boundary between this group and the geographically adjacent CMP group is not entirely clear in all details (Tryon 1995).

The first one to recognize the Oceanic languages as a distinct Austronesian subgroup was Otto Dempwolff (1937), who proposed that a proto-language that he called *Urmelanesisch* was the mother of the languages of Melanesia, Micronesia and Polynesia, thus corresponding to today's proto-Oceanic. The Oceanic subfamily primarily consists of the Western Oceanic, Eastern Oceanic (also referred to as Central-Eastern Oceanic) and Admiralties subgroups. The two related languages of the St Matthias Island (north of New Ireland) as well as Yapese in Micronesia are also considered to be two first order groups (Lynch et al 2002). The Admiralties languages appear to be a subfamily originating from a distinct proto language, while Western Oceanic is a linkage stemming from an earlier dialect chain (Ross 1988). The status of Eastern Oceanic as a linkage or a subfamily is unclear. The boundary between Western and Eastern Oceanic is, however, rather sharp in its delineation through the central Solomon Islands (Ross 1988).

24 Until recently, however, Yapese was often classified (though tentatively) as a WMP language (see e.g. Ruhlen 1987 (mirrored in Grimes 2000); Pawley & Ross 1995).
This may be the result of a separation of two Proto-Oceanic dialect chains that subsequently met again: Western Oceanic seems to have encroached on the western parts of the Eastern Oceanic area (Lynch et al 2002). The Western Oceanic linkage consists of the Meso-Melanesian, North New Guinea, and Papuan Tip sublinkages.

Originating with a suggestion by Pawley (1972), the Eastern Oceanic languages have often been subdivided into a Southeast Solomonic group and a Remote Oceanic group. Remote Oceanic also occurs (with various sister groups) in many later publications (Ruhlen 1987; Grimes 2000; Gordon 2005). But in a classification by Lynch et al (2002), this group is done away with altogether, and the Eastern Oceanic languages are instead grouped into: (i) the Southeast Solomonic subfamily, (ii) the Utupua-Vanikoro languages (also called the Eastern Outer Solomon Islands group), (iii) the Southern Oceanic linkage, (iv) the Micronesian subfamily, and (v) the Central Pacific linkage. Their Southern Oceanic linkage is a novel classification, based primarily on work by Lynch, who has specialized in Southern Vanuatu languages (cf. Lynch 2001). It consists of families and linkages that were all previously regarded as higher order subgroups at different levels within the Eastern Oceanic subfamily (see e.g. Grimes 2000).

Finally, the Central Pacific linkage is commonly seen as consisting of two sublinkages: one made up of Rotuman and the West Fijian dialects, the other of the East Fijian dialects and the Polynesian subfamily (see e.g. Geraghty 1983; Gordon 2005).

An outline of the Oceanic languages as described above is found in Figure 4. The maps of Figure 5 and Figure 6 show the location of the major subgroups of the Oceanic subfamily.

---

25 Pawley (1972) originally called this subgroup North Hebridian-Central Pacific but later renamed it Remote Oceanic (Pawley 1977 (cited in Bowden 1993)), which has become its established name.
Figure 4. Higher order subgroups of the Oceanic languages (after Lynch et al 2002)
3. AUSTRONESEAN FAMILY AND SOURCES OF DATA

Figure 5. The geographic locations of the Western Oceanic subgroups, and the primary Oceanic groups Admiralties and St Matthias (adapted from Lynch et al 2002)
Figure 6. The geographic locations of the Eastern Oceanic subgroups (Lynch et al. 2002)
3.1.2 Typological characteristics

The Austronesian languages display a wide variety of morphological types, from very analytic to agglutinative and even to those with many fusional tendencies. As is true for languages in general, however, these types are associated with subsections of grammar rather than with languages as a whole. In addition, it must be kept in mind that agglutination and fusion are not simple properties, as noted by Plank (1999). For instance, as is often the case in the WMP area, morphology may be agglutinative in terms of ease of morpheme segmentability, while each morpheme may be fusional in terms of the number of categories it covers.

As a rule, Austronesian languages tend to have more bound morphology associated with verbs than with nouns, a trait that is probably shared with most languages in the world. The most complex morphology is found in the Formosan and Philippine languages, with their intricate voice systems, and in the Oceanic languages of Melanesia, which often fuse TMA markers with markers for person and number of the subject into portmanteau forms prefixed to verbs. Bound morphemes in Austronesian include prefixes as well as suffixes; however, of the two, prefixes are more common, which is an unusual typological trait. Some co-occurrences of prefix plus suffix, found particularly in WMP, are quite idiosyncratic and may be analyzed as cases of circumfixation. In the Formosan and Philippine languages, productive infixation is a distinctive feature. The different affixation strategies are exemplified below by a Paamese pronominal-TMA prefix (90), a Lenakel pronominal suffix of inalienable possession (91), an Indonesian circumfix deriving abstract nouns (92), and a Tagalog actor focus voice infix (93).

90. Paamese (Southern Oceanic)

\[ \text{mo-matilu} \]
1.SG.IMMFUT-sleep

'I am about to sleep.' (Crowley 2002a, p 31)

91. Lenakel

\[ \text{ner-k miin petimw} \]
child-1.SG PL all

'all my children' (Lynch 1978, p 79)
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

92. Indonesian

\textit{ke[sulit]an}

NZR[difficult]

'difficulty' (Sneddon 1996, p 35)

93. Tagalog

\textit{b[um]ili ang lalake ng isda sa tindahan}

\textit{buy[AF.PFTV] TOP man GEN fish LOC store}

'The man bought fish in the store.' (MJ)

At the opposite end of the morphological complexity scale are the Polynes-
ian languages. Together with the Micronesian and Fijian languages, and
some of the languages of Melanesia, they predominantly use free particles to
distinguish TMA categories, as in the Niuean and Ulithian examples below.

94. Niuean (Polynesian)

\textit{ne fano a ia ki alofi neafi}

PST go PROP 3.SG to Alofi yesterday

'He went to Alofi yesterday.' (Kaulima & Beaumont 1994, p 38)

There are also Austronesian languages scattered throughout Southeast Asia
with very analytic morphology, such as many of the Land Dayak languages
(WMP) of Western Borneo, and Kéo and Waima’a (CMP) of the Lesser
Sunda Islands (Eastern Indonesia) (Himmelmann 2005a).

One of the most conspicuous properties of the Austronesian family as a
whole is the use of reduplication for a number of functions, with very few
Austronesian languages lacking this feature. Reduplication is often used to
indicate plural reference with nouns; repeated or ongoing aspect, or plural
agreement, with predicates; and intensity or emphasis in general. Such iconic
functions are very common for reduplicated forms in the languages of the
world. In WMP, reduplication of numerals is also very widespread, indicat-
ing such categories as restrictives (e.g. Ilokano \textit{dù-dua}, 'only two'), ordinals
(e.g. Siraya \textit{ka-ra-ruha}, 'second'), or distributivity (e.g. Malay \textit{dua-dua}, 'two
at a time') (Himmelmann 2005a). Reduplicative forms for diminutives or
their opposite, augmentatives, are quite common as well. Some languages
use reduplication, sporadically or productively, for inflection or in the deri-
vation of lexemes. The reduplication of the lexical base may be total or par-
tial, and in the latter case, base initial, base internal or base final. A few ex-
amples are presented below.
3. AUSTRONESIAN FAMILY AND SOURCES OF DATA

95. Balinese (Bali):  
   luh, 'female'
   luhluh, 'females'

96. Indonesian (Indonesia):  
   mata, 'eye'
   matamata, 'spy'

97. Tagalog (Philippines):  
   kain, 'eat' (root)
   kakain, 'will eat' (fut.)

98. Tahitian (Polynesia):  
   maita'i, 'good'
   maitata'i, 'good' (pl. agr.)

99. Thao (Taiwan):  
   shnara, 'burn' (?)
   pashnaranara, 'burn repeatedly'

(Examples from Himmelmann 2005a, pp 121-3, except 97 and 98, which are provided by the present author)

Another property that is widespread in the Austronesian area is the multifunctionality of words. The word class status of lexical items is determined by syntactic function rather than by inherent lexical properties. For instance, most languages lack an open class of adjectives. Usually, adnominal lexical modifiers may function just as well as main predicates. Compare the two Kéo examples below, in which the word bhugé, '(be) fat', is used attributively and predicatively, respectively.

100. Kéo

   a. 'ana bhugé ké ka woso ngata  
      child fat that eat much INTEN
   
      'That fat child eats very much.' (Baird 2002, p 120)

   b. 'ana 'imu bhugé réréréé  
      child 3.SG fat very

      'Her children are very fat.' (Baird 2002, p 119)

Even the difference between nouns and verbs is not clear-cut in parts of the Austronesian area. This is particularly the case in the WMP and Polynesian languages but is also found elsewhere. In general, languages with very analytic morphology seem to allow for somewhat greater distributional flexibility with lexical items. Compare the Tongan examples below (see Vonen (1994) and Broschart (1997) for two very different approaches to analyzing such multifunctionality in Polynesian languages, and in general).

101. Tongan (Polynesian)

   a. na'e lele e kau fefiné  
      PST run ART PL woman.DEF

      'The women were running.' (Broschart 1997, p 134)
b. na'e fefine katoa e kau lelé
     PST woman all ART PL run.DEF

'The ones running were all female.' (Broschart 1997, p 134)

The content words fefine, '(be) woman', and lele, '(be) running', seem not to be determined for word class lexically, although functional words such as articles and TMA markers allow us to identify verb phrases and noun phrases syntactically. However, in the morphologically more complex Philippine languages, words that are obligatorily inflected for focus and TMA categories and conveniently function as main predicates may also be used as arguments with their voice-TMA inflections still in place if paired with case markers. Compare the Tagalog examples below.

102. Tagalog

a. mag-sa-salita' si rosa
     AF.FUT-RD-talk NOM.PROP Rosa

'Rosa will speak.' (Schachter & Otanes 1972, p 154)

b. si rosa ang mag-sa-salita'
     NOM.PROP Rosa TOP AF.FUT-RD-talk

'The one who will speak is Rosa.'
     (Schachter & Otanes 1972, p 154)

In example 102 above, the word magsasalita, 'will talk', is inflected for future tense and actor topic voice, even when used as an argument. Words inflected for other tense and voice categories are also available for multiple functions. Thus, while there is a distinctly verbal inflectional paradigm in Tagalog on a morphological level, there is still multifunctionality on a syntactic level. This is, of course, very different from the Tongan situation. As for phrasal syntagms, however, function words in Tagalog, such as case marker ang, permit NPs and VPs to be postulated, similar to the Tongan case. Reid (2002) proposes an analysis of Tagalog ang as a demonstrative with the following content word(s) – whether morphologically verbal or nominal – constituting a relative clause. Given that Tongan kau historically is a noun, a similar analysis could be employed for Tongan.

The examples from Tongan and Tagalog illustrate that we have to distinguish, on the one hand, between morphological (inflectional) and syntactic (distributional) criteria for determining word class, and on the other hand, a paradigmatic (lexical) and syntagmatic (phrasal) level of analysis. In many European languages, multifunctionality is much more restricted both mor-
phologically and syntactically, so that both lexical items and phrases are easily determined for class.

A common typological distinction is presence versus absence of an alienability split in adnominal possessive constructions in languages. That is, languages differ in whether or not they have a specific construction for expressing possessive relations between inalienably possessed entities, typically body parts and/or kin terms, as in *my father* and *my hand*, and another for expressing possessive relations between alienably possessed entities, which may be acquired and disposed of, such as *my pen* and *my radio*. This feature divides the Austronesian languages in two very distinct groups. The Formosan and WMP languages (except those of southern Sulawesi) generally show no alienability distinctions in adnominal possessive constructions, while the CMP, SHWNG and Oceanic languages (as well as the languages of southern Sulawesi) to a great extent do have alienability or similar distinctions in adnominal possessive constructions. Note the different adnominal possessive constructions from A'ara, an Oceanic language, below.

103. A'ara (Meso-Melanesian)

a. *kma-nya* 
   father-POSS.3.SG
   'his father' (Lichtenberk 1985, p 104)

b. *no-gu mola iara* 
   CL-POSS.1.SG canoe 1.SG
   'my canoe' (Lichtenberk 1985, p 99)

On the other hand, in Tagalog, which is a WMP language, both of the above kinds of relations are expressed by the same construction.

104. Tagalog

a. *ang tatay niya*
   TOP father 3.SG.GEN
   'his father' (MJ)

b. *ang bangka ko*
   TOP canoe 1.SG.GEN
   'my canoe' (MJ)

The details of the alienability distinctions in the Austronesian languages that display them differ considerably, both with regard to form and with regard to
usage. In Buru, for instance, the inalienable construction is used for body parts (and other part-whole relations) but not for kin terms. Polynesian languages may also be described as having an alienability split, as they have two different constructions for adnominal possession, the use of which generally overlaps with the use of typical alienable and inalienable construction, respectively. At closer inspection, however, as attested in many studies (e.g. Wilson 1982), the determining factor in Polynesian possessive constructions pertains to the amount of control exercised by the possessor over the possessive relation rather than alienability. In some cases, the two different constructions are possible with the same possessor and possessee, resulting in a difference in relational meaning (e.g. 'the boy's picture' represented by a control construction indicates that the boy owns the picture and thus controls the relation, while 'the boy's picture' represented by a no control construction indicates that the boy's image is in the picture).

Finally, looking at word order, three word order types are particularly common in Austronesia: SVO, VSO and VOS. Other combinations do occur but are not as common. SVO is predominantly found in Southeast Asia (except the Philippines) and Oceania (except Polynesia). VSO is found mainly in the Philippines and in Polynesia, while VOS is relatively common in Taiwan. VOS is also found in Madagascar (Malagasy). Some of the Austronesian languages in New Guinea display SOV word order, being under the influence of nearby Papuan SOV languages. The situation is well reflected by word order distributions among the sample languages in the present study, as shown in the table below.

Table 6. Word order distribution among sample languages

<table>
<thead>
<tr>
<th>word order</th>
<th>no. of lgs in sample</th>
<th>percentage</th>
<th>predominant for Austronesian lgs in</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVO</td>
<td>25</td>
<td>58.1</td>
<td>Oceania (except Polynesia), Southeast Asia (except Philippines)</td>
</tr>
<tr>
<td>VSO</td>
<td>11</td>
<td>25.6</td>
<td>Polynesia, Philippines</td>
</tr>
<tr>
<td>VOS</td>
<td>5</td>
<td>11.6</td>
<td>Relatively common in Taiwan (Malagasy goes here too)</td>
</tr>
<tr>
<td>SOV</td>
<td>2</td>
<td>4.7</td>
<td>Rare, some in New Guinea</td>
</tr>
<tr>
<td>TOT</td>
<td>43</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
3.2 Sources, sample and database

3.2.1 Sources of data

The data used for this study was collected in different ways. The bulk of the data comes from descriptive grammars or books or articles on certain languages. However, for some of the languages, these materials were complemented by information obtained from questionnaires, which were filled out both by native speakers and by linguists familiar with the languages in question. The questionnaire for native speakers primarily consisted of sets of clauses to translate, while that for linguists also included questions about grammatical details (see appendices 1 and 2, respectively). It would have been ideal, of course, to have been able to work with native speakers to a greater extent in order to obtain first-hand data, but since my access to speakers of a variety of Austronesian languages was limited, and since collecting data in this way is extremely time consuming, I had to rely mostly on grammars and other written descriptions (see appendix 3).

Although information is usually easily accessed in reference grammars, there are several problems with collecting data in this way. One obvious problem is that a certain grammar may not include the information that is being sought. For instance, even when a grammar has a thorough treatment of the area of clause combining, information on how specific subordination tests (see section 2.1.2 above) can be applied to different constructions is often lacking.

Another problem is that different grammars have different resolution of detail in the description of TC clause combining. Counting the number of morpho-syntactic constructions featured for the relevant TC relations in a certain grammar is a way of measuring the resolution of detail for the language described therein. The extent, however, to which different results of such counts correspond to actual variations in detail of data or to variations between languages in the size of their pools of constructions available for the expression of TC relations is not known. Nevertheless, such figures may serve as a rough guide and will be presented in section 3.2.4. We can assume that each grammar covers the constructions most commonly used in the respective languages for description, and therefore we can also be relatively certain that the most relevant constructions are covered in this study, though complete coverage is not obtainable. The selection base could of course be fine-tuned by using corpora for each language in order to determine what constructions are more frequently used – i.e. are more conventional – in a language. But however desirable such an approach would be, there are insurmountable methodological problems involved in the approach, as there are few rich corpora of Austronesian languages and assembling a well-
balanced corpus from scratch even for one language, not to mention 43, would be far too time-consuming for the present study.

A more serious problem is that of Eurocentrism on the part of grammarians describing "exotic" languages. That is, are the TC categories described for a certain "exotic" language established categories in that language, or do they tend to be perceived by linguists because they have them in their own languages? Most writers of descriptive grammars are speakers of European languages or else trained in the European grammar writing tradition. There is a very real possibility that the constructions matching the semantic relations familiar to speakers of European languages, such as reason and concession, may not be commonly used by native speakers or are possibly even non-existent in a specific language. Even so, constructions with similar meaning or one of several non-conventionalized variants may be produced if speakers are asked about them (see further section 4.1.2). In fact, some grammars hint that the language they describe prefers more implicit constructions when TC relations – as well as other relations between states of affairs – are intended, even when more specific constructions are also in use (see section 6.1.3).

However, there is little to be done about these problems apart from being aware of them. We will move ahead as best we can using the data that is available.

### 3.2.2 Sample of languages

Creating a well-balanced genealogical sample from the Austronesian language family is notoriously difficult. The difficulty lies in the way branches have developed in this family. Large subfamilies are often related to much smaller linkages, or small subfamilies stemming from earlier linkages. Let us look at the first order branches of Austronesian, for instance. There are nine very small branches of languages spoken in Taiwan – collectively referred to as the Formosan languages – altogether comprising only 20 languages, and one very large branch comprising the more than 1,200 languages spoken throughout the rest of the Austronesian world – i.e. the Malayo-Polynesian (MP) branch of languages. The Formosan languages, of course, represent the "stay homes", while the MP family represents the "move aways" on this level of the family tree. This pattern is repeated at various levels throughout the Austronesian family. Thus, if we were to collect a sample of Austronesian languages solely based on genealogical distribution, we would have to choose nine languages from Taiwan (one from each branch) for every Malayo-Polynesian language we choose. Now, following such a procedure would of course be quite absurd, as the sample would be highly unbalanced areally. Even if comparative evidence points to a split between the Formosan branches that is as ancient as the split between any of these branches and the vast MP branch, the areal proximity of the Formosan languages has kept the
overall dissimilation rate down among these as compared to the dissimilation rate having occurred between many of the MP sub-branches at various levels.

Dryer's (1989) method of sampling languages from more-or-less equally sized genera, i.e. subgroups, regardless of the level at which they are found in a family tree, does not work well either, because of the large difference in size often found between subgroups at different levels. Take the Atayalic subgroup in Taiwan, for instance, which consists of two languages. This size is obviously too small to be useful, as it would be hard to find enough equally sized subgroups evenly spread out across the Austronesian area. However, moving up to the next level in the family tree, we are already at the top level of Austronesian. Similarly, we find sub-families throughout the tree of ten or so individual languages that are immediate sisters to very large groups, perhaps comprising a hundred languages. It is not obvious how to select equally sized genera from this configuration.

The only solution seems to be a compromise between areal and genealogical considerations. It is along such lines that the core sample of 30 languages used in this study was assembled. The starting point was the commonly recognized top-level family structure of Austronesian comprising the following five major groupings (in bold), some of which are areally defined (in brackets) and some of which are genealogically defined:

Austronesian

[Formosan]
Malayo-Polynesian - MP

[West Malayo-Polynesian – WMP]
Central-East Malayo-Polynesian – CEMP

Central Malayo-Polynesian – CMP

East Malayo-Polynesian – EMP

South Halmahera - West New Guinea – SHWNG

Oceanic

All genealogically defined subgroups with fewer than one hundred languages (an arbitrary figure) subsumed under each of these five major groupings were listed. Subgroups constituting branches of a higher order subgroup on the list were deleted. The 15th edition of the Ethnologue (Gordon 2005) was used as a basis for the genealogical classification, except in the case of Oceanic for which Lynch et al. (2002) was used. From this list of sub-

---

26 Or, to be more precise, when the work on this study began, only the 14th edition of Ethnologue (Grimes 2000) was available, which was used to list the subgroups with fewer than a hundred languages. The list was later recreated and updated with the information in Gordon.
groups, thirty languages were selected, each from a different subgroup, based on the availability of grammars and other usable sources. The distribution of individual languages over the five major Austronesian groups was determined according to the size of these groups in terms of number of languages. Thus, one language was chosen from a Formosan subgroup, thirteen languages were chosen from WMP subgroups, etc., as displayed in Table 7 below.

Table 7. Distribution of the core sample languages across the major Austronesian subgroups (proportional distribution in parentheses)

<table>
<thead>
<tr>
<th>Austronesian</th>
<th>30 (30.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formosan</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>MP</td>
<td>29 (29.5)</td>
</tr>
<tr>
<td>WMP</td>
<td>13 (12.7)</td>
</tr>
<tr>
<td>CEMP</td>
<td>16 (16.8)</td>
</tr>
<tr>
<td>CMP</td>
<td>3 (4.0)</td>
</tr>
<tr>
<td>EMP</td>
<td>13 (12.8)</td>
</tr>
<tr>
<td>SHWNG</td>
<td>1 (1.0)</td>
</tr>
<tr>
<td>Oceanic</td>
<td>12 (11.8)</td>
</tr>
</tbody>
</table>

In this way, languages were selected from subgroups together representing 1,069 languages, which is 84.3 % of the total number of Austronesian languages (1,268 according to Gordon 2005).

An additional 13 languages, for which information was readily available, were selected for an extended sample. These languages are more-or-less randomly distributed over the Austronesian family. The total distribution of the languages in the extended sample is displayed in Table 8.
Table 8. Distribution of the extended sample languages across the major Austronesian subgroups (proportional distribution in parentheses)

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austronesian</td>
<td>43</td>
<td>(43.0)</td>
</tr>
<tr>
<td>Formosan</td>
<td>2</td>
<td>(0.7)</td>
</tr>
<tr>
<td>MP</td>
<td>41</td>
<td>(42.3)</td>
</tr>
<tr>
<td>WMP</td>
<td>17</td>
<td>(18.2)</td>
</tr>
<tr>
<td>CEMP</td>
<td>24</td>
<td>(24.0)</td>
</tr>
<tr>
<td>CMP</td>
<td>4</td>
<td>(5.7)</td>
</tr>
<tr>
<td>EMP</td>
<td>20</td>
<td>(18.3)</td>
</tr>
<tr>
<td>SHWNG</td>
<td>1</td>
<td>(1.4)</td>
</tr>
<tr>
<td>Oceanic</td>
<td>19</td>
<td>(16.9)</td>
</tr>
</tbody>
</table>

Some of the additional languages added previously unrepresented subgroups, thereby increasing overall coverage, while others split already represented subgroups. In the latter case, if the split subgroups had more than two immediate branches (as was often the case), it meant that some of these branches became unrepresented, thereby decreasing the overall coverage. All in all, this resulted in a slightly lower coverage for the extended sample, as 1,018 languages, or 80.3 %, are represented out of the total number of Austronesian languages (for a detailed outline, see appendix 4). The sample languages are listed in Table 9 overleaf. Care has been taken not to include languages spoken geographically close to each other, although the density of languages (and subgroups) in some areas made relative proximity unavoidable in some cases. The geographic distribution of the languages in the core and extended samples are shown in the map of Figure 7 (overleaf). The 30 core sample languages are indicated by red dots, while the additional 13 languages of the extended sample are indicated by yellow dots.
Table 9. Sample languages (bracketed languages are in the extended sample only).

<table>
<thead>
<tr>
<th>ID</th>
<th>Language</th>
<th>Group</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amis</td>
<td>Form</td>
<td>Taiwan</td>
</tr>
<tr>
<td>2</td>
<td>[Seediq]</td>
<td>Form</td>
<td>Taiwan</td>
</tr>
<tr>
<td>3</td>
<td>Palauan</td>
<td>WMP</td>
<td>Micronesia</td>
</tr>
<tr>
<td>4</td>
<td>Central Cagayan Agta</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>5</td>
<td>Tagalog</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>6</td>
<td>Western Subanon</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>7</td>
<td>[Sarangani Manobo]</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>8</td>
<td>Tboli</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>9</td>
<td>Yakan</td>
<td>WMP</td>
<td>Philippines</td>
</tr>
<tr>
<td>10</td>
<td>Sangir</td>
<td>WMP</td>
<td>Sulawesi</td>
</tr>
<tr>
<td>11</td>
<td>[Tondano]</td>
<td>WMP</td>
<td>Sulawesi</td>
</tr>
<tr>
<td>12</td>
<td>Muna</td>
<td>WMP</td>
<td>Sulawesi</td>
</tr>
<tr>
<td>13</td>
<td>Coastal Konjo</td>
<td>WMP</td>
<td>Sulawesi</td>
</tr>
<tr>
<td>14</td>
<td>Ma'anyan</td>
<td>WMP</td>
<td>Borneo, Indonesia</td>
</tr>
<tr>
<td>15</td>
<td>Labuk Kadazan</td>
<td>WMP</td>
<td>Borneo, Malaysia</td>
</tr>
<tr>
<td>16</td>
<td>Karo Batak</td>
<td>WMP</td>
<td>Sumatra</td>
</tr>
<tr>
<td>17</td>
<td>[Acehnese]</td>
<td>WMP</td>
<td>Sumatra</td>
</tr>
<tr>
<td>18</td>
<td>Eastern Cham</td>
<td>WMP</td>
<td>Vietnam</td>
</tr>
<tr>
<td>19</td>
<td>[Indonesian]</td>
<td>WMP</td>
<td>Indonesia</td>
</tr>
<tr>
<td>20</td>
<td>Kambera</td>
<td>CMP</td>
<td>Nusa Tenggara, Indonesia</td>
</tr>
<tr>
<td>21</td>
<td>Tetun</td>
<td>CMP</td>
<td>Nusa Tenggara, Indonesia</td>
</tr>
<tr>
<td>22</td>
<td>[Leti]</td>
<td>CMP</td>
<td>Maluku</td>
</tr>
<tr>
<td>23</td>
<td>Buru</td>
<td>CMP</td>
<td>Maluku</td>
</tr>
<tr>
<td>24</td>
<td>Taba</td>
<td>SHWNG</td>
<td>Maluku</td>
</tr>
<tr>
<td>25</td>
<td>Loniu</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>26</td>
<td>Manam</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>27</td>
<td>Mbula</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>28</td>
<td>Yabem</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>29</td>
<td>[Iwal]</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>30</td>
<td>Mekeo</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>31</td>
<td>Nakanai</td>
<td>Oceanic</td>
<td>PNG</td>
</tr>
<tr>
<td>32</td>
<td>[Hoava]</td>
<td>Oceanic</td>
<td>Solomon Islands</td>
</tr>
<tr>
<td>33</td>
<td>Longgu</td>
<td>Oceanic</td>
<td>Solomon Islands</td>
</tr>
<tr>
<td>34</td>
<td>Erromanga</td>
<td>Oceanic</td>
<td>Vanuatu</td>
</tr>
<tr>
<td>35</td>
<td>[Lenakel]</td>
<td>Oceanic</td>
<td>Vanuatu</td>
</tr>
<tr>
<td>36</td>
<td>Araki</td>
<td>Oceanic</td>
<td>Vanuatu</td>
</tr>
<tr>
<td>37</td>
<td>[Big Nambas]</td>
<td>Oceanic</td>
<td>Vanuatu</td>
</tr>
<tr>
<td>38</td>
<td>Tinrin</td>
<td>Oceanic</td>
<td>New Caledonia</td>
</tr>
<tr>
<td>39</td>
<td>Kusaiean</td>
<td>Oceanic</td>
<td>Micronesia</td>
</tr>
<tr>
<td>40</td>
<td>Samoan</td>
<td>Oceanic</td>
<td>Polynesia</td>
</tr>
<tr>
<td>41</td>
<td>[Rapanui]</td>
<td>Oceanic</td>
<td>Polynesia</td>
</tr>
<tr>
<td>42</td>
<td>[Hawaiian]</td>
<td>Oceanic</td>
<td>Polynesia</td>
</tr>
<tr>
<td>43</td>
<td>[Maori]</td>
<td>Oceanic</td>
<td>Polynesia</td>
</tr>
</tbody>
</table>

\[27\] This language is referred to as 'Labuk-Kinabatangan Kadazan' in Lewis (2009) and according to Hope Hurlbut (personal communication), this is the name that the speakers themselves prefer. The most common names used in the linguistic literature for the language seems to be 'Labuk-Kinabatangan Kadazan' (e.g. Lewis 2009), 'Eastern Kadazan' (e.g. Hurlbut 1990) and occasionally 'Labuk Kadazan' (e.g. Hurlbut 1981). I have chosen to use 'Labuk Kadazan' throughout this thesis both because it includes part of the name the speakers prefer and because it is less cumbersome than the entire sequence 'Labuk-Kinabatangan Kadazan'.
Figure 7. Geographic distribution of the sample languages
3.2.3 The database

All the different constructions found in the sources were listed in my database with information on a number of parameters. In some languages, the same construction was used for different semantic relations. In such cases, the relevant constructions were listed as two (or more) records in the database, each time with a different value for a semantic relation. The parameters in the database primarily relate to morpho-syntactic features and relational semantics of each construction, but also to general information about the sample languages, such as genealogy, geographic location and number of speakers. The principal groups of parameters are listed below.

- General information (language name, record ID, construction ID, core or extended sample, basic word order, etc.)
- Semantic relation (general and specific information about the semantic relation)
- Relation marker (information about form, meaning, class, segmentation and possible cognates of the relation marker and, if relevant, any additional cohesive device)
- Construction (information on morpho-syntactic deviations of any of the clauses, syntactic relation and explicitness of semantic relation)
- Morpho-syntactic details (general grammatical description of the construction, possible orders between clauses (and relation marker))
- Additional information (sources of data, references to representative examples, genealogical information, geographic location, number of speakers (and source), etc.)

In total, the database consists of 1,397 records, describing the details of 1,180 constructions in 43 Austronesian languages along 20 parameters for 10 different general semantic relations.

3.2.4 Factors possibly distorting the results

A greater number of sample languages would, of course, have been desirable, in order to obtain more solid results statistically. However, some of the

---

28 Recall the discussion in section 1.4.3 about how the term "construction" is used in this study.

29 In fact, there are several additional parameters in the database providing information for analyses that never actually made it into the final version of the dissertation.
branches of Austronesian are not that well covered by descriptive material, which means that a larger sample would risk being less well balanced. Indeed, the difference between my core and extended samples already reflects this situation. The extended sample is biased towards the Oceanic languages, as sources are readily available for many of these languages. Sources for the languages of Borneo and Sulawesi, on the other hand, are scarce and relatively hard to come by. Collecting data for a larger sample would, of course, also have been more time consuming. So the number of languages in the samples was limited to 30 and 43 languages, respectively.

The problem of determining the universality of semantic relations between states of affairs has already been discussed, as well as the fact that sources vary in detail in their grammatical descriptions (see section 3.2.1). As an evaluation of the resolution of detail for each language, Table 10 below shows the number of records that was registered for each language in this study.

Table 10. Number of records registered for each language (bracketed languages are in the extended sample only).

<table>
<thead>
<tr>
<th>ID</th>
<th>Language</th>
<th>Group</th>
<th>Number of records</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amis</td>
<td>Form</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>[Seediq]</td>
<td>Form</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>Palauan</td>
<td>WMP</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Central Cagayan Agta</td>
<td>WMP</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>Tagalog</td>
<td>WMP</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>Western Subanon</td>
<td>WMP</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>[Sarangani Manobo]</td>
<td>WMP</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>Tboli</td>
<td>WMP</td>
<td>35</td>
</tr>
<tr>
<td>9</td>
<td>Yakan</td>
<td>WMP</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>Sangir</td>
<td>WMP</td>
<td>43</td>
</tr>
<tr>
<td>11</td>
<td>[Tondano]</td>
<td>WMP</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>Muna</td>
<td>WMP</td>
<td>63</td>
</tr>
<tr>
<td>13</td>
<td>Coastal Konjo</td>
<td>WMP</td>
<td>42</td>
</tr>
<tr>
<td>14</td>
<td>Ma'ananyan</td>
<td>WMP</td>
<td>25</td>
</tr>
<tr>
<td>15</td>
<td>Labuk Kadazan</td>
<td>WMP</td>
<td>39</td>
</tr>
<tr>
<td>16</td>
<td>Karo Batak</td>
<td>WMP</td>
<td>83</td>
</tr>
<tr>
<td>17</td>
<td>[Acehnese]</td>
<td>WMP</td>
<td>22</td>
</tr>
<tr>
<td>18</td>
<td>Eastern Cham</td>
<td>WMP</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>[Indonesian]</td>
<td>WMP</td>
<td>82</td>
</tr>
<tr>
<td>20</td>
<td>Kambera</td>
<td>CMP</td>
<td>27</td>
</tr>
<tr>
<td>21</td>
<td>Tetun</td>
<td>CMP</td>
<td>54</td>
</tr>
<tr>
<td>22</td>
<td>Leti</td>
<td>CMP</td>
<td>14</td>
</tr>
<tr>
<td>23</td>
<td>Buru</td>
<td>CMP</td>
<td>33</td>
</tr>
<tr>
<td>24</td>
<td>Taba</td>
<td>SHWNG</td>
<td>25</td>
</tr>
</tbody>
</table>
Another factor possibly distorting the results is that, in some cases, an unambiguous classification is not possible. Even with the best sources, the data is sometimes not detailed enough to determine whether we are dealing with one construction only or if there are morpho-syntactic differences that warrant the postulation of more than one construction.
In this chapter, I will introduce and provide examples of all the semantic relations investigated in this study. For each semantic relation, I will discuss in some detail what the category includes and present some of the subdivisions that have been attested to occur in Austronesian languages, e.g. the distinction between intentional and realized purpose. The chapter is divided into three parts: first, there is a general discussion on semantic relations (4.1); then some practical considerations in the actual semantic classification of Austronesian constructions are considered (4.2); and finally, the semantic relations as represented in the sample languages, including the nuances of meaning that are commonly expressed, are presented (4.3).

4.1 General discussion

4.1.1 Background and the semantic relations used in the present study

When clause combining constructions are investigated, the semantic relation between the states of affairs depicted by the clauses involved are commonly described in terms of co-occurrence, condition, reason, purpose etc. Halliday & Hasan (1976) classified these into four sets of relations based their expression in English: additive, adversative, temporal and causal. Similar sets have been found useful also for other languages (e.g. Rudolph 1996 for Indo-European).

In one of the first typological studies of adverbial clauses, Thompson & Longacre (1985) describes eleven basic semantic relations, which are listed in (105) below:

105. time ('when', 'after', 'before')
     place ('where')
     manner ('as', 'as if')
     purpose ('in order that')
     reason ('because')
     circumstance ('by')
simultaneity ('while')
condition ('if', 'even if')
concession ('although')
substitution ('instead of', 'rather than')
additive ('besides')

The relations in (105) are claimed to have salient representative constructions in most of the world's languages. Kortmann (1997), in a study of adverbial subordinators in the European languages, sees splitting some of the relational categories above (e.g. **time** into **simultaneity overlap** 'when', **anteriority** 'after' and **posteriority** 'before') as being justifiable and adds negated counterparts to some of the categories (e.g. **negative condition** 'unless'). He ends up with 32 semantic relations commonly marked by what he refers to as 'ideal adverbial subordinators' in the languages of Europe, and he organizes them into four overarching types:

106. **time** (e.g. simultaneity overlap, posteriority, anteriority)
    **CCC** (cause, condition, concession and related relations)
    **mode** (e.g. similarity, instrument, proportion)
    **other** (e.g. place, substitution, addition)

There are plenty of other suggestions for categorizations of semantic relations between clauses, or rather the state of affairs they depict, both in grammar and discourse studies, more or less similar to those discussed above (see for example Mann & Thompson 1983, 1988; Givon 1990; Hengeveld 1993, 1998). However, as already illustrated in the discussion above, there is no real consensus among scholars on how to categorize the relevant semantic relations between two (or more) states of affairs, or how best to organize the categories. The categorization grid could be arbitrarily detailed, since some languages express distinctions that other languages lack. Therefore, I have found it best to investigate relations on a relatively general level. In this study, the ten semantic relations presented in (107) below were selected for investigation in the Austronesian languages. These can be divided into those denoting temporal relations and those denoting what I call co-varying relations, depicting states of affairs in which one is in some sense co-varying with the other.

---

30 Kortmann 1997, as well as this study, employs the terms 'anteriority' and 'posteriority' in a way that may seem counterintuitive. This is, however, consistent with the use of some other relational terms (for further discussion, see section 4.2.3).

31 See Kortmann 1997, pp 71-77, for his definition of an ideal adverbial subordinator.
The reason that these specific semantic relations are singled out in this study is that they are semantically relatively distinct from each other and previously well described in the general linguistic literature. They are also well represented by distinct clause combining constructions in most of the Austronesian languages, as confirmed at the onset of the investigation by consulting all the Austronesian language grammars at my disposal. Below are English representative examples of each of the ten relevant relations (they will be further illustrated by Austronesian examples in section 4.3 below).

108. a. They took care of his cat when he was on vacation.  
    b. We played a game of cards before we went into town.  
    c. They went to the airport after they packed their bags.  
    d. He read her stories until they both fell asleep.  
    e. We have been close since we met at Woodstock in 1969.  
    f. We will surely celebrate, if he gets the job.  
    g. He plays the guitar, although he never took lessons.  
    h. You must pay for it, in order for them to send it.  
    i. I fell over because I tripped on a loose rug.  
    j. We got stuck in traffic, so that we were late.

4.1.2 The universality of semantic relations

As mentioned above, the semantic relations listed in (107) above are all well represented by clause combining constructions in most of the Austronesian
languages and described in grammars of these languages. Whether this reflects the universality of these relations, or a certain degree of Eurocentricity on the part of the authors of grammatical descriptions is a problematic question that cannot be resolved easily. On the one hand, it can be argued that human societies will always find it necessary to be able to communicate relations such as co-occurrence, condition, concession, etc., and any language will have more or less conventionalized constructions for them. On the other hand, although constructions expressing these relations indeed seem to be described in descriptive grammars for most languages of the world, it can be argued that their prominence in the widely explored European languages make them convenient tools to use also for describing other languages, whether or not they actually correspond to morpho-syntactic patterns of the languages in question. Perhaps constructions expressing the semantic relations presented here constitute an areal or genealogical trait of Europe or Indo-European rather than a universal one. It should be kept in mind that most grammarians are native speakers of a European language as well as schooled in the European grammar writing tradition. The fact that the Austronesian languages display a tendency to use much less explicit means of conveying these relations, relying much more heavily on contextual information and propositional content than, for instance, English does, is indeed an indication that the relational categories are not represented in the same way in these languages.32

Also, as evident from well-studied languages, different relation meanings may be expressed by the same relational construction in a given language. Sometimes, this is because they are used in a particular context that affects the unmarked reading of a relational construction. At times, several readings are even possible at once. Consider, for instance, a case in which someone calls a friend who is just about to give her baby a bath. Even if the caller has been told that her friend is about to start this activity, she might choose to say something along these lines, using a conditional clause:

109. **If you are giving the baby a bath, I'll call back later.**

The relation primarily conveyed by the conditional clause here is actually one of reason; the caller is explaining the reason for why she will call back later. At the same time, the conditional format represents the state of affairs as hypothetical, i.e. as something assumed rather than known to be real. This might be a polite and subtle way for the caller to hint at the possibility of her friend perhaps postponing the bathing of her child, in which case condition

32 There are indications that other languages without much contact with European languages also tend to use implicit clause combining constructions to a greater extent, a case in point being Pirahã (Everett 2005), a language isolate in the Amazonian forest of South America.
and reason is wrapped up in the same parcel. Such context sensitivity gives Harris (1988, 1989, 1990) some of his arguments for describing various semantic relations (such as condition and reason) as focal points in the semantic space covered by combined clauses, rather than discrete categories. He does not, however, commit himself to whether these focal points correspond to universal conceptual categories.

There are also examples of constructions that plainly do not make the distinction between two or more of the traditional semantic relational categories. The most famous case pertains to constructions that neutralize the distinction between open conditions and future co-occurrence. This is quite common in Austronesian languages (as well as in other language families), while in most European languages, the speaker must decide whether the state of affairs in question is regarded as one that will eventually become realized (future co-occurrence: when + future tense) or merely as a potential state of affairs (open condition: if). Since in both cases the clauses express non-realized states of affairs (or in some cases states of affairs not confirmed to be realized), the same construction can be used in many Austronesian languages without commitment to any certainty of realization of the states of affairs, as illustrated by the translations of the Samoan example below.

110. Samoan

\[
\begin{align*}
\text{ʻafai} & \quad \text{ʻoleʻā} \quad \text{ʻoulua} \quad ā, \\
\text{if/when} & \quad \text{FUT} \quad 2.\text{DU} \quad \text{go.PL} \\
\text{ia, tautuanā ma} & \quad \text{ʻoulua e} \quad \text{ʻaumai popo} \\
\text{well remember for} & \quad 2.\text{DU} \quad \text{GNR} \quad \text{bring coconut}
\end{align*}
\]

'If you go, well, remember to bring the coconuts.' or
'When you go, well, remember to bring the coconuts.' (AJ)

The difference between the two interpretations is that the former denotes non-reality with potential (though not certain) future realization, while the latter denotes non-reality with presupposed future realization. In Samoan, only the non-realized part is coded in these cases. This does not mean that the semantic categories of condition and co-occurrence do not exist in Samoan and other languages with similar constructions. In most of the sample languages with a collapsed condition/co-occurrence construction, there are other constructions for which co-occurrence is indeed formally distinguished from condition (see further discussion in section 5.2.1 below). This is normally the case for past time co-occurrence and counter-fact conditions. The former expresses realized state of affairs, for which the temporal relation to another state of affairs is valid information, while the latter expresses an imaginary state of affairs. Cf. the Samoan clauses below.
111. Samoan

a. *ina 'ua malama le taeao o le*
when INCH break the morning of the

*isi aso sa matua mafatia lava*
other day PST very be.exhausted indeed

tamaiti i le fia 'a'ai
child.PL at the want eat.PL

'When morning broke of another day, the children were totally exhausted from hunger.' (Mosel/Hovdhaugen 1992, p 632)

b. *'ana 'e lē sau,*
if.CF 2.SG NEG come

*semanū 'ou te alu atu*
probably 1.SG GNR go DIR

'Had you not come, I would probably have gone to see you.'
(Mosel/Hovdhaugen 1992, p 656)

In these examples, the certainty of realization on the part of the speaker is also very different. In fact, they are polar opposites. In the (a) example, the speaker is certain that the state of affairs expressed did happen, and in the (b) example, the speaker is certain that the state of affairs expressed did not happen.

Obviously, there are difficulties in assuming that the semantic categories traditionally used are universal (or at least that they are universally uniform) when investigating clause combining constructions. Nonetheless, all evidence suggests that they are very widespread and that most languages have ways of conveying them, even though the details of how they do so may vary.
4. SEMANTIC RELATIONS

4.2 Practical considerations in the semantic classification

4.2.1 Principles and procedures

With regard to the designated terms for the relations investigated in this study, it can be noted that although they describe the semantic character of the relation, the argument for identifying the relation that holds between the clauses in an example like *I fell over because I tripped on a loose rug* as a relation of reason is in fact morpho-syntactic. Reason clauses express the reason for a certain result, so a reason relation between clauses does, by necessity, include both a clause denoting a reason and a clause denoting a result. Only if there are some structural criteria indicating that the clause denoting reason is special in some way and deviates from an unmarked simplex clause, the reason clause, rather than the result clause, may be singled out as defining the semantic relation that holds for the construction as a whole. In the example just given, the reason clause is syntactically subordinate, and formally it includes the relation marker *because*, and thus, the sentence can be taken to denote a reason relation.

This principle could have been consistently used for all constructions found in the data, but in practice, it was only relevant when separating constructions representing anteriority relations from those representing posteriority relations as well as reason relations from result relations (and to some extent, purpose relations from means relations, although means relations were not part of this study). For all other relations considered here, few or no constructions were marked for their opposite relations in the sample languages.

An outline of the procedure followed when classifying less transparent cases is in place here, since some constructions were less straightforward to classify semantically than others. For example, if asyndetic juxtaposition of clauses was reported to be conventionally employed for one or more semantic relations in a language, the relation of neither of the clauses to the other would be an obvious candidate for characterizing the relation holding for the construction as a whole, since neither of the clauses had an explicit relation marker, or a deviating structure. In those cases, the construction was semantically classified in accordance with the relation of the second clause vis-à-vis the first. The rationale behind this is that it is not until one gets beyond the first clause that one can interpret the relation between the two clauses.

In a few cases, constructions of asyndetically juxtaposed clauses were not classified in this way. This happened if one of the clauses was restricted
in its morpho-syntax, e.g. locked to a certain TMA form, which made it strongly associated with a certain relational category vis-à-vis the other clause. An example is given below from Rapanui.

112. Rapanui

ko oho 'a nua, a ia i tu'u mai ai
PFT go RES NAME PROP 3.SG PST come here ANA

'Nua had gone by the time he arrived.' (Du Feu 1996, p 50)

Both clauses in this example could be used independently, and the fronted subject of the second clause is not a consequence of its occurrence in a clause combining construction. However, when two clauses are juxtaposed in Rapanui clause combining, perfect tense marking in the first one is associated with an anterior state of affairs (Du Feu 1996) – any other TMA marker in the first clause would entail a co-occurrence interpretation. Thus, the construction above was classified as one of anteriority, since the restrictions occur in the clause denoting the anterior state of affairs.

Another potentially troublesome case pertains to constructions with clauses linked by coordinators, whether semantically explicit or not. Traditionally, coordinators are not regarded to be constituents of either of the clauses in clause combining, but many recent studies (e.g. Haspelmath 2004b) assume that coordinators universally are in constituency with one of the coordinands – most commonly the one it precedes – rather than extra structural elements. If there is no indication to the contrary, Austronesian coordinators in the present study have been analyzed to be in constituency with the clause they precede. That is to say, in Austronesian examples equivalent to I took my jacket for it was cold, for is regarded to be a reason relation marker since it precedes the clause denoting a reason for the state of affairs in the other clause. In a handful of cases in my sample, however, there was reason to analyze a coordinator to be in constituency with the preceding clause. The clearest example is found in Manam (an SOV language). This language has a coordinator that is unquestionably clause final. As reported in Lichtenberk (1983), it may be used enclitically to the first of two coordinate clauses but also (though only rarely) enclitically to the last of two coordinate clauses. It is the general coordinator be, 'and', which is often also used to express causation. Consider the examples below.

---

33 I regard for in this example to be a coordinator rather than an adverb because of its distributitional patterns, as it seems to pattern more closely with coordinators in English. Also, clauses linked by for have the same intonation pattern as do clauses linked by prototypical coordinators, such as and, but or or, while deviating from the intonation pattern of linked clauses of which the second has an initial coherence adverb, such as then.
4. SEMANTIC RELATIONS

113. Manam

a. *nóra* malípi né-gu *di-lába-be*
   yesterday work POSS-SG.GEN 3.PL.REAL-be.big-and

   *tágo* *u-púra*
   NEG 1.SG.REAL-come

   'Because I had a lot of work yesterday, I did not come.'
   or 'I had a lot of work yesterday, so I did not come.'
   (Lichtenberk 1983, p 524)

b. *ηáu* *ʔéu* ú-*n-i*
   1.SG dog 1.SG.REAL-hit-3.PL

   *bóro* né-gu *i-taotaon-i-be*
   pig POSS-1.SG.GEN 3.SG.REAL-chase-3.SG.ACC-and

   'I hit the dog because he was chasing my pig.'
   (Lichtenberk 1983, p 491)

Since the coordinator *be* can be shown to be enclitic to the previous clause, these constructions have been classified according to the relational properties of that clause. That is, both examples above were classified as involving reason relations.

Coherence adverbs or other optional morphemes were considered relation markers designating the semantic relation of the clause combining construction only if there were no obligatory explicit relation markers in either of the clauses. For instance, constructions analogous to the pattern in (114a) were classified as expressing posteriority, since the coherence adverb is the only relation marker present, and it marks the second clause as posterior in time, while constructions analogous to the pattern in (114b) were classified as expressing condition, as the subordinator in the first clause marks it as a condition, although the second clause contains the same coherence adverb as the second clause in the (a) example.

114. a. *We had a quick lunch, then we started building it.*

b. *If you tell him, then everything will be all right.*

Note, however, that coherence adverbs were taken to be the designator of the relation even if the other clause was structurally marked, as long as its marking did not indicate a relational meaning explicitly. That is, a construction including a subordinate nominalized reason clause and a result clause with a
coherence adverb was classified as a result construction rather than a reason construction. Amis provides an example below.

115. Amis

\[\text{pi-palu' ni aki ci panay-an,}\]
\[\text{?beat GEN NAME ACC NAME-ACC}\]
\[\text{sa ma-palu' ni mama cingra}\]
\[\text{so PF-beat GEN father 3.SG.TOP}\]

'Aki beat Panay, so he was beaten by father.' (Wu 1995, p 136)

The actor of the first clause, Aki, takes the genitive case marker indicating that the clause is nominalized. The prefix \textit{pi-} of the verb may contribute to the nominal flavor as well, although this prefix also has other functions (Wu 1995). Thus, even though the first clause is morpho-syntactically deviating, the construction was classified as a result construction, because the only semantically explicit relation marker occurs in the second clause.

In cases where both clauses were structurally identical, and both contained a semantically explicit relation marker, the relation of the first clause vis-à-vis the second was taken to hold for the entire construction, since the relation in such cases is established already in the first clause. There were a paucity of examples of this in the data, but one used in Tboli is provided below.

116. Tboli

\[\text{uni-hen l[em]wót, uni-hen m-ton du}\]
\[\text{RM-3.SG.GEN [AF]leave RM-3.SG.GEN AF-see 3SG.TOP}\]

'The minute he leaves he sees it.' (Porter 1977, p 132)

\[\text{34 An interesting fact about this example is that it appears to have originated from a construction that could more literally be translated 'The immediacy of his leaving (was) the immediacy of his seeing it.' As hinted at by Forsberg (1992, p 64), the relation marker \textit{uni} seems to stem from a noun meaning 'immediacy'. Note that the second clause verb has actor focus marking, even though the object is topic marked (which would normally have triggered object focus marking in Tboli; see Forsberg 1992). The locking of the verb to its AF form is a property of some types of complement clauses in several of the Formosan languages (both genealogically and areally close to Tboli) and would seem to suggest that \textit{mton} is (or historically was) the complement of \textit{uni}. In both clauses, the actor/experiencer of the verb is attracted to \textit{uni}, which in Tboli means they occur in the genitive form (regardless of the form they would have had otherwise).}\]
The first clause denotes a state of affairs that occurs immediately before the state of affairs that is denoted by the second clause, so the construction was classified as immediate anteriority.

4.2.2 Discourse levels of interpretation

Another issue potentially affecting the semantic classification of clause combining constructions is the level of discourse at which they may be interpreted. As has been discussed in several studies, semantic relations expressed by adverbial clauses, just as is the case for adverbials in general, may operate on various levels of discourse (Schiffrin 1987; Sweetser 1990; Hengeveld 1993; Crevels 2000b). In general, three levels\(^{35}\) are distinguished: the content level (fact-based), the epistemic level (knowledge-based), and the illocutionary level (speech-act-based). In many languages, the same relation marker may be used for a certain relation on all three levels, as illustrated by the following examples (from Sweetser 1990, p 77).

117. a. John came back because he loved her. CONTENT
    b. John loved her, because he came back. EPISTEMIC
    c. What are you doing tonight, because there is a good movie on. ILLOCUTIONARY

In the (a) example, the fact that John loved someone was the actual reason for his coming back to her. In the (b) example, however, John's coming back is the reason – not for his feelings towards his beloved, but – for the speaker coming to know of these feelings. In the (c) example, then, the fact that there is a good movie on is the reason for posing the question in the main clause, i.e. the relation has to do with the speech-act rather than its content. Note that if presuppositions are explicitly indicated in epistemic and illocutionary level constructions, the rationale behind their relational coding becomes clear. In the cases of the reason clauses in (117b) and (117c), they denote the reason for the presupposed segment, as indicated by example (118) below.

118. a. John loved her, (I know it) because he came back.
    b. What are you doing tonight, (I'm asking) because there is a good movie on.

Also in many of the Austronesian languages, the same relation marker may be used at different levels of discourse. Consider, for instance, the two ex-

\(^{35}\) Some studies, however, have introduced additional levels (e.g. Hengeveld 1993; Crevels 2000b).
amples from Central Cagayan Agta below, intended to be interpreted at the content level and the epistemic level, respectively.

119. Central Cagayan Agta

a. *nag-bilag na hapa ya ugta-en*
   AF-PST-run now also TOP deer-DEF

   *te ne-tappalān na atu-en*
   because TF-bump GEN dog-DEF

   'Now, the deer ran too, because the dog bumped into him.'
   (Mayfield 1972, p 61)

b. *e āk sangaw mag-kalimag kun-na*
   go 1.SG.TOP later AF.NPST-cut.grass say-3.SG

   *ottuhu nag-babāwi te um-an ge nag-lagap*
   then AF.PST-change.mind because AF-go AF.PST-fish

   '"I'm going to cut grass later on", he said and then changed his mind because he went fishing.' (Mayfield 1972, p 41)

Austronesian constructions with relation markers functioning on various levels of discourse have been classified only according to the relation expressed at the content level.

4.2.3 Some terminological notes with regard to temporal ordering

The terms 'anteriority' and 'posteriority' indicate that they refer to something that comes before (i.e. anterior to) and after (i.e. posterior to) something else, respectively. However, when one is labeling subordinate clauses expressing anterior or posterior states of affairs with respect to the main clause, confusion easily arises. For instance, when used in a clause combining construction, the meaning of the subordinator *after* as 'later in time' applies to the main clause predicate, since *after* (and its complement) is a subordinate constituent of that predicate. In other words, it is the state of affairs of the main clause that occurs 'later in time', posterior to something else, while the complement of *after*, the subordinate clause, denotes this "something else" that occurs anterior in time. The logical consequence of this is to refer to the subordinate clause marked by the subjunction *after* as an anteriority clause, although, indeed, the meaning of the subjunction does clash with the label for the relation. The same is true of the subordinator *before*, meaning 'earlier...
in time' but marking posteriority. It is a part of the main clause predicate describing its state of affairs as anterior in time, while its complement depicts what is posterior. This is the labeling convention that has been used in the present study.

Despite adopting it for use, I am well aware that this labeling might seem misleading. However, using 'anteriority relations' and 'posteriority relations' in this way are consistent with the use of other terms – such as 'terminal boundary' for until-clauses and 'result' for so (that)-clauses. The clauses introduced by until and so (that) denote the terminal boundary and the result state of affairs, respectively, of these relations. Thus, in employing the method used in this study for classifying clause combining constructions semantically by which each construction is labeled by the semantic relation denoted by the clause with the distinctive marking (see section 4.2.1 above), it follows logically that after-constructions should be classified as denoting anteriority relations and before-constructions should be classified as denoting posteriority relations, just as until-constructions are classified as denoting terminal boundary relations.

This procedure also makes comparison between subordinate constructions and coordinate constructions more straightforward and consistent. Compare, for instance, the two constructions of temporal ordering in (120) and the two constructions of causality in (121) below in which the (a) constructions involve coordination, and the (b) constructions involve subordination.

120. **ANTERIOR > POSTERIOR**
   a. *I ate a sandwich and then I went to bed.*
   b. *I ate a sandwich before I went to bed.*

121. **REASON > RESULT**
   a. *The traffic was heavy so he was late to the meeting.*
   b. *The traffic was heavy so that he was late to the meeting.*

If the second clause of each sentence in (121) is classified as result semantically, i.e. regardless of whether the subordinator so that or the coordinator/adverb so is used, then the second clause of each sentence in (120) must be classified as posteriority.

But this is a matter of some controversy and linguists obviously do not agree on the subject. Kortmann (1997) and Cristofaro (2005), for instance, use the terms anteriority and posteriority in the same way as I do in this study, while Haspelmath (1997b) uses them in the opposite way.
4.2.4 Semantic relations reference map

In concluding section 4.2, I would like to organize the ten semantic relations investigated here along two cross-cutting axes: (i) level of abstraction, and (ii) positional characteristics. The following categories on each axis can be identified:

LEVELS OF ABSTRACTION (from less to more abstract):
- space
- time
- co-variation

POSITIONAL CHARACTERISTICS:
- co-occurrence = same position
- ordering = different position, relative order
- boundary = different position, relative order, limitation

For the purpose of illustrating some connections between sets of relations in further detail (see chapter 5), I include space on the axis indicating levels of abstraction, although these relations are not the principal subject of investigation in the present study. The resulting table (Table 11) will make up a useful reference map of the relevant relations.
Table 11. Reference map of semantic relations

<table>
<thead>
<tr>
<th>LEVEL OF ABSTRACTION</th>
<th>CO-VARIATION</th>
<th>TIME</th>
<th>SPACE</th>
<th>SAMENESS</th>
<th>ORDERING</th>
<th>BOUNDARY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• concession (although)</td>
<td>• co-occurrence (when, while)</td>
<td>• spatial same-ness (at, with)</td>
<td>• back (in front of)</td>
<td>• front (behind)</td>
<td>• goal (to)</td>
</tr>
<tr>
<td></td>
<td>• condition (if)</td>
<td>• posteriority (before)</td>
<td>• anteriority (after)</td>
<td>• terminal boundary (until)</td>
<td>• initial boundary (since)</td>
<td></td>
</tr>
</tbody>
</table>

POSITIONAL CHARACTERISTICS
The dotted lines in the table are meant to indicate that the relations on each side of the division are mirror relations. For instance, two states of affairs, one occurring before the other in time, could in a clause combining construction principally be coded either for posteriority, with an explicit marker on the clause representing the posterior state of affairs, or for anteriority, with an explicit marker on the clause representing the anterior state of affairs. It will be evident in the next section (4.3) that for most semantic relations, Austronesian languages – as perhaps languages in general – normally select one and the same of two related states of affairs for specific coding in clause combining. For instance, it is always the clause representing the terminal boundary that receives coding, and not the clause representing the state of affairs that is terminated. The two exceptions noted in the data are posteriority-anteriority (as mentioned) and reason-result, also illustrated in the table.

Naturally, the table is a generalization and a simplification for the sake of displayability. For instance, some of the relations in the table are distinguished from other relations in more complex ways than indicated by the parameters in the shaded cells. Concession and condition are two examples in that the former is characterized by counter-expectancy and assertion (something did happen in spite of not being expected), while the latter is characterized by hypothesis (under the hypothesis that something is true, something else is also true). None of these parameters figures in the table. Furthermore, not all co-variational relations in the table are connected to their temporal and spatial counterparts in the same way, which the table can be taken to indicate. Purpose relations, for instance, share the characteristic of a destination with terminal boundary relations and goal relations – their temporal and spatial counterparts. In a similar way, reason relations share the characteristic of an origin with initial boundary and source. But while purpose has the additional modal component of intentionality, this is lacking for reason relations. Nonetheless, Table 11 is useful as a reference map, and as we shall see in chapter 5, it also reflects some diachronic patterns of development.

4.3 Relational categories in the Austronesian languages

In this section, the semantics of the relations investigated will be discussed, as well as shades of meaning attested in the sample languages. Several Austronesian examples will be given.
4. SEMANTIC RELATIONS

4.3.1 Co-occurrence

Co-occurrence relations hold in clause combining constructions in which at least one of the clauses is explicitly marked as denoting a state of affairs co-occurring in time with the state of affairs of the other clause. Co-occurrence is taken to include complete or partial temporal overlap between states of affairs as well as temporal connections between states of affairs occurring within the same temporal setting (for further details, see below). Co-occurrence relations are also taken to hold in clause combining constructions in which the semantic relation is implicit but commonly interpreted to denote states of affairs co-occurring in time. Austronesian examples of constructions representing a co-occurrence relation are provided from Central Cagayan Agta and Kambera.

122. Central Cagayan Agta

\[ mag-adadua \text{ } \bar{a}k \]
\[ \text{NPST.AF-vomit } \text{1.SG.TOP} \]

\[ y\tilde{a}ga \text{ } mag-atattay \text{ } \bar{a}k \text{ } ta \text{ } mabasa \]
\[ \text{COOC } \text{NPST.AF-defecate } \text{1.SG.TOP } \text{OBL } \text{wet} \]

'I vomited and at the same time I had diarrhea.'
(Mayfield 1972, p 43)\textsuperscript{36}

123. Kambera

\[ d\text{edi}, \text{ } \emptyset \text{ } meti-ma-a-nanya \text{ } na \text{ } ina-na \]
\[ \text{be.born } \text{die-EMP-B-3.SG.CNT } \text{ART } \text{mother-3.SG.GEN} \]

'(When he) was born, his mother died.' (Klamer 1998, p 281)

Constructions commonly or exclusively used to express a co-occurrence relation are attested in all but one of the sample languages.\textsuperscript{37} There was some variation in the semantics of the co-occurrence constructions in the data. In

\textsuperscript{36} Non-past tense, which is used in the example, is occasionally used in Central Cagayan Agta instead of narrative past tense to add vividness (cf. Healey 1960, p 20), thus, the translation into past tense.

\textsuperscript{37} The only language for which a simultaneity construction is not attested is Iwal (North New Guinea) from the extended sample. My source, however (Bradshaw 2001), mentions a coordinator be, 'and', with several uses, of which he exemplifies temporal succession only. It seems very likely that it can also be used for temporal co-occurrence given the relevant context.
order to present a more detailed picture of what was counted as a co-
occurrence relation, the most common varieties are presented in the re-
mainder of this subsection.

Several Austronesian languages have relation markers expressing tem-
poral co-occurrence between states of affairs over an extended period of time
(normally in addition to relation markers expressing co-occurrence in a more
general sense). In Samoan, for instance, the relation marker 'a'o, 'while', is
used when a continuous durative temporal interpretation is intended for both
states of affairs. If the clause following the relation marker denotes a state of
affairs that is punctual or very short in temporal extension, or else easily
interpretable as such, ina 'ua is preferred instead. If 'a'o is used in such cas-
es, it imposes a durative reading on the clause following it (it may, for in-
stance, make a punctual state of affairs iterative). Compare the two examples
below.

124. Samoan

a. sa faitau la'u tusi 'a'o tatagi le telefoni
   PST read 1.SG.POSS book while ring the telephone
   'I was reading the paper while the phone rang.' (AJ)

b. sa faitau la'u tusi ina 'ua tatagi le telefoni
   PST read 1.SG.POSS book when INCH ring the telephone
   'I was reading the paper when the phone rang.' (AJ)

The preferred reading of the (a) example is that the phone call was never
answered, while the preferred reading of the (b) example is that it was.38

General co-occurrence markers are more sensitive to contextual informa-
tion and the propositional content of the clauses, and are also often used
even when the states of affairs co-occur without a temporal overlap but within
the same temporal setting. In these cases, rather than indicating simultane-
ity, the marked clause normally provides a general framework relevant for
the other clause, and the state of affairs takes place slightly anterior to that of
the other clause but in close connection to it. Palauan provides an example.

38 This is very similar to the distinction between while and when in English, but while while
and when are synonymous when connecting two durative clauses, this is not quite the case
when 'a'o and ina 'ua connect two durative clauses (apparently because of the inchoative
meaning of 'ua).
125. Palauan

\[
\text{ak } \text{milechêr} \text{ar } a \text{ hong} \\
1.\text{SG} \text{ buy.PST} \text{ PHR book}
\]

\[
\text{er } \text{se } \text{êr } a \text{ kbo } \text{êr } a \text{ stoang} \\
\text{at that at PHR go.PST to PHR store}
\]

'I bought a book when I went to the store.' (Josephs 1975, p 445)

Since many languages seem to have the option of using a general co-occurrence marker in this way (also in languages where there is no specific durative co-occurrence marker), I have made no distinction between co-occurring overlap and co-occurring sequences in cases in which the same relation marker is used for both.

Another type of co-occurrence relation, which is sometimes represented by a specific construction in the sample languages, is a relation that is often called 'generic co-occurrence' or something similar. Kortmann (1997) calls it 'contingency'. In this study, I have used the term 'indefinite time' for this relation, and classified it as a subtype of co-occurrence relations. In English, it is often expressed by the subordinator whenever as in Whenever I see that movie, I cry. In some Austronesian languages, as in English, the relation marker denoting indefinite time is modeled on a relation marker denoting general co-occurrence. This is the case in Lenakel. One of the relation markers for general co-occurrence in this language is derived from the noun for 'time'. To express indefinite time, it is used in the plural.

126. Lenakel

a. \[
\text{nian } \text{io } \text{apwa } \text{lenakel, pwia-k } \text{r-va} \\
\text{time 1.SG LOC Lenakel older.brother-1.SG 3.SG-come}
\]

'While I was in Lenakel, my older brother came.' (Lynch 1978, p 108)

b. \[
\text{nian } \text{miin } \text{i-ak-vin } \text{iifila,} \\
\text{time PL 1.SG-PRES-go Vila}
\]

\[
\text{pukas } \text{taha-k } \text{r-ak-mis} \\
\text{pig POSS-1.SG 3.SG-PRES-be.sick}
\]

'Whenever I go to Vila, my pig gets sick.' (Lynch 1978, p 108)
CLAUSE COMBINING IN AUSTRONESEAN LANGUAGES

Less common in terms of their semantics but also classified as representing co-occurrence are, for instance, the Buru constructions involving the relation markers *bama* and *mama*, and the Muna constructions involving *paka-* and *kirakira*. In Buru, *bama* and *mama* both express co-occurrence, but while the former also implies continuity with a prior discourse context (~’and when… then…’), the latter implies discontinuity (~’but when… then…’). The meaning of *paka-* and *kirakira* in Muna can be illustrated by the examples below.

127. Muna

a. *paka-gaa-ndo sadhia do-pogira*
   
   INCOOC-marry-3.PL.POSS always 3.PL.REAL-fight

   'When they were just married, they were always fighting.'
   (Berg 1989, p 246)

b. *kirakira40 no-maho-mo*
   
   about.to 3.SG.REAL-nearly.happen-PFTV

   *na-rumako-da patu*
   
   3.SG.IRR-catch-3.PL bamboo

   *garaa no-tumbu-mo patu*
   
   suddenly 3.SGEAL-grow-PFTV bamboo

   'When she was about to catch them, suddenly there grew a bamboo bush.' (Berg 1989, p 248)

In the (a) example, the marked clause specifically indicates that the initial stages of the state of affairs co-occur with the state of affairs denoted by the other clause. In the (b) example, the marked clause indicates that the state of affairs is just about to materialize when the state of affairs of the other clause occurs (and intervenes).

---

39 These relation markers can also be used in conditional constructions, with the same continuity/discontinuity meanings (see section 4.3.6).

40 The word *kirakira* cannot be used in simplex clauses; it must always occur in clause combining constructions. It often co-occurs with the verb *maho* (‘nearly happen’), but not obligatorily so (Berg 1989).
4.3.2 Posteriority

Posteriority relations hold in clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs occurring posterior in time to the state of affairs of the other clause. Posteriority relations are also taken to hold in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause, or if structurally identical, the second clause, is conventionally interpreted as denoting a state of affairs occurring posterior in time to the state of affairs of the other clause. Austronesian examples are provided below.

128. Erromangan

pehnur-i kakle-nteihep ra nmap
before-3.SG 1.PL.EX.FUT-throw.down to ground

kakeml-antipe nulgo-n uloki
1.PL.EX.PRES-put.down leaf-3.SG.POSS coconut

'Before we throw it down into the ground, we put down the coconut leaves.' (Crowley 1998, p 275)

129. Hoava

[...] ada ria tige buki-a ria sa buki
wake 3.PL then blow-3.SG 3.PL ART.SG conchshell

'...they woke up, then they blew the conchshell.' (Davis 2003, p 262)

The last one of these examples represents a very common way of manifesting posteriority relations in the sample languages, which can be called temporal succession. Such constructions are characterized by being obligatorily linear-iconic in terms of temporal order, so that the second clause is always interpreted as the posterior one. Normally (but not necessarily), these constructions involve coordination syntactically. Linear-iconic constructions can consist of asyndetically juxtaposed clauses, but often a relation marker is present, renderable as 'and', 'and then', or 'then', etc.

In some Austronesian languages, a relation marker meaning 'not yet' or similar is used to indicate posteriority. The inference is that the state of affairs indicated as not yet having taken place will or did take place at some later point. Buru provides an example.
130. Buru

\[ da \ mata \ \textit{mohede}, \ da \ stori \ gam \ naa \ [...] \]

3.SG  die  not.yet  3.SG  speak  like  this

'BBefore he died, this is what he said...' (Grimes 1991, p 421)

In addition to constructions denoting general posteriority, some of the sample languages also have posteriority constructions with relation markers specifically expressing immediate posteriority. In Taba, the relation marker \textit{turus}, a loanword from a nearby Malay variety, expresses that the posterior state of affairs occurs immediately after the other state of affairs, or after a very short interval of time.

131. Taba

\[ ulan \ kwat \ \textit{turus} \ ni \ kihis \ n-sopang \]

rain  EMPH  POST  3.SG.POSS  flood  3.SG-descend

'There was strong rain and straight away a flood descended.'

(Bowden 1997, p 438)

Constructions commonly or exclusively used to express a posteriority relation are attested in all the sample languages.

4.3.3 Anteriority

Anteriority relations are expressed by clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs occurring anterior in time to the state of affairs of the other clause. Anteriority relations are also taken to be present in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause is conventionally interpreted as denoting a state of affairs occurring anterior in time to the state of affairs of the other clause. Anteriority relations are the opposite of posteriority relations in that an anterior state of affair presupposes a posterior state of affairs, and vice versa, regardless of which relation is marked in clause combining. Anteriority constructions are exemplified from Tagalog and Western Subanon below.

---

41 The Indonesian cognate is \textit{terus}, 'directly' (Interactive dictionary, Indonesian-English, Northern Illinois University).
4. SEMANTIC RELATIONS

132. Tagalog

'a-alis  kami  
RD:AF.IRR-leave  1.PL.EX.TOP

ma-tapos na k[um]ain ng hapunan ang mga bata  
AF-finish COMP [AF]eat GEN lunch TOP PL child

'We'll leave after the children eat lunch.'  
(Schachter & Otanes 1972, p 475)

133. Western Subanon

pok-polaus ni pitandang dun koyon  
ANT-take.out POSS NAME 3.SG.OBL that

b[in]otang non sog glupaq 42  
put[FOC.TMA] 3.SG.POSS to dirt

'Pitandang having taken that one out, he put it on the ground.'  
(Hall 1973, pp 6-7)

For linear iconic reasons, which are naturally important with temporal relations, there are no examples of implicit clause combining constructions with structurally identical clauses for which the second is regularly interpreted as representing the anterior state of affairs. However, there are clause combining constructions in which neither clause deviates from a regular simplex clause pattern, which still have been considered to involve an anteriority relation. This pertains to cases where one of the clauses has restrictions on some morpho-syntactic variable. When the clause with restrictions – usually to some perfect/perfective TMA category – is conventionally interpreted to denote an anterior state of affairs in relation to the state of affairs of the other clause, it has been classified as representing an anteriority relation; cf. the example from Rapanui below (reproduced from example 112 above).

134. Rapanui

ko oho 'a nua, a ia i tu'u mai ai  
PFT go RES NAME PROP 3.SG PST come here ANA

'Nua had gone by the time he arrived.' (Du Feu 1996, p 50)

42 The morphematic transcription is a little uncertain in this example, since Hall (1973) does not indicate morpheme boundaries.
Similarly to the use of 'not yet' to indicate posteriority (see previous section), equivalents to the adverb 'already' is sometimes used in Austronesian languages to indicate anteriority. Note the example from Tetun below.

135. Tetun

\[ kawen \ ti'a, \ tur \ iha \ ne'e \ dei \]
mARRY \ already \ sit \ LOC \ this \ only

'After (we) are married, (we) must live here.' (Klinken 1999, p 236)

Just as many European languages (see e.g. Kortmann 1997), some Austronesian languages have constructions in which the relation marker indicates that the anterior state of affairs immediately precedes the other state of affairs. Note the Acehnese example below.

136. Acehnese

\[ ban \ ji \ teu-pue \ na \ gopnyan \ di \ rumoh \]
just \ 3.SG \ know-what \ be \ 3.SG \ at \ house

\[ ji \ jak \ lê \ keu \ dêh \]
3.SG \ go \ then \ to \ there

'As soon as he knew that he was at home, he went there immediately.' (Durie 1985, p 259)

Sangir has an anteriority relation marker indicating close, though not immediate, anteriority. That is, although there is a time interval between the states of affairs, it is indicated to be very short.

137. Sangir

\[ buhèng \ télu-wullang \ i \ kami \]
soon.after \ three-month \ TOP \ 1.PL.IN

\[ mésé-sikollà-e, / ni-songong-ke-ng \]
FOC.TMA-attend.school-now \ FOC.TMA-invade-now-TOP

\[ japaéng \ su \ manaro \]
Japanese in Manado

'Shortly after we had been attending school three months, we were invaded by the Japanese in Manado.' (Maryott 1979, p 190)

43 The -ng is a clitic topic marker indicating the following noun phrase as a topic. It is used after vowel-final words (Maryott 1977, p 108). Other than for -ng, the placement of morpheme boundaries is a bit uncertain for this word, since Maryott (1979) does not indicate them in his example.
Some of the sample languages lack a specific anteriority construction, preferring instead constructions of temporal succession (linear-iconic) to express that two states of affairs are temporally ordered (see previous section, 4.3.2). The overwhelming majority of constructions of the temporal succession format have been classified as manifestations of a posteriority relation, either because an explicit relation marker occurs in the clause denoting the posterior state of affairs, or, when a relation marker is lacking, because the relational reading of the construction is not apparent until the second clause, expressing the posterior state of affairs, is reached. The only exceptions are those few cases where a neutral coordinator could be decided to occur in constituency with the first of two linear-iconic clauses. They were instead classified as constructions of anteriority since the distinctive marking could be argued to occur in the anteriority clause. Manam is a case in point.

138. Manam

\[
\text{áine moarépi } i-pási-Ø-be \\
\text{woman sweet.potato 3.SG-dig.out-3.PL.OBJ-and} \\
\text{di-moamoá-mwa-Ø} \\
\text{3.PL.REAL-roast-RD-3.PL.OBJ}
\]

'The woman had dug out the sweet potatoes, and then they were roasting them.'

or 'After the woman had dug out the sweet potatoes, they were roasting them.' (Frantisek Lichtenberk, p.c. 2010)

As discussed earlier (section 4.2.1), the coordinator -be can be shown to be in constituency with the clause it follows (Lichtenberk 1983). Constructions specifically used for anteriority relations are attested in 25 of the 30 core sample languages and in 32 of the 43 extended sample languages.

4.3.4 Terminal boundary

Relations of terminal boundary are expressed by clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs that defines the end point of the state of affairs of the other clause. Terminal boundary relations are also taken to be present in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause, or if structurally identical, the second clause, is conventionally interpreted to denote a state of affairs ending the progression of the state of affairs of the other clause. Examples are provided below.
139. Rapanui

\[ he \ \etaatu \ hau \ rima \ ata \ ka \ hiohio \ ro \]
ACT press INSTR hand until MOM be.hard REAL

'You work it with your hands until it is quite hard.' (Du Feu 1996, p 51)

140. Big Nambas

\[ a-v-rp-i \ \ da-v'a \ ti \ i-valau \]
3.PL.REAL-PL-hit-3.SG TMA-go that 3.SG.REAL-cry

'They hit him until he cried.' (Fox 1979, p 87)

141. Muna

\[ ne-late \ bhe \ awa-no \ ini-a \]
3.SG.REAL-live with grandparent-3.SG.POSS this-CL

\[ ta-no-bhala-hi-mo \]
TB-3.SG.REAL-be.big-HI-PFTV

'He lived with his grandmother until he had grown up.' (Berg 1989, p 246)

A substantial portion of the attested terminal boundary relation markers are clearly derived from verbs (typically meaning 'go', 'reach', 'arrive', 'be sufficient'), which have grammaticalized to a greater or lesser extent (cf. the Big Nambas example above).

The semantic differentiation among terminal boundary relation markers is not very great. Some languages, however, have special relation markers to indicate that the length of time to the terminal boundary is possibly extended. This is the case in Rapanui, for instance. Compare example (142) below with example (139) above.

\[ \]

---

44 The -\(a\) is a clitic that in this context indicates that the speaker has not yet finished his sentence (cf. Berg 1989, p 269).

45 The meaning of -\(hi\) is difficult to pinpoint, but it might be a nominalizing morpheme in this context (cf. Berg 1989, p 284).
142. Rapanui

\[ ki \ tiaki \ 'ahar\dot{a} \ ki \ tu'u \ mai \ a \ nua \]
MOM wait until.EXT MOM come DIR PROP Nua

'Let's wait until Nua comes.' (Du Feu 1996, p 51)

The use of the relation marker 'ahar\dot{a} here indicates that the wait might be long and that it is not even certain that Nua will eventually turn up (Du Feu 1996).

In many Formosan languages (see e.g. Holmer 2006), there is a construction worthy of special mention in connection with terminal boundary relations; among the sample languages, it occurs in both Amis and Seeqiq. The construction consists of two juxtaposed clauses in which certain morpho-syntactic restrictions are imposed. First, the verb of the second clause must occur in the active focus voice form, with other voice forms resulting in ungrammaticality. Furthermore, the actor participant must be interpreted to be identical in both clauses, and can only be overtly expressed in one of the clauses, the first or the second, but not in both. Compare the two examples below.

143. Amis

a. \[ ma\text{-}fucal \ tu \ ci \ aki \ k[um]aqen \ tu \ tali \]
AF-full INCH NOM Aki eat[AF] ACC taro

'Aki ate taros until he became full.' (Liu 2003a, p 63)

b. *\[ ma\text{-}fucal \ ci \ aki \ ma\text{-}kaqen \ ku \ tali \]
AF-full NOM Aki PF-eat NOM taro

Intended: 'Aki ate taros until he became full.' (Liu 2003a, p 64)\(^{47}\)

This construction is actually used most commonly for manner expressions – in cases where languages like English use adverbs – with the first clause expressing the manner in which the state of affairs of the second is carried out, as exemplified in (144).

\(^{46}\) The construction comes in several variants with slightly different structural and semantic properties in the different languages (see e.g. Chang 2006a; Holmer 2006; Liu 2003a; Wu 1995), but for our purposes, a broad brush description will suffice.

\(^{47}\) The presence or absence of the inchoative aspect marker has nothing to do with the (un)grammaticality of these two examples (Liu 2003a).
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

144. Amis

mi-palifud tu ci aki mi-palu’ ci kacaw-an
AF-violent INCH NOM Aki AF-hit ACC Kacaw-ACC

'Aki was violent to Hit Kacaw.' (Liu 2003a, p 68) or
'Aki hit Kacaw violently.'

For this reason it seems natural to interpret the first clause as the adverbial clause and significant for semantic classification. However, the restrictions in morphology, locking the verb to its actor focus form, applies to the second clause, not the clause representing the terminal boundary. This leads Chang (2006a) to analyze the second clause of a similar construction in the Formosan language of Kavalan as dependent on the initial verb rather than the other way around. In actual fact, however, these Formosan language constructions are the only ones found in the sample languages for which the clause representing a terminal boundary is structurally less marked than the other clause. So, in spite of their structural marking, it seems natural to group them together with the other terminal boundary constructions semantically, rather than in a group of their own. A further complication lies in the fact that in all the examples found of this construction in which a terminal boundary interpretation is possible, the context is such that the state of affairs encoded in the second clause is also the reason for that encoded in the first. Reason and terminal boundary relations are closely intertwined in these constructions. Therefore, I have classified them both as representing a terminal boundary relation and a reason relation in this study. A reason relation was chosen (rather than a result relation) because the structurally marked clause – the second one – represents a reason. A terminal boundary relation was chosen (rather than an opposite relation), in spite of the fact that the structurally unmarked clause – the first one – represents a terminal boundary relation, for practical reasons, since the Formosan constructions were the only ones of this type.

In the core sample, constructions exclusively or commonly associated with relations of terminal boundary were attested in 22 of 30 languages, while in the extended sample, they were attested in 31 of 43 languages.

---

48 Inferring a terminal boundary relation for constructions explicitly conveying a reason or result relation is probably quite common. So, a justified question is: why not classify those constructions as both terminal boundary and reason or result as well? However, in those cases, the terminal boundary relation is based on inference, while the reason or result relation is explicit, whereas in the relevant Formosan language constructions, neither the terminal boundary relation nor the reason relation seem to take precedence over the other, which justifies a double semantic classification in the latter case but not in the former.
4.3.5 Initial boundary

Relations of initial boundary are expressed by clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs that defines the starting point of the state of affairs of the other clause. There are no examples of implicit clause combining constructions involving an initial boundary relation in the data.

The initial boundary relation is a counter-relation to the terminal boundary relation, since it defines a starting point instead of an end point to some state of affairs. However, initial boundary and terminal boundary are not opposite relations in the sense that posteriority and anteriority are, since the one does not presuppose the other. The initial boundary state of affairs constitutes the start of another state of affairs, and that other state of affairs is of course not a terminal boundary in relation to the initial boundary. But if some state of affairs is anterior to another, that other state of affairs is by necessity posterior in relation to the anterior one.

In general, grammaticalized constructions coding initial boundary appear to be rather uncommon in Austronesian languages. It is probably the case that, in quite a few Austronesian languages, initial boundary is expressed using lexical items in various constructions instead of grammaticalized items in specialized and designated constructions. Such constructions are not always represented in descriptive grammars. The relation markers that indicate initial boundary are often derived from verbs meaning 'begin' and 'start'. Some examples of Austronesian initial boundary constructions found in the data are given below.

145. Tetun

\[ \text{aikanoik né hori rai moris n-ó kedas} \]

\text{story this since earth live 3.SG-exist immediately}

'This story has existed since the earth came into being.'

(Klinken 1999, p 163)

146. Coastal Konjo

\[ \text{mulai ri si-itte-na i-Ali na i-Baco' ri bira} \]

\text{begin at REC-see-3.POSS PROP-Ali and PROP-Baco' at Bira}

\[ \text{a'-kai' lantang-i} \]

\text{INTR-connect deeply-3.ABS}

'Since Ali and Baco met in Bira, they have been fast friends.'

(Timothy Friberg, questionnaire)
147. Tagalog

\[ \text{mula nang na-matay ang awawa niya} \]
start.from when PFTV.AF-die TOP wife 3.SG.GEN

\[ \text{hindi na siya na-ka-gawa ng trabaho} \]
be.not already 3.SG.TOP AF.PFTV-?-make GEN work

'Since his wife died, he hasn't been able to do any work.'
(Schachter & Otanes 1972, p 476)

In the Coastal Konjo example, the verb stem \textit{mulai}, 'begin', occurs without any argument markers, and in the Tagalog example, the cognate verb stem \textit{mula}, from \textit{magsimula}, 'start from' (Large Tagalog dictionary, online resource, Northern Illinois University), occurs without its focus marker, indicating that these verbs have traveled some distance on the path of grammaticalizing into relation markers.

Constructions expressing initial boundary are attested in only 7 of the 30 languages of the core sample as well as in an additional language of the extended sample, i.e. in only 8 of the 43 languages.

4.3.6 Condition

Conditional relations are expressed by clause combining constructions in which one of the clauses is explicitly marked as denoting a hypothetical state of affairs representing a condition on which the state of affairs of the other clause is dependent for its realization, as in \textit{If you brush your teeth, I'll read you a bedtime story}. Although the states of affairs are often in sequence and imply a consequential link between them (as in the example just mentioned), a conditional relation in a logical sense does not include the component of sequentiality, nor consequentiality, since the fulfillment of the condition need not bring about the realization of the other state of affairs, neither directly nor indirectly; cf. for instance, the sentence \textit{If a language has trial grammatical number, it also has dual grammatical number}. It merely states that in case the conditional clause is (was or will be) true, so is (was or will be) the other clause.

In logic, the conditional clause is called the protasis, while the accompanying clause is often called the apodosis. In the sample languages, the protasis is most often explicitly marked as a condition, although some constructions (optionally or obligatorily) have marking on both clauses, indicating both the protasis and the apodosis, respectively (cf. 'if...then' in English). Very few languages have explicit relational marking on the apodosis clause
alone. However, there are examples of asyndetic constructions commonly used with a conditional meaning (cf. You come one step closer, I'll shoot in English) as well as constructions with a neutral coordinator used in the same way (cf. You come one step closer and I'll shoot). Since these always have the conditional clause first, they should perhaps be classified by the last clause as a kind of apodosis relation rather than a conditional relation, either in keeping with the method of classifying asyndetic constructions by the relation of the last clause to the first, or by the coordinator being in constituency with the last clause. Nonetheless, I have classified all such examples as instances of conditional relations. There are basically two reasons for this: (i) In all Austronesian constructions of this kind, the conditional clause is strongly perceived of as the topic of the construction (as are conditional clauses in general, cf. Haiman 1978), and is therefore, discourse-wise, marked as the conditional clause. And (ii), although prosodic information for specific constructions is scant in most grammars, it seems likely that, in Austronesian, as in English and other European languages, non-explicit constructions with a conditional reading of the first clause, as well as those with an explicit marker for apodosis, the conditional clause is prosodically marked, reflecting its topical status.

Constructions used for conditional relations are attested in all sample languages. Some examples are provided below for illustration. Note that in the Mbula example, both the protasis and the apodosis are marked, while in the Mekeo example, neither the protasis nor the apodosis is marked.

148. Araki

\begin{verbatim}
aru mo usa, co pa re ai
\end{verbatim}

if 3.SG.REAL rain 3.SG.IRR then be.some water

'It it rains, we will have water.' (François 2002, p 178)

149. Mbula

\begin{verbatim}
sombe ti-posop uraata, so aŋ-giimi zin \footnote{Both relation markers in this example are derived from the verb for 'say', which is -so. When used as a conditional clause marker, as the initial element of the clause, it has lost the (typically verbal) ability of being prefixed by a pronominal subject marker, though it is often suffixed by 'mbe' or its allomorph 'be', analyzed variously by Bugenhagen 1995 as a non-factual auxiliary (p xi), a non-factual complementizer (p 157) and a non-factual adverb (p 159). When so is used as a marker for the apodosis, it indicates counter-factuality, and cannot take any morphology (Bugenhagen 1995, p 276). The restrictions in morphology suggest that so((m)be) in these cases has grammaticalized to relation markers to some extent.}
\end{verbatim}

\begin{verbatim}
if 3.PL-finish work then 1.SG-buy 3.PL.ACC
\end{verbatim}

'If they had finished work, I would have paid them.' (Bugenhagen 1995, p 277)
150. Mekeo (northwest dialect)

\[
\begin{align*}
iu & \quad \text{agia-u} & \quad n\text{-}io & \quad \etaa, \\
1.\text{SG} & \quad \text{friend-1.SG.POSS} & \quad 3.\text{SG.FUT-go} & \quad \text{ASS}
\end{align*}
\]

\[
\begin{align*}
iu & \quad \text{aga} & \quad na\text{-}io \\
1.\text{SG} & \quad \text{too} & \quad 1.\text{SG.FUT-go}
\end{align*}
\]

'If my friend goes, I will go too.' (Jones 1998, p 514)

In the Mekeo example, the future tense marking accompanied by a high rising tone on the verb of the first clause strongly suggests a conditional interpretation (Jones 1998, p 514). The assertive particle \(\etaa\) seems to contradict the conditional reading but instead signals a request for acceptance from the hearer of the clause as a premise. Jones (1998) describes the underlying proposition as "It is a fact that X will happen: OK? – (then) Y will happen".

Traditionally, it is common to distinguish conditional relations along three axes: (i) whether they express open or counter fact conditions, (ii) whether they express concessive or non-concessive conditions, or (iii) whether they express affirmative or negative conditions. Although all these distinctions are not attested in all languages, most languages have distinct constructions for at least two of them, either by designated grammatical marking in the conditional clause (such as inflections or particles also used in other contexts) or specific conditional markers, or a combination of the two. These various conditional meanings, and some less common ones, will be described and exemplified for the Austronesian languages in the following subsections.

4.3.6.1 Open and counterfactual conditionals

The difference between open conditionals and counterfactual conditionals has to do with truth-value of the state of affairs coded as a conditional clause. In both cases it is hypothetical, but for open conditionals it is pending, assumed but not known, while for counterfactual conditionals it is known to be false, assumed for the sake of argument. Open conditionals are about states of affairs in the past, present or future, which could potentially happen (or potentially have happened); counterfactual conditionals are about states of affairs in the past or present, which are known not to have happened or not to be happening, but are put forth as a hypothesis. Compare the two Samoan sentences below of which (a) is an open conditional and (b) is a counterfactual conditional.
151. Samoan

a. ole’ā 'ata le tama
FUT laugh the boy

pe 'āfai va'ai mai ia te a'u
Q if see DIR to PROP 1.SG

'The boy will laugh if he sees me.' (AJ)

b. semanū e 'ata le tama
probably GNR laugh the boy

pe 'ana va'ai mai iā te a'u
Q if see DIR to PROP 1.SG

'The boy would probably have laughed if he had seen me.' (AJ)

For open conditionals, context and/or formal markers may indicate different degrees of doubt on the part of the speaker about the potential truthfulness of the conditional state of affairs.

152. a. He surely laughed if he saw me. high likelihood
b. He would laugh if he saw me. low likelihood
c. He would laugh if he were to see me. very low likelihood

A scale of conditionals can thus be constructed with pending truth-values at one end (open conditionals) with truth-values of falling degrees of likelihood down to negative truth-values at the other end (counterfactual conditionals). It can be observed that co-occurrence relations fit in nicely at the top of this scale with an affirmative truth-value. Against this background, it is easy to understand that some languages have the same construction for some types of co-occurrence relations and some types of conditional relations, as mentioned in section 4.1.2.
The distinction between open and counterfactual conditional constructions is attested in 15 of the 30 core sample languages, and in 19 of the 43 extended sample languages. The number may be suspected to be somewhat higher, since some languages distinguish between open and counterfactual conditionals only by using certain combinations of TMA markers in the two clauses, which is the case, for instance, in Erromangan. Grammars do not always give an account of such details in their description of conditional clauses.

Languages have various means of differentiating between degrees of potential truthfulness of open conditionals. An Austronesian example of a very stratified and systematic differentiation of open conditional constructions can be taken from Mekeo.

153. Mekeo

**VERY PROBABLE:**

a. *isa a-ke-mai *aisa-ma, lau a-la-lao  
   3 FUT-3.PL-come time-INT 1.SG FUT-1.SG-go

'If/when they come, I will go.' (East Mekeo dialect)  
(Jones 1998, p 513)

**MILDLY IMPROBABLE:**

b. *isa a-ŋe-lao koà, lau isava a-la-lao  
   3 FUT-3.SG-go if 1.SG also FUT-1.SG-go

'If she should go, I would go too.' (East Mekeo dialect)  
(Jones 1998, p 515)

**QUITE IMPROBABLE:**

c. *isa fe-lao koà, iʔa a-loa  
   3 OBLG.3.SG-go if 1.PL.IN 1.PL-go

'Should she go, we all would go.' (East Mekeo dialect)  
(Jones 1998, p 515)

**REMOTELY PROBABLE:**

d. *isa afe-lao koà, lau isava afa-lao  
   3 HYP.3.SG-go if 1.SG also HYP.1.SG-go

'If she went, I might also go.' (East Mekeo dialect)  
(Jones 1998, p 515)
4. SEMANTIC RELATIONS

VERY IMPROBABLE:

e. *ika ama-ao goa, iza fe-gi-oa*
   1.PL.IN IMP.1.PL-go if 3.PL OBLG-3.PL-go

'If we all went, they could go.' (North Mekeo dialect)
(Jones 1998, p 516)

HIGHERY IMPROBABLE:

f. *oi a-ŋo-lao koà aisa-ma lau isava a-la-lao*
   2.SG FUT-2.SG-go if time-INT 1.SG also FUT-2.SG-go

'If you were to go, I would go too.' (East Mekeo dialect)
(Jones 1998, p 516)

When *aisama* is used on its own for a relation marker, the truth-value is very high. When *koà* is used for a relation marker, the truth-value reading is dependent on the presence/absence of *aisama*: in its presence, the truth-value is very low, and in its absence, the truth-value varies with the TMA marking of the conditional clause. (See Jones 1998, pp 512-16.)

4.3.6.2 Concessive conditionals

In a concessive conditional, a component of unexpectedness is added to the conditional meaning (normally expressed by *even if* in English, as in *They will fire him even if he does a good job*). Unexpectedness is the key ingredient of concessive relations (see section 4.3.7 below), and for concessive conditionals, it is blended with the ingredient of hypothesis. Regular non-concessive conditionals, on the other hand, do not express the value of unexpectedness at all.

Concessive conditionals may be expressed with either an open or a negative truth-value, in the same way as plain conditionals. An open concessive conditional clause thus denotes a past, present or future hypothetical state of affairs, the fulfillment of which is not normally expected to co-occur with the state of affairs of the other clause, which in turn is expected to occur (or have occurred) regardless of the fulfillment of the concessive condition. An example is provided from Coastal Konjo.
154. Coastal Konjo

mang-na   ma-mo   ni-itte   ri   pulisi-a  
even.if-3.POSS  just-PFTV  PASS-see  by  police-DET

la-na-lukka'-ji  capi-na  para-na  tau  
FUT-3.ERG-steal-3.LIM  cow-3.POSS  fellow-3.POSS  person

'Even if he is watched by the police, he will steal his neighbor's cow.'  
(Timothy Friberg, questionnaire)

A counterfactual concessive conditional clause denotes a past hypothetical 
state of affairs that is known not to have occurred, but the fulfillment of 
which would not normally have been expected to co-occur with the state of 
affairs of the other clause, which in turn is actually known to have occurred.

155. Coastal Konjo

mang-na   ma-mo   ni-itte   ri   pulisi-a  
even.if-3.POSS  just-PFTV  PASS-see  by  police-DET

na-lukka'-i   kedde'   sihangngi  
3.ERG-steal-3.ABS  suppose  last.night

capi-na  para-na  tau  
cow-3.POSS  fellow-3.POSS  person

'Even if he had been seen by the police, he would have stolen his 
neighbor's cow last night.'  (Timothy Friberg, questionnaire)

The distinction between open concessive conditional (We will have a picnic 
even if it rains tomorrow) and counterfactual concessive conditional (We 
would have had a picnic even if it had rained yesterday) is something many 
grammars do not describe (although obviously, at least some languages have 
principled ways of expressing the distinction). Constructions commonly or 
exclusively associated with a concessive conditional relation are attested in 
11 of the 30 core sample languages and in 13 of the 43 extended sample 
languages. As with open and counterfactual conditionals, however, the ac-
tual figure is probably higher.
4.3.6.3 **Negative conditional relations**

Some of the sample languages have a specific relation marker for negative condition equivalent to *unless* in English. The only relational categories that are attested to have special negative relation markers in the sample are negative condition and negative purpose (see section 4.3.8.2). An example of negative condition is provided from Muna.

156. Muna

\[
\begin{align*}
\text{pa} & \quad \text{nae-} \text{mbali} & \quad \text{deki} & \quad \text{na-moni} & \quad \text{telo} & \quad \text{lambu} \\
\text{NEG.FUT} & \quad 3.\text{SG.IRR-} \text{can} & \quad \text{first} & \quad 3.\text{SG.IRR-} \text{go} & \quad \text{in} & \quad \text{house}
\end{align*}
\]

\[
\begin{align*}
\text{tabea} & \quad \text{na-tumisa} & \quad \text{deki} & \quad \text{kapa-no} \\
\text{unless} & \quad 3.\text{SG.IRR-} \text{plant} & \quad \text{first} & \quad \text{cotton}-3.\text{SG.POSS}
\end{align*}
\]

'He cannot go up into his house, unless he first plants his cotton.'
(Berg 1989, p 251)

A negative conditional relation marker is attested in 6 of the 30 core sample languages and in 8 of the 43 extended sample languages.

4.3.6.4 **Other variants of conditional relations in Austronesian languages**

There are some other conditional relation markers among the sample languages with distinctive meanings that are apparently much less common, however. A few of these are presented in this section.

Some languages have designated relation markers indicating an exclusive condition, i.e. only the fulfillment of the conditional state of affairs, and no other, would entail the realization of the hypothetical concomitant state of affairs (cf. English *only if*). Note the Mbula example below, in which the exclusive conditional marker *bela* is a verb meaning 'go' prefixed with a non-factuality marker.

---

50 The Muna relation marker *tabea* has a range of uses and can also be taken to mean 'only' (Berg 1989). It seems to derive from an exclusive conditional (see next section) that could be rendered 'Q will happen; only if P happens, Q will not happen' which has led to 'Q will happen, unless P happens'. The same overlap between a negative conditional marker and an adverb meaning 'only' is found in Karo Batak (Woollams 1996).
The exclusive conditional relation is the inverse of an ordinary conditional relation, since an exclusive condition is a necessary condition for the other state of affairs to be true, while an ordinary condition is a sufficient condition for the other state of affairs to be true. In the Mbula exclusive conditional, the letter may have a 60 toea stamp on it for other reasons as well (for instance, to reach some other part of the world), but if it does not have a stamp to this value, we know that it will not get to America (by regular postal means, anyway). In logic, a condition the fulfillment of which is both necessary and sufficient for the other state of affairs to be true is called equivalence, but no such conditional constructions have been found in the Austronesian languages. However, in Austronesian, as in any language, such interpretations could arise from both ordinary and exclusive conditional constructions given the right context. In general, purely logical renditions of conditionals are often overridden in natural language by pragmatic context-dependent considerations.

Another conditional relation marker found among the sample languages indicates a strong desire on the part of the speaker for the conditional and the hypothetical concomitant state of affairs to be true (cf. 'if only' in English). An example is provided by Palauan.

In Tagalog, it is interesting to note that the combination of the imperative negator and the general sentence negator yields a diminutive concessive conditional meaning (~however little, in English).
'However little I eat, I still get fat.' (Schachter & Otanes 1972, p 480)

And in Buru, the relation markers mama and bama mentioned in section 4.3.1 can be used both for co-occurrence and for condition; mama implies continuity with a prior discourse context (~'and if'), while bama implies discontinuity (~'but if'). Compare the clauses below.

160. Buru

a. fene, mama gam dii, do
   say and.if be.like that well
   ku ego uka ture-n ang dii
   2.SG get bamboo short-GEN IMM that

   '[He] said, "And if that is the case, well, you go fetch that piece of bamboo."' (Grimes 1991, p 413)

b. bama sira ep-mata-h moo,
   but.if 3.SG CAUS-die-3.SG NEG

   sira benihi-k mohede
   3.SG be.happy-APPL not.yet

   'But if they have not yet killed it, they are not yet satisfied.'
   (Grimes 1991, p 413)

4.3.7 Concession

Concessive relations hold in clause combining constructions in which one of the clauses is explicitly marked as denoting a concession, i.e. a state of affairs that is asserted to be true and that has an expected implication running counter to the state of affairs denoted by the other clause, which is, however, also asserted to be true. No implicit constructions commonly associated with concessive relations have been found.
In the sample languages, the clause representing the concessive state of affairs is most often marked as such, although some constructions (optionally or obligatorily) have marking on both clauses (cf. 'although…still' in English) or only on the clause representing the counter-expected state of affairs. Because all constructions with these different kinds of marking emphasize the same kind of relation, they were all classified in the study as instances of concessive relations. Some Austronesian examples are presented below.

161. Ulithian (Micronesian)

\[ \text{ilamou ho te dipli skul ngo ho-be dabey} \]
even though 2.SG not like school and 2.SG-will attend

'Even though you don't like school, you must attend' (John A. Walsh, questionnaire)

162. Yabem

\[ \text{kamoc ké-taŋ enŋ, mago ké-taŋ atom} \]
sore 3.SG.REAL-burn 3.SG still 3.SG.REAL-cry NEG

'Although the wound burned him, he didn't cry.'
(Dempwolff 2005 [1939], p 107)

Constructions explicitly marking concessive relations are attested in 19 of 30 languages in the core sample and in 26 of the 43 languages of the extended sample.

Relation markers with slightly nuanced concessive meanings were attested in some Austronesian languages although the variation is not considerable. In Samoan, there is a special concessive relation marker expressing a situation perceived of as negative for the actor of the concessive clause. This relation marker derives from the verb for 'be painful'.

163. Samoan

\[ \text{tīgā ona 'ou ole atu i lo'u tinā} \]
be.painful COMP 1.SG beg DIR to 1.SG.POSS mother

\[ \text{'ae tasi lava lana 'upu [...]} \]
but one EMPH 3.SG.POSS word

'However much I begged my mother, she only said…'
(Mosel & Hovdhaugen 1992, p 663)
The regular concessive relation marker in Samoan is *ui*, derived from the verb meaning 'grant'.

**164. Samoan**

\[
\begin{align*}
    ma & \quad lo'\quad māfaufau \\
\text{and} & \quad 1.SG.POSS \quad \text{thinking}
\end{align*}
\]

\[
\begin{align*}
    'ua & \quad māfa\quad tū \quad i \quad le \quad su'\quad ga \\
\text{INCH be.\quad heavy \quad stand \quad to \quad DEF \quad search-NZR}
\end{align*}
\]

\[
\begin{align*}
    e & \quad ui \quad lava \quad ina \quad le'i \quad taitai \quad le \quad taimi \\
\text{GNR be.\quad granted \quad indeed \quad COMP \quad not.yet \quad be.near \quad DEF \quad time}
\end{align*}
\]

'And my thinking was entirely on the exam, although it was not yet time.' (Mosel & Hovdhaugen 1992, p 662)

In Tagalog, the use of either of the relation markers *gayong* or *ganoong* implies disapproval of the counter-expected state of affairs on the part of the speaker.

**165. Tagalog**

\[
\begin{align*}
    hindi & \quad raw \quad nila \quad ga-gaw\quad in \quad ang \quad trabaho \\
\text{be.\quad not \quad it.is.said \quad 3.PL.GEN \quad IRR\quad make-PF \quad TOP \quad work}
\end{align*}
\]

\[
\begin{align*}
    gayong & \quad b[in]\quad ayar\quad an \quad mo \quad sila \\
\text{even.\quad though \quad [PFTV]pay-PF \quad 2.SG.GEN \quad 3.PL.TOP}
\end{align*}
\]

'They say they won't do the work even though you paid them' (Schachter & Otanes 1972, p 479)

Sometimes the term 'adversative relation' is used interchangeably with the term 'concessive relation'. Conversely, scholars may attribute the difference between them to syntax alone, claiming that concession is subordination (often coded by *although* in English) and adversativity is coordination (coded by *but* in English). I take the difference to be semantic in nature. While concessive relations indicate that given one state of affairs, the occurrence of another state of affairs runs counter to general expectations, adversative relations do not provide this component of counter-expectancy (although it is often inferred pragmatically from the discourse context).\(^{51}\) Compare the examples below.

---

\(^{51}\) See Crevels (2000a) and Malchukov (2004) for elaborations of this view.
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

166. Adversative (general contrast):
*Bananas are yellow but pears are green.*

167. Concessive (counter expected contrast):
*Although he lost the match, he was celebrated as the winner.*

168. Adversative (general contrast, counter-expectancy inferred):
*He lost the match, but he was celebrated as the winner.*

169. Concessive (counter expected contrast):
*Although bananas are yellow, pears are green.*

Since adversative and concessive relations are semantically similar (coding contrast), they are also sometimes morpho-syntactically marked in similar ways. In English, for instance, the coordinator *but* alone indicates adversativity (general contrast), while the coordinator-adverb sequence *but still* can be said to indicate concession, since *still* adds a component of counter-expectancy. Adversative relations have not been considered in this study, as they do not involve counter-expectancy, merely general contrast, and are therefore not really co-variational. A change in one state of affairs does not imply a change in the other. Concessive relations, on the other hand, have a clear component of implication, which make them unidirectional and, therefore, co-variational.

4.3.8 Purpose

4.3.8.1 Regular purpose

Purposive relations hold in clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs aimed at being attained by means of the state of affairs outlined in the other clause. Purposive relations are also assumed to hold in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause, or if structurally identical, the second clause, is conventionally interpreted as denoting a purposive state of affairs. In the example *He stood up in order to see better,* the first clause encodes the means by which the actor intends to realize the purpose, encoded by the second clause.

Often the actor in the means clause is the same as a core argument participant in the purpose clause, since the means is usually aimed at enabling the actor to do or experience whatever is encoded in the purpose clause. Therefore, many Austronesian languages have specific purpose constructions in which the actor in the means clause is obligatorily the same as a salient participant in the purpose clause, and in which reference to this participant is optionally or obligatorily suppressed in the purpose clause. This is
comparable to non-finite purpose constructions in languages with a clear finiteness distinction. Austronesian examples of purpose relations are provided from Paluan and Hoava below.

170. Palauan

\[ a \ droteo \ a \ ngiluu \ a \ mlai \]
\[ PHR \ NAME \ PHR \ take.PST \ PHR \ car \]
\[ \text{çl mo mëchar a biang} \]
\[ \text{in.order.to go.PRES buy.PRES PHR beer} \]

'Droteo took the car to go buy beer.' (Josephs 1975, p 301)

171. Hoava

\[ la \ ria \ va-bubui-a \ se \ amina \]
\[ go \ 3.PL.SBJ CAUS-forget-3.SG.OBJ ART NAME \]
\[ de \ hiko-a \ lisa \ sa \ baeka \ tabu \]
\[ PURP \ steal-3.SG.OBJ NAME ART bag tabu.shell \]

'They go and make Amina forget in order for Lisa to steal the tabu shell.' (Davis 2003, p 280)

Constructions commonly or exclusively used to express purpose relations have been attested in all but one of the sample languages. In Acehnese, no combined clause purposive construction is attested, although my source (Dubre 1985) mentions a serial verb construction in which the second verb sometimes has a purposive sense. Serial verb constructions, however, do not fall under the definition of clause combining constructions used in this study.

4.3.8.2 Negative purpose

In several of the sample languages there is at least one construction denoting a semantic relation that is often called 'caveat' (e.g. Hill 1992 for Longgu) or 'aversive' (e.g. Crowley (1998) for Erromangan). This denotes an unfavorable state of affairs that is preferably avoided and is often translated into English by means of 'lest', 'otherwise' or 'in case'. One of the shades of meaning that is subsumed under this relation is negative purpose, i.e. the purpose is to avoid the realization of the state of affairs denoted in the marked clause. The examples below illustrate.
172. Mekeo (east dialect)

\[
[ \text{laueu} \quad \text{kapia} \quad \text{fe-}\text{ŋa} \quad \text{foʔa,} \quad \text{la-}\text{pa-pua-Ø} ]
\]
1.SG.POSS plate 3.SG.OBLG-lose lest 1.SG-CAUS-hide-3.SG

'In case my plate goes missing, I will hide it.' (Jones 1998, p 531)

or: 'In order for my plate not to go missing, I will hide it.'

173. Erromangan

\[
\text{ko-etwo-nomonki} \quad \text{orog}
\]
2.SG.HOD-NEG.FUT-drink much

\[
\text{k-ante} \quad \text{m-agku} \quad \text{m-amah}
\]
2.SG.FUT-stay ES.SG-say ES.SG-die

'Don't drink too much, in case you die.' (Crowley 1998, p 258)

or: 'Don't drink too much, so that you won't die.'

These constructions usually also denote unfavorable states of affairs in a more general way in the languages in which they occur. Compare the two Tinrin examples below. In the (a) example, the state of affairs in the first clause does not clearly constitute the means by which the state of affairs in the second clause is intended to be avoided. In the (b) example, on the other hand, it does.

174. Tinrin

a. \[
\text{u} \quad \text{barri} \quad \text{traiki} \quad \text{pa} \quad \text{nrîî} \quad \text{eghe} \quad \text{rò}
\]
1.SG be.afraid dog lest 3.SG.FUT bite 1.SG

'I'm afraid of the dog, as it might bite me.' (Osumi 1995, p 276)

b. \[
\text{savaa} \quad \text{pa} \quad \text{rri} \quad \text{nrî} \quad \text{ho} \quad \text{nrî} \quad \text{nrâ} \quad \text{mêrrê} \quad \text{mwâ}
\]
defend lest 3.PL IRR eat 3.SG.OBJ SBJ PL DEM

'Pay attention lest others eat it.' (Osumi 1995, p 276)

In many languages, the negative purpose clause can be used on its own as a warning of an unfavorable state of affairs, which may be the consequence of a certain activity.
175. Erromangan

\[ k-ante \ m-agku \ m-amol \]
2.SG.FUT-stay  ES.SG-say  ES.SG-fall

'You might fall!' (Crowley 1998, p 258)

Of the relevant constructions, only those for which one of the readings is compatible with a negative purpose meaning have been considered in the present study. Constructions commonly or exclusively used to express a negative purpose relation are attested in 9 of the 30 core sample languages and in 15 of the 43 extended sample languages.

**4.3.8.3 Fulfilled purpose**

Generally, in the languages of the world, the state of affairs denoting the means in purposive relations is presented as asserted information, while the purpose is usually presented as non-asserted information. That is, no information is provided as to whether the purposive state of affairs was actually realized or not. In some Austronesian languages, however, there are also constructions in which the purpose is presented as denoting a realized purpose. Such constructions are not common, but they are attested in three of the sample languages (Buru, Longgu, Manam). Compare the Longgu examples below.

176. Longgu

a. \[ ma \ mwane-na \ ii'o \ mola \]
and man-that stay only

'\[ania \ wa'iwa'i-a \ burung-a-na-i-na \]
PURP.REAL  hit-3.SG  spouse-3.SG-SG-that

'And that man just stayed to beat his wife [and he did beat her].' (Hill 1992, p 272)

b. \[ ara \ la \ hou \ ni \ voli-u \ malau \ nina \]
3.PL  go thither PURP.IRR  buy-1.SG  PLACE  then

'They went to buy me [for marriage] at Malau then.' (Hill 1992, p 267) [There is no information as to whether they actually did "buy me" or not.]
In the (a) example, the use of 'ania indicates that a beating actually took place, while in the (b) example, the use of ni gives no information about whether any deal was made or not. I have found no constructions in which the purpose relation is specifically presented as denoting a frustrated purpose.

4.3.8.4 Anticipatory purpose

Another interesting purposive relation, attested only in Muna and Central Cagayan Agta, is one that I have called anticipatory purpose. Clauses marked for anticipatory purpose denote a purpose that is desired but far from certain, i.e. the realization of the intended result may depend on other circumstances than the potential of the actor, while it is nevertheless anticipated and hoped for by the actor. The Muna sentence below provides an example.

177. Muna

\[
a-mangkafi-da-mo \quad bhahi \quad kaawu \quad a-rafo-da
\]

2.SG-IRR.follow-3.PL-PFTV \hspace{0.5cm} PURP \hspace{0.5cm} only \hspace{0.5cm} 1.SG-REAL.catch-3.PL

'I will follow them so that perhaps I will find them.'

(Berg 1989, p 261)

Muna also has negative anticipatory purpose constructions. In these, the marked clause denotes a state of affairs that is undesirable, and that although precautions are taken (the state of affairs encoded by the other clause), it is highly uncertain what the outcome may be.

178. Muna

\[
feka-taa \quad ne \quad ini \quad bhahi \quad ta-do-wora-e \quad mie
\]

CAUS-be.good at this \hspace{0.5cm} PURP \hspace{0.5cm} TB-3.PL-REAL-see-3.SG \hspace{0.5cm} people

'Put it there, so that perhaps people will not see it.' (Berg 1989, p 261)

The verb of the negative anticipatory purpose clause in Muna is prefixed by \(ta\)-, which, according to Berg (1989), is homonymous and probably shares the etymology with the Muna terminal boundary prefix \(ta\)- (cf. example (141) above), although together with \(bhahi\), it is used to indicate negative anticipatory purpose.
4.3.9 Reason

Reason relations hold in clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs that is directly causing or in some way constituting the grounds for another state of affairs to emerge. Reason relations are also taken to hold in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause, or if structurally identical, the second clause, is conventionally interpreted as denoting a causing state of affairs in relation to the state of affairs of the other clause. Reason relations, as intended in the present study, include both relations involving external motivation (direct causes) and those involving internal motivations (in the mind of an agent). Austronesian languages seem to use the same construction for both, as indeed most languages do (Givón 2001). Austronesian examples are provided below.

179. Erromangan

\[
\begin{align*}
nimo & \quad y\text{-omol} \\
& \quad \text{house} \quad 3.\text{SG.DISTPST-fall} \\
ra & \quad nemetagi \quad y\text{-elimsi} \\
& \quad \text{because} \quad \text{cyclone} \quad 3.\text{SG.DISTPST-blow.3.SG}
\end{align*}
\]

'The house fell over because the cyclone blew it.'
(Crowley 1998, p 273)

180. Yakan

\[
\begin{align*}
ga'i & \quad \{kite\ bi\} \quad maka-atu \\
& \quad \text{not} \quad 1.\text{PL.IN.ABS} \quad \text{ABIL-challenge} \\
pegge' & \quad dik'i\ dik'i' \quad du \quad \{kite\ bi\} \\
& \quad \text{because} \quad \text{be.very.small} \quad \text{certainly} \quad 1.\text{PL.IN.ABS}
\end{align*}
\]

'We cannot challenge them because we are very small.'
(Brainard & Behrens 2002, p 197)

Three specific reason constructions are worth mentioning briefly, singling out a specific shade of meaning found in only one sample language each.

In Central Cagayan Agta, a clause following either of the relation markers \textit{awá} or \textit{bakawá} indicates that the following clause denotes a state of affairs that provide reason for doubting the truth of the state of affairs denoted by the other clause. The preceding clause has to express this doubt (either
lexically or phonologically) by interrogative, dubitative or negating features (Mayfield 1972, p 49).

181. Central Cagayan Agta

a. maga-tahabáku hud
   AF.NPST-work DUB

   awá mag-yán lá ta bali i-maqidda
   because AF.NPST-stay only in house IF-lie.down

   'It isn't true that he works since he just stays in the house lying down.' (Mayfield 1972, p 50)

b. páqpaqnun na maka-tugut
   how 3.SG AF.NPST.ABIL-walk

   bakawá na-tahengkányun
   because AF.PST-be.supported.by.wood

   'How can he walk as he is propped up on wood [i.e. crutches]?' (Mayfield 1972, p 50)

In these cases, the reason relation clearly emerges from the speech act level rather than the content level. The reason clause provides the reason to maintain a certain attitude (doubt) toward the content of the other clause; it does not give the reason for the content state of affairs of the other clause. The word awá seems to be related to the general negator awe in Central Cagayan Agta (for examples, see Healey 1960, pp 38 and 50). Speculatively, awá may well originate from the negator awe in predicate position linked to the following clause by the linker á. Such head-dependent linkers (often called ligatures) are found in some Formosan and Philippine languages, where it sometimes takes the form a, as in Amis (Liu 2003a) and Ilokano (Foley 1976).

Another reason construction with negative connotations is found in Samoan, where the word for '(be) bad' has grammaticalized into a reason relation marker meaning 'because unfortunately'. It indicates that both the reason and the entailing result are unwanted states of affairs.

182. Samoan

   'ua 'ou fa’anoanoa leaga 'ua lēiloa la'u maile
   INCH 1.SG be.sad because INCH loose 1.SG.POSS dog

   'I am sad because I lost my dog.' (Mosel & Hovdhaugen 1992, p 627)
The position of *leaga* and the lack of a TMA particle preceding it indicate that it is no longer a verb in the above construction but a relation marker.

Finally, Indonesian has a specific relation marker to denote exclusive reason, i.e. focusing on the fact that a certain state of affairs, and no other, is the reason for a certain result.

183. Indonesian

\[\text{mentang-mentang dia orang kaya,} \]
\[\text{extend-extend 3.SG person rich} \]
\[\text{dia berbuat se-enak-nya} \]
\[\text{3.SG do DER-please-3.SG} \]

'Just because he is a rich person he does what he likes.'
(Sneddon 1996, p 343)

The word *mentang* is a verb meaning 'extend', 'spread out' (Sederet Online dictionary).

4.3.10 Result

Result relations hold in clause combining constructions in which one of the clauses is explicitly marked as denoting a state of affairs occurring as the direct or indirect consequence of the state of affairs of the accompanying clause. Result relations are also taken to hold in relationally implicit clause combining constructions for which the morpho-syntactically deviating clause, or if structurally identical, the second clause, is conventionally interpreted as denoting a state of affairs occurring as a consequence of the state of affairs of the other clause.

Result relations are the opposite of reason relations in that a result presupposes reason, and vice versa, and similar to posteriority and anteriority – but unlike most other opposing pairs of relations – either the result or the reason may be explicitly marked as such in clause combining in most Austronesian languages when two states of affairs relate to each other in this way. Result relations are very commonly encoded in the Austronesian languages but are not quite as widespread as constructions encoding reason relations. Constructions commonly or exclusively used to express a result relation are attested in 28 of the 30 core sample languages and in 36 of the 43 extended sample languages. Most languages have at least one construction explicitly denoting result, although some are attested conventionally to use implicit constructions in addition, such as asyndetic juxtaposition and
neutral coordination, for result relations. Examples of Austronesian constructions representing result are provided below.

184. Taba

\[
\text{ni} \quad \text{reng} \quad \text{ta-dopas} \\
\text{3.SG.POSS} \quad \text{seal} \quad \text{PASS-perish}
\]

\[
\text{ndadi} \quad \text{ni} \quad \text{ol} \quad \text{n-sopal-ik} \\
\text{so} \quad \text{3.SG.POSS} \quad \text{oil} \quad \text{3.SG-grow.out-APPL}
\]

'It's seal is perished so oil leaks out of it.' (Bowden 1997, p 458)

185. Kambera

\[
\text{ta-pakiring} \\
\text{1.PL.IN.NOM-start}
\]

\[
\text{ka} \quad \text{ta-tinu-nya} \quad \text{na} \quad \text{lau} \quad \text{haromu} \\
\text{so} \quad \text{1.PL.IN.NOM-weave-3.SG.DAT} \quad \text{ART} \quad \text{sarong} \quad \text{tomorrow}
\]

'We start (with something else) so that we'll weave the sarong tomorrow' (Klamer 1998, p 143)

In Leti and Samoan, implicit constructions are the only attested constructions used for result relations. In Leti, an initial clause, nominalized by an enclitic deictic particle, becomes the topic of the sentence, and depending purely on contextual cues, it may be construed as a result clause or a conditional clause (Engelenhoven 1995). An example of the former is provided below.

186. Leti

\[
\text{r-salinn-e-de}, \quad \text{na-tjerkyèran-o} \\
\text{3.PL-queue-DX-DEIC} \quad \text{3.SG-be.thirsty-IND}
\]

'That they queued up was because he was thirsty.' (Engelenhoven 1995, p 241) (or 'He was thirsty, so they queued up.' Note that the result clause is structurally marked.)

In Samoan, a neutral coordinator is used as relation marker for result relations, and here too, the construction is dependent on contextual information for its interpretation as a result construction, although an anaphoric particle is often present in the result clause, which guides the reader/hearer toward a result interpretation.
4. SEMANTIC RELATIONS

187. Samoan

\[ sa \quad pa'ū \quad le \quad popo \quad i \quad le \quad ulu \]
\[ PST \quad fall \quad SPEC \quad coconut \quad at \quad SPEC \quad head \]

\[ o \quad le \quad tamāloa, \quad ma \quad 'ua \quad ia \quad matua'i \quad manu'a \quad ai \]
\[ of \quad SPEC \quad man \quad and \quad INCH \quad 3.SG \quad be.severe \quad be.hurt \quad ANA \]

'The coconut fell in the man's head so that he was badly hurt.' (AJ)
Lit. 'The coconut fell in the man's head and he was badly hurt (by it).'

There is no considerable variation in meaning among the Austronesian result constructions. One distinction, however, formally present by means of a designated relation marker in a few of the sample languages – Indonesian, Kambera and Tboli – is made between a general result and a graded result. In English, a graded result can be expressed by the relation marker to the extent that. A graded result clause denotes the result of the extent to which a certain quality of a state of affairs holds true, rather than the result of the state of affairs itself. Note the Indonesian example below.

188. Indonesian

\[ dia \quad lari \quad begitu \quad cepat \]
\[ 3.SG \quad run \quad like.that \quad fast \]

\[ sehingga \quad sangat \quad sukar \quad untuk \quad mem-otret-nya \]
\[ to.the.extent \quad very \quad be.difficult \quad PURP \quad ACTV-take.photograph-3.SG \]

'She was running so fast that it was very difficult to photograph her.'
(Sneddon 1996, p 351)

In this case, it is not merely the state of affairs of running quickly that makes it difficult to photograph her, but the actual speed at which she is running.
In this chapter, polysemic patterns between Austronesian clause combining constructions will be explored. Under the assumption that polysemic patterns in synchronic data reflect paths of semantic development diachronically, the extent to which relation markers are polysemic between two or more semantic relations may suggest trends of development. Consistent and regular patterns for such developments have been suggested in the literature and these will be examined from the viewpoint of the Austronesian data.

I use the term polysemy broadly to refer to situations in which linguistic items – in our case, relation markers – display multiple meanings, regardless of whether they arise in interaction with the context (different uses) or constitute different conventional meanings (different senses). Sometimes the difference between the two is difficult to determine, which is why I prefer not to commit my use of 'polysemy' to either one.52

5.1 Three paths of development well-attested cross-linguistically

Interesting historical ties between relational markers for spatial, temporal and co-variational relations have been found to follow similar patterns of development in many unrelated languages of the world (see e.g. Genetti 1986, Ohori 1996, Kortmann 1997). Specifically, markers of spatial co-occurrence (at) tend to develop into markers of temporal co-occurrence (when, while), which tend to develop into conditional markers (if). Markers of spatial goal (to) tend to develop into markers of (temporal) terminal boundary (until), which tend to develop into markers of purpose ((in order) to). Also, markers of spatial source (from) tend to develop into markers of (temporal) initial boundary (since), which tend to develop into markers of

52 Haspelmath (2003) prefers to use the term 'multifunctionality' as a general term to cover multiple meanings arising from both different 'uses' and different 'senses'. I would prefer, however, to use 'multifunctionality' for situations in which a linguistic item may have different grammatical/syntactic functions, e.g. a relation marker functioning either as a subordinator or an adverbial, such as Kambera hàla, 'after'/afterwards'. In this case, the multifunctionality also entails relational polysemy (see section 5.2.5 below).
reason (*because*). This is summarized in Table 12 below (derived from similar tables in Genetti 1986 and Ohori 1996).

**Table 12.**

<table>
<thead>
<tr>
<th>DEVELOPMENT</th>
<th>Sameness</th>
<th>Goal</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>SPATIAL SAMENESS</td>
<td>SPATIAL GOAL</td>
<td>SPATIAL SOURCE</td>
</tr>
<tr>
<td></td>
<td><em>at</em></td>
<td><em>to</em></td>
<td><em>from</em></td>
</tr>
<tr>
<td>Time</td>
<td>CO-OCCURRENCE</td>
<td>TERM. BOUNDARY</td>
<td>INIT. BOUNDARY</td>
</tr>
<tr>
<td></td>
<td><em>when, while</em></td>
<td><em>until</em></td>
<td><em>since</em></td>
</tr>
<tr>
<td>Co-variation</td>
<td>CONDITION</td>
<td>PURPOSE</td>
<td>REASON</td>
</tr>
<tr>
<td></td>
<td><em>if</em></td>
<td><em>(in order) to</em></td>
<td><em>because</em></td>
</tr>
</tbody>
</table>

The direction of development from spatial via temporal to co-variation markers follows the general pattern of semantic development from a concrete to a more abstract meaning. Space is stable and concrete, time is always ongoing and less concrete than space (at least as experienced by the human mind), and co-variation is the way states of affairs affect each other. Relationally, each step is more abstract than the previous, and it is easy to appreciate that semantic change tends to occur in the direction from concrete to abstract notions. It also seems quite natural for the meaning of relational markers to spread along the paths of development represented by the three columns in Table 12 more often than cross-path-wise. Spatial sameness, temporal co-occurrence and condition all indicate sameness in some sense; spatial source, initial boundary and reason all share the component of source; and spatial goal, terminal boundary and purpose share the component of goal. They seem to form natural classes, each sharing some component of meaning. And as we can see, the conceptual reference map of semantic relations that was presented in the previous chapter (see Table 11) is perfectly compatible with Table 12, based on diachronic data.

The Austronesian data on clause combining constructions strongly support the close connection between co-occurrence and conditional relations at the bottom of the first column of Table 12. In 21 of the 30 core sample languages and 27 of the 43 extended sample languages (70 % and 63 %, respectively), at least one relation marker was found to be used both in constructions explicitly expressing co-occurrence and constructions explicitly expressing conditionality.\(^{53}\) The following examples illustrate polysemic rela-

---

\(^{53}\) Constructions for which the relation remains implicit, such as juxtaposed clauses and clauses coordinated by a neutral coordinator (equivalent to English *and*), were excluded in this analysis.
tion markers between spatial sameness and temporal co-occurrence (189), and between temporal co-occurrence and condition (190).

189. Central Cagayan Agta

\[
\text{sangaw } ta \ naka-dagut \ kid \ ta \ weh-en,
\]
then at/when AF.PST-descend 3.PL.TOP at creek-DEF

\[
\text{weh } na \ \text{danglig}, \ yen \ ya \ \text{nagafuyan} \ da
\]
creek POSS NAME that TOP cooking.place 3.PL.POSS

'Then, when they arrived down at the Danglig creek, that's their cooking place.' (Mayfield 1972, p 26)

190. Erromangan

a. \text{nempgon} \ kem-agkol-i \ nacave
\[
\text{time} \quad 2.SG.PRES-dig? \ kava
\]
\[
\text{ko-agkol-i-veh} \quad \text{ko-etwo-anduntvi} \ ov-nowatni-n
\]
2.SG.FUT-see.?-good 2.SG-NEG.FUT-break PL-root-3.SG.POSS

'When you dig up the kava, be careful not to break its roots.'
(Crowley 1998, p 271)

b. \text{nempgon} \ kokum-ampculac
\[
\text{time} \quad 1.DU.IN.PRES-get.married
\]
\[
\text{cw-ampai} \ nulgo-n \ ntal \ enyau
\]
3.PL.FUT-take leaf-3.SG.POSS taro 1.SG.POSS

'If we get married, they will take my taro leaves.'
(Crowley 1998, p 271)

These examples of polysemy do not really say anything of the direction of development in Austronesian languages for the relevant relation markers. However, looking at some of the cases in which my sources indicate more than one meaning for co-occurrence and conditional relation markers, it seems reasonable to assume that the relation markers have extended their semantics in the direction toward more abstract notions, in accordance with Table 12. Note, for instance, the relation markers in Table 13 below.
Another reasonably salient polysemic pattern in Austronesian languages – which is, however, not reflected in Table 12 but is apparent in my data – is that between verbs for 'say' or 'speak' and conditional relation markers. This connection has been noted to occur in several other language families as well (e.g. Saxena 1988; Heine et al 1991). Note the Austronesian examples in Table 14 below.

Table 13.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>LITERAL MEANING</th>
<th>TC MEANING</th>
<th>SUGGESTED DEVELOPMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaiian</td>
<td>ʻā</td>
<td>at</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Erromangan</td>
<td>nempgon</td>
<td>time</td>
<td>when, if</td>
<td>time &gt; co-oc., cond.</td>
</tr>
<tr>
<td>Hoava</td>
<td>pa</td>
<td>at</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Mekeo</td>
<td>miangai</td>
<td>at beginning</td>
<td>while</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Leti</td>
<td>lera</td>
<td>day</td>
<td>if</td>
<td>time &gt; condition</td>
</tr>
<tr>
<td>Coastal Konjo</td>
<td>ri</td>
<td>at</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Coastal Konjo</td>
<td>naia ri</td>
<td>as.for at</td>
<td>if</td>
<td>space &gt; condition</td>
</tr>
<tr>
<td>Muna</td>
<td>welo</td>
<td>in</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Muna</td>
<td>bhe</td>
<td>with</td>
<td>while</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Palauan</td>
<td>(er) se ʻer</td>
<td>at that</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
<tr>
<td>Tagalog</td>
<td>sakali</td>
<td>moment</td>
<td>in case</td>
<td>time &gt; condition</td>
</tr>
<tr>
<td>Tinrin</td>
<td>rugi ōō</td>
<td>at that</td>
<td>when</td>
<td>space &gt; co-oc.</td>
</tr>
</tbody>
</table>

The word rugi, 'at', is a complex preposition made up of constituent prepositions ru, 'on', 'in', and gi, 'at' (Osumi 1995, pp 84-5). The word ōō is a complementizer introducing complement clauses (Osumi 1995, p 264).

Table 14.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>LITERAL MEANING</th>
<th>TC MEANING</th>
<th>RELATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Araki</td>
<td>co de</td>
<td>it would say</td>
<td>if</td>
<td>condition</td>
</tr>
<tr>
<td>Big Nambas</td>
<td>-vr ti</td>
<td>say that</td>
<td>if</td>
<td>condition</td>
</tr>
<tr>
<td>Mbula</td>
<td>so</td>
<td>say</td>
<td>if</td>
<td>condition</td>
</tr>
<tr>
<td>Indonesian</td>
<td>bicara</td>
<td>speak</td>
<td>if</td>
<td>condition</td>
</tr>
</tbody>
</table>
The literal meaning of relation markers is not altogether consistently noted in my data, and a qualified guess would be that several more examples of both these patterns (co-occurrence-conditional polysemy and say/speak-conditional polysemy) could be found in Austronesian languages. The areal and genealogical distribution, at least in Table 13, is fairly wide and seems to underpin such an assumption, in any case for the co-occurrence-conditional pattern of polysemy.

There is less of a basis for firm conclusions in my Austronesian data for the path of development in the second column of Table 12 – i.e. the cross-linguistic tendency for relation markers to extend their meaning from spatial goal to terminal boundary to purpose. The share of languages for which explicit constructions for both terminal boundary and purpose relations were attested amounted to slightly less than 2/3 in both samples (19/30 in the core sample and 26/43 in the extended sample). By comparison, the vast majority of the sample languages had explicit constructions for both co-occurrence and conditional relations (28/30 in the core sample and 38/43 in the extended sample). However, the connection between terminal boundary and purpose can be seen in a few of the languages. In 3 of the 30 core sample languages and in the same 3 of the 43 extended sample languages (10 % and 7 %, respectively), at least one principal relation marker was found to be used both in constructions explicitly expressing terminal boundary and constructions explicitly expressing purpose. There are also a few examples of polysemy between spatial goal and either or both of the relations of terminal boundary and purpose. The following examples from Loniu and Muna illustrate such a pattern.

191. Loniu

a. ɔ-pɔ an [ ki-lɛ sih pelet... ]
   2.SG-do water 3.SG.POT-go one plate

   'Put some water into a dish...' (Hamel 1994, p 128)

b. ɲanɛ suʔu i-tɔ tuwani [ i-lɛ meʔisan ]
   mother 3.DU 3.SG-be.at cook 3.SG-go be.done

   'Their mother would cook it until it was done.'
   (Hamel 1994, p 130)

c. iy i-puti iy [ i-lɛ čani pute-n ]
   3.SG 3.SG-take 3.SG 3.SG-go cut umbilical.cord-3SG.POSS

   'She took him in order to cut the umbilical cord.'
   (Hamel 1994, p 130)
5. PATTERNS OF POLYSEMY

192. Muna

a. *a-gholi-e* *so ihintu*
   1.SG.REAL-buy-3.SG.OBJ  for 2.SG

'I bought it for you' (Berg 1989, p 139)

b. *no-sangke pandanga-no*
   3.SG.REAL-lift  spear-3.SG.POSS

*so na-lumogha-e-ghoo*
   for 3.SG.IRR-stab-3.SG.OBJ-PURP

'He lifted his spear to stab him.' (Berg 1989, p 260)

The Loniu examples also hint at another polysemic pattern involving terminal boundary and purpose relations, viz. of that between verbs of directed motion and relation markers for these relations. Where literal meanings are described in the sample languages, motion verbs seem to dominate for terminal boundary and/or purpose relation markers. Some examples are provided in the table below.

Table 15.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>LITERAL MEANING</th>
<th>TC MEANING</th>
<th>RELATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Nambas</td>
<td>dav’a</td>
<td>go</td>
<td>until</td>
<td>term. boundary</td>
</tr>
<tr>
<td>Lenakel</td>
<td>aroatis</td>
<td>arrive</td>
<td>until</td>
<td>term. boundary</td>
</tr>
<tr>
<td>Loniu</td>
<td>mä</td>
<td>go here</td>
<td>in order to</td>
<td>purpose</td>
</tr>
<tr>
<td>Loniu</td>
<td>la</td>
<td>go away</td>
<td>in order to</td>
<td>purpose</td>
</tr>
<tr>
<td>Tetun</td>
<td>dudaun</td>
<td>continue</td>
<td>until</td>
<td>term. boundary</td>
</tr>
<tr>
<td>Tetun</td>
<td>hodi</td>
<td>bring</td>
<td>in order to</td>
<td>purpose</td>
</tr>
</tbody>
</table>

While the data do not indicate anything concerning actual directions of development between these relations in Austronesian, the findings are certainly compatible with universal tendencies (as suggested in Table 12).

With regard to the path of development represented by the third column in Table 12 – i.e. the tendency cross-linguistically for relation markers to extend their meaning from spatial source to initial boundary to reason – there is even less of a basis for definite conclusions for Austronesian in my data sample. This is because of the fact (as noted in section 4.3.5 above) that con-
strucutions associated with initial boundary are rarely described for Austron-
sian languages. Speculatively, Austronesian languages have non-
conventionalized ways of expressing these relations, mostly involving lexi-
cal items to make the relation explicit, that do not always find their way into
descriptive grammars. The number of languages for which explicit construc-
tions for both initial boundary and reason relations were attested was quite
small (only 7 out of 30 in the core sample and 8 out of 43 in the extended
sample). Not surprisingly then, it was considerably more difficult to find
connections between initial boundary and reason. In fact, none of the lan-
guages showed any identical principal relation markers that were attested to
be used in both constructions explicitly expressing initial boundary and con-
structions explicitly expressing reason relations. However, if we loosen the
requirement that the relation markers should be identical, it is possible to
find some formally close relation markers that are used for initial boundary
and reason relations, respectively, such as Samoan talu ona, 'since', and talu
ai, 'because', and Tinrin ghegi bee, 'since', and ghegi wa drae rra, 'because'.
Tinrin ghegi featured in these complex relation markers is derived from ghe,
which is a spatial source preposition meaning 'from'.

193. Tinrin

a. \textit{nra} ta \textit{nri} \textit{nra} \textit{trar}u \textit{ghe} \textit{aroa} \textit{giwe}
   \hspace{1cm} 3.SG hit 3.SG.OBJ SBJ person from over.there mountain

   'The person from the mountain hit him.' (Osumi 1995, p 81)

b. \textit{nra} \textit{tru}u \textit{aroa} \textit{nra} \textit{hoowi}
   \hspace{1cm} 3.SG stay there SBJ fir

   \textit{ghegi} bee \textit{murru} \textit{nra} \textit{ro}
   \hspace{1cm} from COMP be.small POSS 1.SG

   'The fir tree has been over there since I was small.'
   (Osumi 1995, p 263)

c. \textit{ghegi} wa \textit{drae} \textit{rra} \textit{nra} \textit{vaju} \textit{nra} \textit{audre-nri}
   \hspace{1cm} from DET thing DEM 3.SG die SBJ mother-3.SG.POSS

   \textit{nra} \textit{nra} \textit{drarr}i
   \hspace{1cm} 3.SG PROG cry

   'Because his mother died, he is crying.' (Osumi 1995, p 278)

\footnote{The \textit{-gi} part of \textit{ghegi} is a locative/dative preposition meaning 'at', 'to', 'with' (Osumi 1995, pp 79-82).}
Some initial boundary relation markers and some reason relation markers are clearly derived from prepositions – mostly with elative ('from') or possessive ('of') meanings – and several initial boundary relation markers have the literal meaning 'begin'. Other literal meanings occasionally found are 'source' or 'trace'. Thus, despite very meager data, there is some indication that the connections between spatial source, initial boundary and reason are also found in the Austronesian languages. However, the most common polysemic pattern for Austronesian reason relation markers seems to be words meaning 'reason' or 'cause'. Further analysis would be needed in order to determine whether these words in turn tend to stem from words meaning 'source'. The table below gives some examples of literal meanings of Austronesian relation markers expressing initial boundary and reason.

Table 16.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>LITERAL MEANING</th>
<th>TC MEANING</th>
<th>RELATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tinrin</td>
<td>ghegi bee</td>
<td>from that</td>
<td>since</td>
<td>initial boundary</td>
</tr>
<tr>
<td>Buru</td>
<td>peltanek fi di</td>
<td>beginning from that</td>
<td>since</td>
<td>initial boundary</td>
</tr>
<tr>
<td>Coastal Konjo</td>
<td>mulai ri</td>
<td>beginning at</td>
<td>since</td>
<td>initial boundary</td>
</tr>
<tr>
<td>Tagalog</td>
<td>mula nang</td>
<td>beginning when</td>
<td>since</td>
<td>initial boundary</td>
</tr>
<tr>
<td>Tinrin</td>
<td>ghegi wa drae rra</td>
<td>from the thing</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Kusaiean</td>
<td>ke srihpen</td>
<td>at its reason</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Mekeo</td>
<td>ngome</td>
<td>source</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Buru</td>
<td>wahan</td>
<td>trace of</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Central Cagayan Agta</td>
<td>gafu ta</td>
<td>reason to</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Karo Batak</td>
<td>iban</td>
<td>caused by</td>
<td>because</td>
<td>reason</td>
</tr>
<tr>
<td>Rapanui</td>
<td>o</td>
<td>of</td>
<td>of</td>
<td>reason</td>
</tr>
<tr>
<td>Tagalog</td>
<td>sa dahilan</td>
<td>at reason</td>
<td>because</td>
<td>reason</td>
</tr>
</tbody>
</table>
5.2 Austronesian polysemic relation markers

The previous section showed that Austronesian data is clearly compatible with the directions of development illustrated for relation markers universally in Table 12 above. In this section, we will analyze all ten semantic relations in pairwise comparison to obtain a more complete picture of Austronesian polysemic relation markers. That is, what relation markers out of all those recorded in my database tend to be used for more than one semantic relation?

Table 17 illustrates the number of languages with identical relation markers for each pair of explicitly indicated relations for the core sample. Constructions for which the relation remains implicit, such as juxtaposed clauses and clauses related by a neutral relation marker (such as equivalents to the English coordinator *and*), have been excluded from the count.

Table 17. No. of languages with identical principal relation marker for each pair of explicitly indicated relations in the core sample (percentages in shaded cells)

<table>
<thead>
<tr>
<th></th>
<th>co-occurrence</th>
<th>posteriority</th>
<th>anteriority</th>
<th>terminal boundary</th>
<th>initial boundary</th>
<th>condition</th>
<th>concession</th>
<th>purpose</th>
<th>reason</th>
<th>result</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-occurrence</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>21</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>posteriority</td>
<td>13%</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>anteriority</td>
<td>10%</td>
<td>17%</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>term. boundary</td>
<td>17%</td>
<td>0%</td>
<td>3%</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>initial boundary</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>70%</td>
<td>10%</td>
<td>13%</td>
<td>13%</td>
<td>0%</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>concession</td>
<td>10%</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>23%</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>purpose</td>
<td>10%</td>
<td>7%</td>
<td>0%</td>
<td>10%</td>
<td>0%</td>
<td>17%</td>
<td>3%</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>reason</td>
<td>10%</td>
<td>7%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>0%</td>
<td>20%</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>result</td>
<td>7%</td>
<td>27%</td>
<td>0%</td>
<td>13%</td>
<td>0%</td>
<td>7%</td>
<td>0%</td>
<td>17%</td>
<td>13%</td>
<td></td>
</tr>
</tbody>
</table>

The figures for the extended sample are quite similar and therefore will not be displayed.
The table is intended to be read in the following way (by way of an example): 21 languages in the core sample (70%) have at least one principal relation marker that is used both in constructions explicitly expressing co-occurrence relations and in constructions explicitly expressing conditional relations.

In the following subsections, we will have a closer look at some of the relation pairs at the top of the list and their manifestations in Austronesian. The discussion will revolve around the top four:

- co-occurrence-condition (core sample freq.: 70%)
- posteriority-result (core sample freq.: 27%)
- condition-concession (core sample freq.: 23%)
- purpose-reason (core sample freq.: 20%)

as well as around two other interesting near-top relation pairs:

- purpose-result (core sample freq.: 17%)
- posteriority-anteriority (core sample freq.: 17%)

Posteriority-anteriority is interesting because of the apparent oddity of finding two opposing relations being represented by the same relation marker so close to the top of the list, and purpose-result is interesting because it ties in with the purpose-reason relation pair among the top four.

5.2.1 Co-occurrence-condition polysemy

As made obvious in Table 17, co-occurrence-condition is by far the most common relation pair to be represented by identical relation markers in the sample languages. This pair of relations has already been discussed above, but in this subsection, we will have a closer look at the ways different shades of meaning in the co-occurrence/conditional domain are expressed by Austronesian relation markers.

Even though most Austronesian languages have a relation marker that is used for both co-occurrence and condition, nearly all sample languages do also distinguish formally between specific types of co-occurrence relations and specific types of conditional relations. Often enough, past time co-occurrence and counterfactual condition are individually distinguished by means of designated TMA or other markers, or even by means of entirely different relation markers. Future co-occurrence and open conditional relations, on the other hand, are commonly expressed by identical constructions in Austronesian languages; that is, the distinction is not made at all. The reason is that neither future co-occurrence relations nor open conditional relations denote states of affairs ascertained to have taken place. Past time
co-occurrence relations, on the other hand, denote states of affairs that actually did occur. And counterfactual conditional relations denote imagined states of affairs that are known for certain not to have taken place. Thus, certainty of realization (or non-realization) of the states of affairs seems to be an important distinction in the domain of co-occurrence and conditional relations in Austronesian languages. However, this is not the whole truth of the matter. It turns out that almost all sample languages have overlapping sets of relation markers for co-occurrence and conditional relations. A pattern that occurs frequently is having a general relation marker for co-occurrence/condition (covering all or parts of this semantic domain), while at the same time having one separate relation marker for co-occurrence in general (or possibly only past co-occurrence) and one for condition in general (or possibly only counterfactual conditionals). The table below shows the patterns from some of the sample languages.

Table 18.

<table>
<thead>
<tr>
<th>language</th>
<th>relation marker</th>
<th>past co-occurrence</th>
<th>future co-occurrence</th>
<th>open condition</th>
<th>counter-fact cond.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erromanga</td>
<td>ndan</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>nempgon</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nagku</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>ina</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>'āfai</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'ana</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maori</td>
<td>i</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hei</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ina</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mehemea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mekeo</td>
<td>gonga</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>aidama</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>goà</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>emia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tinrin</td>
<td>mwa</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nri</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tra</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buru</td>
<td>eta</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bama</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mama</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Taba</td>
<td>polo</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
As the table shows, different relation markers within the same language have evolved to cover different spectra of meaning within the co-occurrence/conditional domain, with hypotheticality and certainty of realization being two important watersheds. Further study on the etymology of these relation markers may shed some light on patterns of development. With regard to the distinction between open conditional relations and counter-fact conditional relations, about half of the sample languages make this distinction overtly, either using different relation markers or by means of other markers in the constructions, such as TMA. Several languages of the remaining half might also use various grammatical means to make the distinction, although this is not stated in the sources. Some Austronesian languages, however, do not make this distinction by any formal means. Note, for instance, the Buru example below, which can be rendered an open conditional or a counter fact conditional according to context (Grimes 1991).

194. Buru

*eta du eru-k, petu du hulu-n*

if 3.PL agree-APPL then 3.PL cheer-DETRANS

Open: 'If they accept, then they cheer in celebration.' or
Counter fact: 'If they had accepted, then they would have cheered in celebration.' (Grimes 1991, p 424)

### 5.2.2 Posteriority-result polysemy

The next polysemic pattern on the list is that of posteriority and result relation markers. This pattern is not nearly as frequent as the co-occurrence-condition pattern. Nonetheless, almost one third of the core sample consists of languages with at least one relation marker used both for constructions explicitly expressing posteriority relations and for constructions explicitly expressing result relations. The polysemy here seems natural considering that any result must necessarily be posterior in time to the reason for its occurrence. Thus, in many situations the speaker can choose whether to stress the temporal posteriority relation or the resultative relation between two states of affairs. If a speaker uses a posteriority relation marker when a resultative relation can also be inferred from context, it is easy to see that the posteriority relation marker may take on a resultative flavor as well. This may well be what has happened in a number of Austronesian languages. An example again from Buru may serve to illustrate. In this language, the relation marker *petu* stresses sequentiality between the states of affairs, but it is
also often used for result relations. Both are construable for petu in the example below.

195. Buru

\[
\begin{align*}
&\text{wae dika-t sa moo,} \\
&\text{water other-NZR one NEG}
\end{align*}
\]

\[
\begin{align*}
&\text{petu geba-ro asu-k wae dii, petu du ino,} \\
&\text{SEQ person-PL dip-APPL water that SEQ 3.PL drink}
\end{align*}
\]

\[
\begin{align*}
&\text{tu du masa-k inaa-n, tu du foi} \\
&\text{with 3.PL cook-APPL food-GEN with 3.PL bathe}
\end{align*}
\]

'There was no other water, so people fetched that water, and then they drank (it), and they cooked their food (with it), and they bathed (in it).' (Grimes 1991, p 406)

Three verbs figure prominently as the origin of relation markers used for both posteriority and result relations in the sample languages: they are the equivalents of 'go', 'become' and 'finish'. Some examples are presented below.

\[
\begin{align*}
\text{Table 19.}
\end{align*}
\]

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>VERBAL ORIGIN</th>
<th>TC MEANING</th>
<th>RELATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetun</td>
<td>dadi</td>
<td>'becomes'</td>
<td>then, so</td>
<td>posteriority, result</td>
</tr>
<tr>
<td>Yakan</td>
<td>ubus</td>
<td>'finish'</td>
<td>and then, and so</td>
<td>&quot;</td>
</tr>
<tr>
<td>Hoava</td>
<td>gila</td>
<td>'and go'</td>
<td>and then, and so</td>
<td>&quot;</td>
</tr>
<tr>
<td>Araki</td>
<td>ale</td>
<td>French 'allez' (^{57})</td>
<td>then, so</td>
<td>&quot;</td>
</tr>
<tr>
<td>Erromanga</td>
<td>isuma</td>
<td>'that is all'</td>
<td>then, so</td>
<td>&quot;</td>
</tr>
</tbody>
</table>

These verbs are all associated with a transition from one state of affairs to another, and so they easily become markers for temporal succession, i.e. indicating the posteriority of the second state of affairs. The verb 'become', in addition, often hints at causation of some sort, which means it is especially prone to developing into a result relation marker. In the Western part of

\(^{57}\) Via Bislama ale, which has the same function and meaning as ale in Araki (François 2002, p 173).
the Austronesian Area (Western Malayo-Polynesian, Central Malayo- Polynesian and South Halmahera-West New Guinea), the word *dadi* has cognates in many languages, most of which are used as result relation markers. The table below displays the cognates found in the sample languages.

Table 20.

<table>
<thead>
<tr>
<th>Language</th>
<th>Region</th>
<th>Cognate</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indonesian</td>
<td>WMP</td>
<td><em>jadi</em></td>
<td>so, thus, therefore</td>
</tr>
<tr>
<td>Muna</td>
<td>WMP</td>
<td><em>dadi</em></td>
<td>so, consequently</td>
</tr>
<tr>
<td>Labuk Kadazan</td>
<td>WMP</td>
<td><em>jadi</em></td>
<td>so (then)</td>
</tr>
<tr>
<td>Sangir</td>
<td>WMP</td>
<td><em>diadi</em></td>
<td>so</td>
</tr>
<tr>
<td>Western Subanon</td>
<td>WMP</td>
<td><em>dadi</em></td>
<td>therefore</td>
</tr>
<tr>
<td>Kambera</td>
<td>CMP</td>
<td><em>njadi</em></td>
<td>so</td>
</tr>
<tr>
<td>Tetun (Fehan dialect)</td>
<td>CMP</td>
<td><em>dadi</em></td>
<td>then, so</td>
</tr>
<tr>
<td>Tabab</td>
<td>SHWNG</td>
<td><em>ndadi, dadi</em></td>
<td>so</td>
</tr>
</tbody>
</table>

Since all these languages are spoken in or in close proximity to Indonesia, it is tempting to think that they are loanwords from Indonesian. Whether this is in fact the case, or whether they have developed independently is not quite clear, although Taba *ndadi* appears to be of local origin (Bowden 1997). Another relation marker for result relations, with cognates in many Austronesian languages of the Southeast Asian area, is *maka*, which is sometimes used for posteriority relations as well (as in Easter Cham, Ma'anyan and Karo Batak).

### 5.2.3 Condition-concession polysemy

Moving on to the relation markers with a condition-concession polysemic overlap, the reason for the high ranking of these in the list is that a number of Austronesian languages have the same relation marker for concessive conditional relations and concessive relations, as is the case, for instance, in Tagalog.
196. Tagalog

a. \textit{maski} b[in]ayar-an mo sila,
\text{CONC} [PFTV]pay-{PF} 2.SG.GEN 3.PL.TOP

\textit{hindi} mo maipapagawa iyon sa kanila
\text{be.not} 2.SG.GEN ask.to.make.{PF} DEM.TOP to 3.PL.OBL

'Even though you paid them, you can't ask them to do that.'
(Schachter & Otanes 1972, p 480)

b. \textit{maski} pag[banta]an din niya ako,
\text{CONC} [threaten]{PF} too 3.SG.GEN 1.SG.TOP

\textit{hindi} ko i-bi-bigay ito sa kaniya
\text{be.not} 1.SG.GEN PF-RD:IRR-give DEM.TOP to 3.SG.OBL

'Even if he should threaten me too, I won't give this to him.'
(Schachter & Otanes 1972, p 480)

Cognates of \textit{maski} are common in concessive and concessive conditional constructions in the western part of the Austronesian area.\footnote{Apparently, its origin is not entirely clear. Morris (1984, p 139 [cited in Klinken 1999, p 313, fn 15]) claims that Tetun masik (which he lists as \textit{maski} for the eastern dialect) derives from Portuguese mais que, which can be used as a concessive relation marker (Rudolph 1996). The fact that the word can be found both in Tetun Fehan (masik, 'although', 'even if') and Indonesian (meski, 'although'), spoken in areas where Portuguese was formerly, or to some extent still is, an influential language, lends some support to this claim. Furthermore, it is also found in some Philippine languages (Tagalog and Central Cagayan Agta of the sample languages) used in areas where Spanish was the dominating colonial language; Spanish \textit{más que} can also, under certain circumstances, be used to indicate concession (Rudolph 1996). On the other hand, mais que and más que are apparently both quite rare as concessive markers in these languages, which of course does not preclude them as the source of borrowed relation markers for concession but reduces its likelihood. Also, Healey (1960, pp 71, 76) claims that Central Cagayan Agta \textit{maski} has a literal meaning 'it doesn't matter', 'it's alright', which is difficult to reconcile with a loan from Spanish \textit{más que}.}
concessive conditionals describe hypothetical states of affairs. Another way of not committing to a temporal reality is by using irrealis TMA markers. And although the grammatical details differ among the sample languages with identical explicit relation markers for concession and concessive condition, irrealis marking appears to be obligatory in some languages for the latter constructions, but never for the former. This is the case in Araki.

197. Araki

a. napdogo co usa,
   CONC 3.SG.IRR rain
   pla-m pera co ce levse īaudu
   CL-2.SG taro 3.SG.IRR NEG know live

   'Even if it rains, your taro won't be able to grow properly.'
   (François 2002, p 180)

b. napdogo mo usa,
   CONC 3.REAL rain
   pla-ku pera mo ce īaudu
   CL-2.SG taro 3.REAL NEG live

   'Although it rained / it is raining, my taro is not growing.'
   (François 2002, p 180)

With napdogo as relation marker, both clauses in the clause combining construction must take the irrealis TMA marker to be interpreted as a concessive conditional relation, while both clauses have to take the realis TMA marker to be interpreted as a concessive relation (François 2002).

5.2.4 Purpose-reason and purpose-result polysemy

The fourth most common relation pair represented by identical explicit relation markers among the sample languages is purpose and reason relations. This is illustrated for Yabem below.

---

60 The same distinction is found elsewhere in the world. An Indo-European example is provided by Spanish in which aunque translates as 'although' with indicative verb forms, but as 'even if' with subjunctive verb forms (Schmitt 1999).
198. Yabem

a. àwê òn kê-sa gê-mêŋ
   woman one 3.SG.REAL-arise 3.SG.REAL-come

   gebe ê-tê bu
   RM 3.SG.IRR-draw water

'A woman came up in order to draw water.'
(Dempwolff 2005 [1939], p 113)

b. ô-tôc intêna,
   2.SG.IRR-show path

   gebe aê ga-jam kauc
   RM 1.SG 1.SG.REAL-lack knowledge

'Show the way, because I don't know it.'
(Dempwolff 2005 [1939], p 114)

In the first example, the relation marker gebe precedes the purpose clause, i.e. representing the state of affairs logically postdating the other state of affairs in time, and in the second example, gebe precedes the reason clause, i.e. representing the state of affairs logically predating the other in time. Note that the relation marker encodes two opposite states of affairs in terms of their logical temporal order. However, purpose clauses marked by gebe in Yabem are obligatorily in irrealis mood, while reason clauses marked by gebe are obligatorily in realis mood, so TMA marking is providing the formal distinction in these cases. The use of identical or similar relation markers for purpose and reason relations is a widespread tendency in the world's languages (Thompson & Longacre 1985; Givón 1990), and as in Yabem, the semantic relations are normally distinguished by means of different TMA markers. Some Austronesian languages, however, extend the formal similarity between these relations and make no distinction at all between them. In Rapanui, for instance, nominalized clauses introduced by benefactive relation marker mo may denote both purpose and reason, depending on the context.61

61 Judging from examples in Du Feu (1996), it does seem that postposed mo clauses tend to get a purpose interpretation, while preposed mo clauses tend to get a reason interpretation, although Du Feu (1996) does not state if the reason and purpose interpretation of mo clauses always go hand in hand with a preposed and postposed distribution, respectively. However this may be, context does play some role in the interpretation of these constructions since mo clauses (both preposed and postposed) can also be interpreted as conditional clauses.
5. PATTERNS OF POLYSEMY

199. Rapanui

a.  \textit{i oho ai mo ho'o 'o'oku i te hare}
PST go PHO BEN buy 1.SG.POSS ACC SPEC house

'I went to buy a house.' (Du Feu 1996, p 53)  
(lit. '(I) went for my buying a house.')</b.  \textit{mo kore o te potu, he tere a nua}
for lack of SPEC cigarette ACT run PROP NAME

'Because there are no cigarettes, Nua has vanished.'  
(Du Feu 1996, p 55)

In a similar manner, some Austronesian languages show interesting interrelations between constructions encoding purpose and result relations. This is the case, for instance, in Araki, as is shown in the examples below.

200. Araki

a.  \textit{nam sovi naivou-ku ro}
1.SG.REAL wait wife-1.SG PROG

\textit{vada co sle naru-na lo hanhan}  
so 3.SG.IRR give child-3.SG OBL food

'I am waiting for my wife to feed the children.'  
(François 2002, p 183)

b.  \textit{mo ūei-a ūa vada uluvo mo le smat}
3.REAL do-3.SG go so young 3.REAL again look.smart

'He went on (rubbing) so the young man became handsome again.'  
(François 2002, p 183)

Also in these constructions, the relation marker is the same, while the marked clauses denote different semantic relations, i.e. purpose and result. Here, the relation marker \textit{vada} marks clauses denoting outcome in some sense in both examples, but one is an intended outcome (purpose), while the other is an actual outcome (result). This distinction is represented formally in a similar way as that between the Yabem examples above, viz. by means of irrealis marking for the purpose clause and realis marking for the result clause (cf. François 2002).

To understand the reason for these two formal overlaps – purpose-reason and purpose-result – it might be a good idea to separate two different strands of meaning inherent within the relations of result, purpose and reason: (i)
Both reason and purpose relations provide answers to the question 'why?'; reason clauses explain why a state of affairs occurs in terms of factual reasons, and purpose clauses explain why a state of affairs occurs in terms of the intentions on the part of an actor (cf. Thompson et al 2007). Result clauses, on the other hand, provide a representation of a development. Conversely, both purpose and result relations represent states of affairs logically postdating other states of affairs, while reason relations represent states of affairs logically predating others. These two strands of meaning may be labeled 'explanatory value' and 'logical order', respectively.

Table 21.

<table>
<thead>
<tr>
<th>RELATION</th>
<th>EXPLANATORY VALUE</th>
<th>LOGICAL ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result</td>
<td>development</td>
<td>postdating</td>
</tr>
<tr>
<td>Purpose</td>
<td>explanation</td>
<td>postdating</td>
</tr>
<tr>
<td>Reason</td>
<td>explanation</td>
<td>predating</td>
</tr>
</tbody>
</table>

When one relation marker covers more than one relation in the semantic domains of purpose, reason and result, it may do so either in adherence to the logical order of the states of affairs in the relations, in which case we get a result-purpose semantic overlap, or in adherence to the explanatory value, we get a purpose-reason overlap. As evident from the sample languages, both kinds exist in Austronesian languages.

5.2.5 Posteriority-antioriity polysemy

Finally, posteriority and anteriority is a relation pair perhaps not expected to be denoted by identical explicit relation markers in any languages, since they are each other's opposite. Still, 5 of the 30 core sample languages (and 6 of the 43 extended sample languages) display precisely this pattern. The reason in most cases seems to be that there is a relation marker that may function both as an adverbial subordinator and a relational adverb. This may occur, for instance, when the verb 'finish' grammaticalizes into a relation marker, since it may happen in several ways. A series of juxtaposed clauses, such as 'I cleaned the house, it finished, I went for a walk', may be the starting point in grammaticalization processes resulting in clause combining constructions equivalent to those in (201).
201. a. \textit{I cleaned the house, then I went for a walk.}\newline
from:\newline
\textit{I cleaned the house, (and when) it finished, I went for a walk.}\newline
b. \textit{I cleaned the house after I went for a walk.}\newline
from:\newline
\textit{I cleaned the house (when) it finished (that) I went for a walk.}\newline

In the (a) example, 'it finished' has developed into a relational adverb (~'then', 'afterwards'), and in the (b) example, it has turned into an adverbial subordinator ('after'). Kambera provides an authentic example, in which the relation marker \textit{hàla}, originating from the verb for 'finish' and often suffixed by perfective marker -\textit{ka}, may take on a meaning construable as either 'afterwards' (approaching an adverbial function) or 'after' (taking on a subordinator function) depending on the context (Klamer 1998).

202. Kambera

a. \textit{mbinu reti kudu, mbinu reti bakul}\newline
\textit{be.full stem be.small be.full stem be.big}\newline
\textit{hàla-ka kahaka, pakindi-nanya-ka}\newline
\textit{finish-PFTV grow.upwards swell-3.SG.CNT-PFTV}\newline

'The stem becomes full of small [leaves and it] becomes full of large [leaves]. Afterwards [the leaves] grow upwards and it [each ear of rice] is swelling.' (Klamer 1998, p 361)

b. \textit{hàla-ka pa-muti-ya}\newline
\textit{finish-PFTV COMP-harvest-3.SG.ACC}\newline
\textit{parina-du-nda-nya-ka}\newline
\textit{trash-EMPH-1.SG.GEN-3.SG.DAT-PFTV}\newline

'After harvesting it, we trash it.' (Klamer 1998, p 362)

In (202a), \textit{hàlaka} indicates that the growing up and swelling takes place posterior to the leaves appearing. Here, \textit{hàlaka} is an adverbial dependent of the verb \textit{kahaka}, 'grow upwards'. In (202b), on the other hand, \textit{hàlaka} indicates that the harvesting takes place anterior to the trashing. Here, \textit{hàlaka} takes the following verb as its complement. That is, the same relation marker is used to mark the two opposite semantic relations of posteriority and anteriority in Kambera.
In this chapter, we will take a closer look at some of the morpho-syntactic features found in the Austronesian clause combining constructions. When classifying clause combining constructions structurally, several different parameters may potentially be used. In RRG, for instance, the parameters of syntactic function (coordination, subordination, cosubordination) and level at which the clauses are linked (nucleus, core, clause, sentence) are used, the first parameter also employing a number of sub-parameters to establish its values (refer to chapter 2). Cristofaro (2005) uses solely formal features to classify subordinate clauses structurally: TMA marking, argument marking, use of case markers on the verb, and coding of verb arguments.

In classifying the Austronesian constructions in my data, the above parameters have been found useful, and parts of them have also been employed in pin-pointing the morpho-syntax of the sample language constructions. However, I found it important also to distinguish between constructions in which the semantic relation was made explicit or kept implicit for my Austronesian data. Therefore, my classification was made with regard to explicitness of semantic relation, to syntactic relation (i.e. the traditional subordination-coordination parameter), and to deviation of morpho-syntactic features from those in a declarative simplex clause (subsuming all of Cristofaro's (2005) features as well as parts of the RRG parameters). This, I feel, takes the most important structural features into consideration, without involving too many parameters. The parameters are discussed in section 6.1. The types resulting from combining the parameter values will be discussed in section 6.2 together with a tightness rank of the types.

6.1 Classifying parameters

The three parameters along which the Austronesian temporal and co-variational clause combining constructions were classified are listed below (with available values for each in parenthesis):
6. MORPHO-SYNTAX

As just mentioned, the parameters were chosen because they encapsulate the most common features associated with combined clauses. In the following subsections, each of the parameters will be discussed, including how their values were determined for constructions in the sample languages.

6.1.1 Deviating morpho-syntax

Clause combining constructions as defined in this thesis consist of two clauses – or clause-like syntagms – with or without one or more relation markers relating the clauses syntactically and/or their propositions semantically to each other. Recall from section 1.1 that, ideally, the two clausal syntagms should

(i) represent individual states of affairs semantically, and not be intimately linked as subcomponents of a single overall happening, and
(ii) constitute individual clauses or clause-like syntagms syntactically, and not be construable as two equal constituents on the same level of a single predicate.

These two points are meant to delineate clause combining constructions from serial verb constructions. Further, clausal syntagms in the clause combining constructions as studied here, should

(iii) be possible to interpret as one sentence, prosodically, syntactically and/or according to normal writing conventions.

This point is meant to delineate clause combining constructions from series of sentences in a discourse.

Together, points (i)-(iii) ensure a certain degree of structural homogeneity in the constructions studied in the different sample languages. Nonetheless, there is still a fair amount of morpho-syntactic variation between the constructions under study. The clauses (or clause-like structures) involved may deviate to a greater or lesser extent from the form of a simplex declarative clause in the language in which they occur. For instance, in the Tagalog (a) example below, the second clause has a structural pattern that cannot be
found in simplex declarative clauses in the language. It is therefore classified as deviating with regard to morpho-syntax. Compare (a) with (b), (c) and (d).

203. Tagalog

a. \( p[\text{in}]\text{arusa-han si juan} \)
   \[ \text{[PFTV]punish-PF TOP NAME} \]
   \[ \text{dahil sa pagka-salita' sa klase} \]

   because DAT NZR-talk DAT class

   'Juan was punished for having spoken in class.' (Schachter & Otanes 1972, p 159)

b. \( p[\text{in}]\text{arusa-han si juan} \)
   \[ \text{[PFTV]punish-PF TOP NAME} \]

   'Juan was punished.' (MJ)

c. *\( \text{pagka-salita' sa klase} \)
   \[ \text{NZR-talk DAT class} \]

   Intended: '[He] spoke in class.' (MJ)

d. \( \text{nag-salita' siya sa klase} \)
   \[ \text{AF.PFTV-talk 3.SG.TOP DAT class} \]

   'He spoke in class.' (MJ)

If, on the other hand, both clauses in a clause combining construction adapt to the structural pattern of simplex declarative clauses in the languages at hand, they display a non-deviating morpho-syntactic pattern. This is the case in the Samoan example below. Compare (a) with (b) and (c).

204. Samoan

a. \( \text{'ole'ā lē toe feiloa'i pale ma maria} \)
   \[ \text{FUT NEG again meet NAME and NAME} \]
   \[ \text{pe 'āfai 'ole'ā ō i niū sila} \]
   \[ \text{Q if FUT go.PL to NAME NAME} \]
   \[ \text{le aiga o maria} \]
   \[ \text{the family POSS NAME} \]

   'Pale and Maria will not meet again, if the family of Maria will go to New Zealand.' (AJ)
b. 'oleʻā lē toe feiloaʻi pale ma maria
FUT not again meet NAME and NAME

'Pale and Maria will not meet again.' (AJ)

c. 'oleʻā ʻō i niū sila le aiga o maria
FUT go.PL to NAME NAME the family POSS NAME

'The family of Maria will go to New Zealand.' (AJ)

Note that the parameter of deviating morpho-syntax does not take a possible relation marker into account, so that the presence or absence of one does not affect the parameter value. Nor does the parameter value have anything to do with whether the clauses are coordinate, or one subordinate to the other. In some studies, structural deviation and syntactic relation between the clauses are interconnected parameters, in that a deviating clause is seen as a sign of subordination (see chapter 2). However, as Haiman & Thompson (1984) point out, morpho-syntactic deviation (and in their case, specifically morpho-syntactic reduction) is not a trait exclusive to subordinate clauses. Conversely, subordinate clauses are not necessarily morpho-syntactically deviant, as stressed, for instance, by Stassen (1985) in his characterization of balanced and deranked constructions.

Cliticized coordinators and subordinators are treated on par with free coordinators and subordinators in this study. As clitics, they are not integrated into the clause and are not considered as producing a deviating clause pattern. An example is the Manam coordinating morpheme be, which is often cliticized to the last constituent of the previous clause.

205. Manam

ŋa tágo u-móre-be màsa m-malípi
1.SG NEG 1.SG.REAL-be.sick-and IRR 1.SG.IRR-work

'If I am not sick, I will work.' (Lichtenberk 1983, p 524)

Clausal nominalizations are fairly common examples of deviating patterns in clause combining. Explicitly nominalizing morphology is often present in these cases, as in the Tagalog example (203) above. But an explicit nominalizer is not needed in all cases to create a nominalized clause. A nominalization may be indicated by the use of typical noun phrase morphology (free or bound), such as possessive pronouns or other determiners. This is also a deviating pattern if it occurs in complex clauses in a way not possible in simplex declarative clauses. Consider the following Maori example.
206. Maori

\[
\text{e haere ana ahau ki te tāone}
\]
CNT move CNT 1.SG to the town

\[
\text{ki te hoko mai i tētahi koha}
\]
to the barter hither ACC one gift

'I am going to town to buy a present' (Bauer 1993, p 66)

In this case, the head of the (would-be) verb phrase, hoko, 'barter', retains its ability to combine with an adverbial particle, mai, and to take an object, i tētahi koha. The phrase is unquestionably nominal, however, because it is determined by the specific article te.

Less common but also found among the sample languages are converb forms, which derive adverbial forms from clauses. Such constructions are, of course, also classified as instances of deviant morpho-syntactic if occurring in ways not possible in simplex clauses. An example is the Labuk Kadazan construction below.

207. Labuk Kadazan

\[
pog-tingga-Ø-ku na[antang]an-ku rumah-sakit ranau
\]
IMM-look.up-PF-1.SG.ACTR RF[see]-1.SG.ACTR house-sick Ranau

'As soon as I looked up I saw "Ranau Hospital"' (Hurlbut 1988, p 73)

The Labuk Kadazan prefix pog- is cognate with similar prefixes in Philippine languages (pag- in Tagalog), in which it is often referred to as a gerund form primarily with a nominalizing function. In Labuk Kadazan, however, its principal function appears to be that of a converb prefix, turning the verb into an adverb with a specific temporal meaning (cf. Hurlbut 1988, 1990).

In connection with converb forms, the so-called 'citation form' in Erromangan should be mentioned. This is a special verb form used as an alternative to regularly inflected verbs in some specific contexts, including, for instance, head-tail linkage. In such contexts, it takes on the flavor of a converb form.

208. Erromangan

\[
kamu-tew-i mavel-i yi-tamul-i.
\]
1.DU.EX.PST-wait.for-3.SG until-CONST 3.SG.PST-send-3.SG

\[
tamul-i kamli-vai.
\]
CIT.send-3.SG 1.PL.EX.PST-take

'The two of us waited until he sent it. After having sent it, we took it.' (Crowley 1998, p 118)
Citation forms of vowel initial verb stems take a prefix -n, while other verb stems remain the same. The citation form prefix is identical to the nominalizing prefix in Erromangan, although they are applied according to slightly different rules (see Crowley 1998, pp 46ff and 116ff). Neither nominalized verbs nor citation forms allow any further prefixes. However, while object suffixes may be used on citation forms (as in the example above), nominalizations take only possessive pronominal marking. Although citation forms may also occur in some simplex clause contexts (such as telegraphic language, lists and corrections), they are not used in regular declarative simplex clauses and are therefore classified as constituting deviating morpho-syntax when used in clause combining.

The lack of grammatical morphology or an argument in clause combining, obligatory in corresponding simplex clauses, also constitutes deviating morpho-syntax. Several Formosan languages, for instance, have constructions in which the actor participant must be interpreted to be identical in the two clauses and can only be overtly expressed in one of them, either the first or the second one, but not in both (Liu 2003a). Examples are given from Amis below.

209. Amis

a. ma-fucal tu ci aki k[um]aqen tu tali
   AF-full INCH NOM Aki eat[AF] ACC taro

   'Aki ate taros until he became full.' (Liu 2003a, p 63)

b. ma-fucal k[um]aqen ci aki tu tali
   AF-full eat[AF] NOM Aki ACC taro

   'Aki ate taros until he became full.' (Liu 2003a, p 63)

The most common grammatical markers left out in clause combining are TMA markers, which in Melanesia also often include reference to person and number of the subject. Consider examples (210) and (211) below.

210. Longgu

mo go oli / kote aluburi-o mai
and OBL.2.SG return lest follow-2.SG hither

mama ngaia vua nau
father 3.SG grandchild 1.SG

'And you must return lest my grandchild's father follows you here.'
(Hill 1992, p 284)
211. Tagalog

\{oras na\} ma-kita ko siya
as.soon.as PF-see 1.SG.GEN 3.SG.TOP

i-ta-tapon ko ito sa mukha niya
PF-RD:FUT-throw 1.SG.GEN this.TOP to face 3.SG.GEN

'The minute I see him, I'll throw this in his face.'
(Schachter & Otanes 1972, p 471)

Any of the TMA particles that are normally available to Longgu clauses (such as modal go in the initial clause above) would be ungrammatical in the second clause of example (210) (Hill 1992). TMA particles are generally optional in the language, but they are obligatorily absent in clauses introduced by kote, 'lest'. Therefore, this and similar constructions have been classified as morpho-syntactically deviant. Tagalog TMA markers interact in complex ways with the focus marker, but a TMA category is normally obligatorily indicated in simplex clauses. In some clause combining constructions, however, the TMA marker is absent from one of the clauses, leaving the focus marker as the sole grammatical marker on the verb, as in the first clause of the Tagalog example above. Such forms are sometimes called infinitive forms. Naturally, they are also classified as morpho-syntactically deviant.

Let us now turn to some clause combining constructions that have not been rendered to display deviating morpho-syntax, although they possess features specific to clause combining. This is the case, for instance, where there is a reduction in paradigmatic potential in one of the clauses in clause combining (e.g. restrictions in the TMA oppositions available) as compared to the corresponding simplex clause. An example can be taken from Kusaiean.

212. Kusaiean

a. fin kom ac tuhkuh, nga ac soanwe-kom
   if 2.SG FUT come 1.SG FUT wait-2.SG

   'If you come, I will wait for you.' (Good 1989, p 168)

b. fin kom *in tuhkuh, nga ac soanwe-kom
   if 2.SG IRR come 1.SG FUT wait-2.SG

   Intended: 'If you come, I will wait for you.' (Good 1989, p 168)
Conditional clauses in this language cannot take the irrealis TMA marker *in*, while other TMA markers may freely occur. Good (1989) suggests that the relation marker *fin* incorporates an irrealis feature that renders irrealis TMA marking superfluous. Here, the structural make up of the clause is identical to a potential simplex clause, but a certain morphological category displays fewer values in one of the clauses in clause combining than does the same category in a simplex clause. Another example of the same type can be taken from many Formosan languages in which the verb is locked to an actor focus form in some clause combining constructions, while other focus forms are ungrammatical. All these cases have been regarded to display non-deviating morpho-syntactic patterns.

Another example of features specific to clause combining that have not been classified as deviating morpho-syntax is when one of the clauses contains a morpheme (lexical or grammatical) exclusive to clause combining constructions but part of a class of markers occupying a regular syntactic slot common to both simplex and complex clauses. This normally pertains to relational adverbs and a small number of particles. The Muna adverb *kirakira*, 'about to', is an example. It may only occur in clause combining constructions, but its syntactic distribution mirrors that of several other adverbs in Muna which occurs in both complex and simplex clauses, for instance *bhahi*, 'maybe' (Berg 1989). Compare the sentences below.

213. Muna

a. *kirakira* no-maho-mo na-rumako-da
   about.to 3.SG.REAL-nearly.happen-PFTV 3.SG.IRR-catch-3.PL

   *garaa* no-tumbu-mo *patu*
   suddenly 3.SG.REAL-grow-PFTV  *bamboo*

   'When she was about to catch them, suddenly there grew a bamboo bush.' (Berg 1989, p 248)

b. *bhahi* a-m-unda we se-wata-no *ini*
   maybe 1.SG.IRR-I RR-jump LOC one-side-IPFV this

   'Maybe I could jump to this side.' (Berg 1989, p 186)

---

62 For this reason, Berg (1989) classifies *kirakira* as a conjunction and suggests a division between dependent conjunctions, which like *kirakira* can only occur in clause combining constructions, and free conjunctions, which like, for instance, *pasino*, 'then', may also occur in simplex clauses, indicating a relation to the previous context. Berg (1989) admits that the latter group could well be alternatively classified as adverbs. I am inclined to view many of the items in both these groups as adverbs because of their distributional patterns. Other items in these groups, however, have morphological properties indicating that they are verbs or nouns, or have a history as such.
Since *kirakira* seems to have no impact on the morpho-syntax of the clause different from that of other clause initial adverbs, *kirakira*-clauses are not regarded as morpho-syntactically deviant.

Finally, the order of a clause's constituent parts may be altered in clause combining. But as long as the shuffling around of items in the clause does not have the effect of altering its categorical status, e.g. nominalizing it, I have not viewed it as a case of deviating morpho-syntax. In Tboli, for instance, the use of a certain class of relation markers causes focused actors – pronominal as well as lexical ones – to occur before the main predicate rather than in their usual position after the main predicate (214) (cf. the word order of the first and second clause of the example).

214. Tboli

```
igò  kasi  d[m]adu  /  kól  walan
while  NAME  [AF]plow  arrive  NAME
```

'While Kasi was plowing, Walan arrived.' (Porter 1977, p 24)

In Tboli simplex clauses, any participant targeted by the focus form of the verb can be optionally preposed for emphasis (Forsberg 1992, p 55). However, in clause combining, with the relevant class of relation markers, a focused actor is obligatorily preposed (Porter 1977, pp 22-3). In the West Malayo-Polynesian area and in some Formosan languages, it is very common in the presence of certain clause initial relation markers that this preposing pattern applies only to pronominal forms, which otherwise normally follow the predicate in these languages. However, in Tboli it applies also to lexical arguments. Pronominal preposing is shown for Tagalog in (215) below (cf. the position of the pronoun *siya*, 'he', in the two clauses in the example). In some constructions, the shift is optional; in others, it is compulsory.

215. Tagalog

```
[mag-a]-aral  siya  ng  liksyon  niya
[FX-RD]:AF,FUT-study  3.SG.TOP  GEN  lesson  3.SG.GEN
```

```
bago  siya  ma-tulog  gabi-gabi
before  3.SG.TOP  AF-sleep  night-night
```

'He will study his lessons before he goes to bed every night.' (Schachter & Otanes 1972, p 473)
In the West Malayo-Polynesian languages, this preposing of pronominal forms often applies in simplex clauses as well in the presence of certain clause initial syntactic items, normally negators and auxiliaries.

6.1.2 Syntactic relation between clauses

Let us now have a look at the next parameter, which concerns the syntactic relation between clauses, i.e. whether they are coordinate or one is subordinate to the other. There are a number of difficulties involved in determining this relation in general, as should be obvious from the discussion in section 2.1 of coordination and subordination in clause combining. The specific Austronesian constructions investigated here are no exception, and it has sometimes been very difficult to pinpoint the syntactic relation between the clauses in the data. Therefore, although a classification into subordination or coordination has been made for all constructions encountered, some of the classifications along this parameter rest on rather shaky ground, either because the source material does not provide enough information to perform a series of syntactic tests or because the tests give ambiguous results. However, in the majority of cases, a classification was in fact readily obtainable from the information in the source materials and the application of syntactic tests. Therefore, I do not think that the uncertain cases will have obscured the overall picture in any serious way.

The following criteria are commonly used as indicators of subordination (see chapter 2 for details) and have been used whenever possible also on the Austronesian data:

- **pronominal cataphoric reference**
  This is normally only possible from a preposed subordinate clause to the following main clause:
  
  Before he went to bed, Peter read for a while.
  *He went to bed and Peter read for a while.

- **flexible clause order**
  Subordinate clauses can often occur both postponed and preposed without a change in meaning, while this is impossible for most coordinate clauses:
  
  Ben left when Bob arrived. When Bob arrived, Ben left.
  Ben left but Bob arrived. *But Bob arrived, Ben left.

- **insensitivity to tense iconicity**
  While the linear order of coordinate clauses tend to mirror temporal order, this is not necessarily so with subordinate and main clauses:
  
  Because we turned up the music, John left the party.
  John left the party and we turned up the music.
• **inapplicability of the coordinate structure constraint (CSC)**
  Complement constituents can be extracted out of a matrix clause with a subordinate clause connected to it, but not out of a coordinate clause:
  
  *What did he say when he saw you?*
  *What did he say and he saw you?*

• **lack of assertive power**
  Subordinate clauses, but not coordinate clauses, tend to lack assertive power, as indicated by tag questions, for instance:
  
  *He left before you fell asleep, didn't he? / *didn't you?*
  *He left and you fell asleep, didn't you? / *didn't he?*

A paradigm Austronesian example can be taken from Samoan. Consider the example below in which the relation marker 'ona,' 'because', indicates that the second clause denotes a reason for the state of affairs denoted in the first.

216. **Samoan**

```
sa alu le tama e 'aumai le fana
 PST go the boy GNR bring the gun

'ona sa ia va'ai-a le pua'a 'ai-vao
because PST 3.SG see-AUG the pig eat-leaf

'The boy went to get his gun, because he saw the wild boar.' (AJ)
```

In (216), the second clause denotes an event that precedes the one in the first clause. Thus, there is no tense iconicity. The order of the clauses may be changed, as in (217) below, in which the reason clause occurs initially, preceded by the relation marker 'ona. The criterion of flexibility in clause order is thus fulfilled. Furthermore, in (217), the third person pronoun *ia* occurs before its antecedent *le tama*, 'the boy', which means that pronominal cataphoric reference is possible.
6. MORPHO-SYNTAX

217. Samoan

"ona sa ia va'ai-a le pua'a 'ai-vao
because PST 3.SG see-AUG the pig eat-leaf

sa alu ai 63 loa 63 le tama e 'aumai le fana
PST go ANA PART the boy GNR bring the gun

'Because hei saw the wild boar, the boyi went to get his gun.' (AJ)

Thus, by the criteria of pronominal cataphoric reference, flexible clause order and insensitivity to tense iconicity, the reason clause can be identified as a subordinate clause. The fact that the relation marker 'ona must be prefixed to the reason clause, and thus that *Sa ia va'aia le pua'a 'aivao, sa alu ai loa le tama e aumai le fana 'ona is ungrammatical tells us that it is indeed the reason clause, and not the result clause, that is subordinate here.

Further confirmation that subordination is involved is supplied by the fact that extraction is allowed out of the first clause in example (216) above, as is evident in example (218).

218. Samoan

'o le ā le mea sa alu le tama e 'aumai
PRES the what the thing PST go the boy GNR bring

'ona sa ia va'ai-a le pua'a 'ai-vao
because PST 3.SG see-AUG the pig eat-leaf

'What did the boy go and get because he saw the wild boar?' (AJ)

In other words, the coordinate structure constraint (CSC) does not apply, which indeed is a sign of subordination.

Finally, we may subject the reason clause to alternative negation, as in example (219), to see whether it has its own assertive power or not.

63 Ai is an anaphoric element referring back to a previous setting, and loa is a discourse particle meaning 'at this point'. They are, strictly speaking, not needed for grammaticality here, but they contribute coherence and make the sentence more idiomatic.
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

219. Samoan

\[sa\ \textit{alu le tama e 'aumai le fana}\]
\[PST\ \textit{go the boy GNR bring the gun}\]

\[e \ \textit{lē 'ona sa ia va'ai-a le maile}\]
\[GNR\ \textit{NEG because PST 3.SG see-AUG the dog}\]

\['ae 'ona sa ia va'ai-a le pua'a 'ai-vao\]
\[but\ \textit{because PST 3.SG see-AUG the pig eat-leaf}\]

'The boy went to get his gun, not because he saw a dog, but because he saw the wild boar.' (AJ)

Evidently, the reason clause can be contrasted with the alternative negated structure, which is only the case if it has no assertive power of its own. As assertive power is a property of main clauses, the reason clause is once again demonstrated to be a subordinate clause.

For illustrative purposes, we may compare the subordinate construction that was just examined with a semantically similar but structurally opposite construction in the same language, opposite in the sense that applying the same subordination-coordination tests as before will yield the opposite result in each case. Thus, the sentence in (220) is a paradigm example of coordination.

220. Samoan

\[sa\ \textit{va'ai-a e le tama le pua'a 'ai-vao}\]
\[PST\ \textit{see-AUG ERG the boy the pig eat-leaf}\]

\[ma\ \textit{sa ia alu ai loa e 'aumai le fana}\]
\[and\ \textit{PST 3.SG go ANA PART GNR bring the gun}\]

'The boy saw the wild boar and (so) he went to get his gun.' (AJ)

The sentence is interpreted according to tense iconic principles, and a reason-result relation is normally inferred from the propositional content. Reversing the order of the clauses with the relation marker \textit{ma}, 'and', remaining between them would result in the states of affairs occurring in the opposite temporal order. Consider example (221) below (with the positions of \textit{ai} and \textit{loa} accommodated to their anaphoric functions).
221. Samoan

\[
\text{sa \ ia \ alu \ e \ 'aumai \ le \ fana} \\
\text{PST \ 3.SG \ go \ GNR \ bring \ the \ gun}
\]

\[
\text{ma \ sa \ va'ai-a \ ai \ loa \ e \ le \ tama \ le \ pua'a \ 'ai-vao} \\
\text{and \ PST \ see-AUG \ ANA \ PART \ ERG \ the \ boy \ the \ pig \ eat-leaf}
\]

'Hei went to get his gun, and (then) the boy saw the wild boar.' (AJ)

Moreover, the third person pronoun cannot be interpreted cataphorically, so \text{ia} and \text{le tama}, 'the boy', would refer to two different persons. Putting \text{le tama} in the first clause and \text{ia} in the second as to allow anaphoric pronominal reference would just emphasize the reversed order of the states of affairs (compared to example (220)) and thus provide further evidence of tense iconicity and of a coordinate structure.

The related criterion of clause flexibility also fails to detect subordination in this case. Interposing the clauses together with the relation marker kept in its linear position (as though it was a subordinator) yields either (222) or (223) (with the positions of \text{ai} and \text{loa} again accommodated to their anaphoric functions).

222. Samoan

\[
\text{?ma \ sa \ ia \ alu \ e \ 'aumai \ le \ fana} \\
\text{and \ PST \ 3.SG \ go \ GNR \ bring \ the \ gun}
\]

\[
\text{sa \ va'ai-a \ ai \ loa \ e \ le \ tama \ le \ pua'a \ 'ai-vao} \\
\text{PST \ see-AUG \ ANA \ PART \ ERG \ the \ boy \ the \ pig \ eat-leaf}
\]

'?And he saw the wild boar, the boy went to get his gun.' (AJ)

223. Samoan

\[
\text{*sa \ alu \ le \ tama \ e \ 'aumai \ le \ fana} \\
\text{PST \ go \ the \ boy \ GNR \ bring \ the \ gun}
\]

\[
\text{sa \ ia \ va'ai-a \ ai \ loa \ le \ pua'a \ 'ai-vao \ ma} \\
\text{PST \ 3.SG \ see-AUG \ ANA \ PART \ the \ pig \ eat-leaf \ and}
\]

'*The boy went to get his gun, he saw the wild boar and.' (AJ)

The second of these examples is definitely ungrammatical, while the first one is, if not ungrammatical too, still distinctly odd. Syntactically, the use of \text{ma} sentence-initially followed by two juxtaposed clauses is quite unconvenc-
tional according to my Samoan informant. If a semantic interpretation is attempted, the clauses tend to be rendered tense iconically, here as in (221), which changes the meaning of the sentence. Neither would cataphoric pronominal reference be possible. In other words, the construction does not allow flexible clause order in the way that many subordinate clause constructions do.

To follow through with the tests applied to the subordinate construction in the previous section, we can also establish that the CSC applies, and that alternative negation is impossible (thereby demonstrating lack of assertive power), as shown by the ungrammaticality of (224) and (225), respectively.

224. Samoan

*'o le 'ā sa va'ai-a e le tama
CASE the what PST see-AUG ERG the boy

ma sa ia alu ai e 'aumai le fana
and PST 3.SG go ANA GNR bring the gun

*'What did the boy see, and he went to get his gun?' (AJ)

225. Samoan

*sa alu le tama e 'aumai le fana
PST go the boy GNR bring the gun

e lē ma sa ia va'ai-a le maile
GNR NEG and PST 3.SG see-AUG the dog

'ae ma sa ia va'ai-a le pua'a 'ai-vao
but and PST 3.SG see-AUG the pig eat-leaf

*'The boy went to get his gun, not and he saw a dog, but and he saw the wild boar.' (AJ)

In other words, these tests too indicate the construction to be coordinate.

In the majority of cases where one or more of the subordinate-coordinate tests could be performed, they did indicate a classification of the Austronesian constructions concurring with the one made in the source materials. But in a few instances, a different classification was indicated. In Kambera, for instance, Klamer (1998) classifies the relation marker jàka, 'if', as a coordinating conjunction, and thus, the sentence below as an instance of coordination.64

64 Klamer (1998) describes only complement clauses and relative clauses in her chapter on subordinate clauses and views all other instances of clause combining as cases of coordina-
226. Kambera

\[
\text{pa-hilu-ya \quad jàka \ na-ruhak}
\]
CAUS-exchange-3.SG.ACC if 3.SG.NOM-be.broken

'Change it if it is torn.' (Klamer 1998, p 111)

However, by the criteria of flexible clause order and insensitivity to tense iconicity, the construction appears to involve subordination, as the conditional clause, together with the relation marker, may also occur initially in the sentence without a change in overall meaning.

227. Kambera

\[
jàka \ nda \ na-mài \ meti-nggunya
\]
if NEG 3.SG.NOM-come die-1.SG.CNT

'If he doesn't come, I'll die.' (Klamer 1998, p 108)

Here, and in similar cases, the implications of the criteria were followed, and Kambera conditional clauses introduced by jàka were classified as instances of subordination.

In a few cases, tests according to the coordination-subordination criteria showed contradicting results. This sometimes happened when the relation marker was a verbal predicate (or could be shown to originate from one). Consider the conventional way of expressing concessive conditionals in Samoan, displayed by the examples below, in which the relation marker e tusa lava, 'even if' (lit. 'is indeed the same'), a verb phrase, accompanies its clause in both initial and final position in the sentence without any change in meaning.

228. Samoan

a. 'ou te alofa 'iā 'oe
1.SG GNR love to 2.SG

e tusa lava pē 'e te ita 'iā a'u
GNR be.same indeed Q 2.SG GNR hate to 1.SG

'I love you, even if you hate me.' or
'I love you, it is the same if you hate me.' (AJ)

This distinction seems to be similar to Halliday's (1994) distinction between embedded clauses (mainly complement clauses and relative clauses) and tactic clauses (mainly adverbial clauses and coordinate clauses). However, whereas Klamer (1998) makes no further division of what she refers to as coordinate clauses, Halliday (1994) separates his category of tactic clauses into hypotaxis and parataxis.)
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

b. e tusa lava pē 'e te ita 'iā a'u
   GNR be.same indeed Q 2.SG GNR hate to 1.SG

'ou te alofa 'iā 'oe
1.SG GNR love to you

'Even if you hate me, I love you.' or
'It is the same if you hate me, I (still) love you.' (AJ)

The criterion of flexible clause order thus indicates that the concessive clause is subordinate in this case. While this is true for the relation between e tusa lava, which takes a complement, and the concessive conditional clause, which fills this complement slot, it is less clear for the relation between the entire e tusa lava syntagm and the other clause, since neither extraction nor alternative negation are possible here. Rather, this relation is most naturally seen to involve coordination by juxtaposed units, as indicated by the alternative translation in both examples. Nevertheless, as the concessive conditional clause in the Samoan examples is the complement of the verbal relation marker e tusa lava, it is still subordinate, and consequently, this construction has been classified as involving subordination.65 The same goes for all similar Austronesian constructions in which it is clear that one of the clauses functions as the complement of a verbal relation marker. This seems natural, as the development from verb to relation marker is not uncommon among the Austronesian languages.66

However, the transformation into subordinators is not the only path available for verbal relation markers. Some relation markers with an apparent verbal origin are also involved in coordinate structures. Typically, they have come to function as adverbs or coordinators. In Mbula, for instance, Bugenhagen (1995) analyzes so as a modal adverb of hypotheticality when used without a pronominal agreement prefix in conditional clauses. With this prefix, it is a verb with the meaning 'say' (compare the (a) and (b) examples of 229 below). And in Hoava, Davis suggests that the verb for 'finish' has grammaticalized into a coordinator (compare the (a) example with the (b) and (c) examples of 230).

65 Mosel & Hovdhaugen (1992) classify the Samoan concessive conditional discussed as dependent but not embedded (i.e. cosubordinated in RRG terms).

66 The fact that e tusa lava is often rendered tusa lava (without the TMA particle) is perhaps an indication that the Samoan concessive conditional construction is heading in this direction too. Another indication is that my Samoan informant, though hesitantly, allowed for cataphoric pronominal reference in the construction in the equivalent of the following sentence: 'Even if you hate him, the boy loves you.'
6. MORPHO-SYNTAX

229. Mbula

a. *to-na kon moori ta-na i-so p-ini...*
then-GIV ghost woman that-GIV 3.SG-say OBL-3.SG

'After that, that ghost woman said to her…'
(Bugenhagen 1995, p 328)

b. *so ti-posop uraata, so aŋ-giimi zin*
supposedly 3.PL-finish work supposedly 1.SG-buy 3.PL

'If they had finished the work, I would have paid them.'
(Bugenhagen 1995, p 277)

230. Hoava

a. *dae puta sa t[in]avete-na*
be.finished indeed ART.SG [NZR]work-3.SG.POSS

*sa gugusu henì*
ART.SG village this

'The work of this village was totally finished.' (Davis 2003, p 260)

b. *mamaqa sa ŋuzu-na sa noki*
open.wide ART.SG mouth-3.SG.POSS ART.SG snake

*dae ŋan-i-a sa kutu*
and(.then) eat.TR-3.SG ART.SG rat

'The mouth of the snake opened wide, and (then it) ate the rat.'
(Davis 2003, p 260)

---

67 In this type of clause, *so* is required in the protasis (always the first clause) but is only one of several possible adverbial items in the apodosis. When *so* is used in the apodosis, the construction is interpreted as a counter fact conditional (Bugenhagen 1995, p 277).

68 Although Davis (2003) translates *dae* as 'and', it seems quite likely that a sequential interpretation is valid here (as well as in the (c) example), given the origin of the word. She does state that the word is used to link two sequential clauses. This seems to suggest a possible alternative analysis of *dae* as an adverb, in which case its development would be parallel to that which Bugenhegen (1995) suggests for *so* in Mbula. Davis (2003), however, consistently ascribes a coordinating function to *dae* (and translates it 'and') when it is not the main predicate of the clause.
In general, to the extent that languages have relation markers originating from verbs, I am under the impression that they are more prone to being analyzed by researchers as having developed into adverbs or coordinators (rather than into subordinators) if the language also makes use of serial verb constructions (as do Mbula and Hoava). If there are no salient serial verb constructions, verbal relation markers tend to be analyzed as developing into subordinators. Whether this correlation can be confirmed, however, is not within the scope of this thesis.

6.1.3 Explicitness of semantic relation

The semantic relation between the states of affairs of the clauses in clause combining may be indicated explicitly, by means of a morpheme with a semantically explicit relational meaning, or it may be kept implicit, when no such morpheme is present. Austronesian languages often have the option of keeping the semantic relation implicit between the clauses in clause combining, either by simply juxtaposing two clauses, or by using a neutral relation marker (equivalent, for instance, to the English coordinator and). Several authors of Austronesian reference grammars remark on the frequent use of semantically implicit clause combining constructions. For Seediq, for instance, Holmer (1996) writes:

"Many situations where subjunctions would be expected in other languages are solved neatly without them in Seediq. […] Serialising subordinate clauses such as those introduced by 'before', 'after' are rendered in Seediq by two main clauses in series." (Holmer 1996, p 59)

Similar statements are made for languages from throughout the Austronesian area. In Araki, François (2002) reports that the most common way of relating clauses in discourse is by juxtaposition, although there are also other more explicit ways of relating clauses. Thurgood (2005) states of Eastern Cham that coherence in clause combining is normally achieved by contextual fac-
tors and the iconicity of sequencing alone, and that explicit relation markers are used only when deemed necessary to avoid misunderstanding. A similar situation is presented for Indonesian, although semantically specific relation markers are becoming the norm in the emerging formal literary tradition (Ewing 2005). For colloquial Indonesian, however, as attested by Englebretson (2003), it is more common than not to leave the relation between clauses implicit:

"Sometimes the relationship between clauses is overtly indicated by conjunctions, connectives, or discourse markers, but more often in the database this relationship is left to contextual inference on the part of the interlocutors." (Englebretson 2003, p 38)

Also in the older Malay varieties on which Indonesian is based, before extensive contact with connector heavy languages such as Hindi, Arabic, and Dutch, juxtapositional constructions seem to have been extensively employed in clause combining (cf. Macdonald 1976).

It could of course be argued that two juxtaposed clauses, for instance, may be used to represent any semantic relation if the propositional content and contextual cues are right, but in light of the above, it seems clear that implicit means of creating clause combining constructions is, on a token basis, a very salient strategy in many Austronesian language. In addition, as we shall see below, some subtypes of semantically implicit clause combining constructions are intimately connected with a specific TC relation in many of the sample languages. And for some languages, the only clause combining construction attested to represent one of the relevant TC relations is an implicit one. For all these reasons, to exclude semantically implicit constructions entirely from the analysis would be to miss a central fact about Austronesian languages.

The implicit constructions included in the study have been limited to those that are specifically indicated to represent one or more TC relations in the sources of data. However, since grammatical descriptions are perhaps especially prone to varying in detail with respect to implicit constructions, the data may not be completely representative for all languages here. It may well be that relationally implicit clause combining constructions are more common than the present study will indicate (cf. chapter 7).

The presence of a semantic parameter to define structural types might seem odd, considering the very purpose of defining structural types here is to investigate the correlation between structure and meaning. One must keep in mind, however, that the parameter only superficially pertains to semantics, since the presence or absence of a semantically explicit relation marker has nothing to do with specific relational meanings. There is no necessary or inherent correlation between the presence or absence of a semantically ex-
plicit relation marker and the actual semantic relation represented in the relevant construction. We have no reason a priori to postulate that the absence of an explicit relation marker entail certain relational meanings (although our investigations may show such a correlation; see further details in chapter 7).

Semantically explicit constructions make use of several kinds of relation markers – some clearly functional in nature (subordinators, coordinators, particles), and some grammaticalizing lexical items (verbal, nominal adverbial). Examples are shown below.

231. Samoan

'o lo'u tama na fa'a-muli
PRS my father PST CAUS-last

se'ia⁶⁹ fo'i ane le malolo-ga a le fōma'i o tomasone
until return DIR the rest-NZR of the doctor of NAME

'My father stayed behind until doctor Tomasone returned from his holiday.' (Mosel & Hovdhaugen 1992, p 361)

232. Tetun

kawen ti'a, tür iha ne'e dei
marry already sit at this only

'After (we) are married, (we) must live here.' (Klinken 1999, p 236)

233. Kusaiean

tuku-n el orekma upac ah, el mas-ack
back-3.SG.POSS 3.SG work hard DET 3.SG sick-become

'After he worked hard, he became sick.' (Good 1989, p 173)

234. Coastal Konjo

a'-tangngaeng-i a-tinro i-Ali
INTR-be.in.middle.of-3.ABS STAT-sleep PROP-Ali

ri tuju-na kaluku / na-angambi' i-Baco'
at base-3.POSS coconut 3.ERG-TR-climb PROP-Baco'

'While Ali slept under a coconut tree, Baco' climbed it.' (Timothy Friberg, p.c. 2004)

⁶⁹ Mosel & Hovdhaugen (1992) describe se'ia as a TMA particle. It occurs in paradigmatic opposition to other TMA-particles but is only used in clause combining constructions indicating terminal boundary.
As illustrated by the examples, some of the nominal and verbal relation markers retain their productive morphology when they are used as relation markers in clause combining. Some, however, show more or less pronounced signs of being reanalyzed as designated functional relation markers. Usually, various grammatical morphology is lost: first optionally, then obligatorily, and finally the resulting form is associated only with its relation marker function. In Palauan, for instance, the conditional relation marker used when strong desire is involved derives quite clearly from a verb.

235. Palauan

\begin{verbatim}
ulékum a sensei ěr kêmam a mo ěr a guam,
if PHR teacher to 1.PL.EX PHR go.PRES to PHR Guam

mę ng mo diak a klas
then 3.SG FUT be.not.PRES PHR class
\end{verbatim}

'If only our teacher would go to Guam, then we wouldn't have any class.' (Josephs 1975, p 389)

It appears to have once been a verb in the so-called hypothetical verb form (primarily used for irrealis states of affairs). The verb in conditional clauses normally occurs in this form, but the main verb of \textit{ulékum}-clauses fails to do so, indicating that \textit{ulékum} itself may have been the hypothetical form in this construction. The -l- (and its position in the word) seems to hint at a third person singular form in the hypothetical verb paradigm. However, as \textit{ulékum} in present-day Palauan occurs only in conditional clauses, and invariably appears in this form, it must be regarded as an indivisible unit – a designated conditional relation marker. Semantically explicit relation markers may also be bound morphemes, as in the Labuk Kadazan example below.

236. Labuk Kadazan

\begin{verbatim}
ko-rikoti siri valai ka iri, mangaino ponsuo om
COOC-arrive there house RS DEM do bathe and.then

mangaino kulovungai do nulou
do REF wrap cloth
\end{verbatim}

'When he returned at the house, they say, he bathed it and wrapped it in cloth.' (Hurlbut 1990, p 105)

Semantically implicit clause combining constructions also come in several subtypes in Austronesian languages. The clearest examples are of course
juxtaposed simplex clauses, lacking any relational morpheme whatsoever, as in the Araki example below (237).

237. Araki

\[
\begin{array}{llllll}
\text{nam} & \text{lito-vi-a} & / & \text{mo} & \text{le} & \text{mle} \\
1.\text{SG.REAL} & \text{spit-TR-3.SG} & 3.\text{REAL} & \text{again} & \text{go.back}
\end{array}
\]

'I insulted him, (so) he went away.' (François 2002, p 172)

The result interpretation of the second clause in relation to the first is made from the propositional content of the clauses alone. Clausal juxtaposition is also commonly used for co-occurrence and posteriority relations in Araki, with the appropriate inferential cues (François 2002). Another salient subtype of implicit constructions involves semantically neutral relation markers. In the examples below, the relation markers do not contribute to establishing the semantic relation between the clauses.

238. Buru

\[
\begin{array}{llllll}
\text{tu} & \text{kam} & \text{hai,} & \text{tu} & \text{kam} & \text{seka-h} & \text{pao} & \text{lopi-n} \\
\text{and} & 1.\text{PL.EX} & \text{follow} & 1.\text{PL.EX} & \text{pierce-3.SG} & \text{down} & \text{bed-GEN}
\end{array}
\]

'And we followed it [the pig], and [then] we stabbed it down in the [stream]bed.' (Grimes 1991, p 404)

239. Big Nambas

\[
\begin{array}{llllll}
\text{pa-v'ah} & \text{prapar} & \text{lera} & \text{ti} & \text{i-haahaa} \\
3.\text{SG.REAL-look.out} & \text{sow} & \text{that} & \text{COMP} & 3.\text{SG.REAL-bite}
\end{array}
\]

'Look out for that sow, because it bites!' (Fox 1979, p 45)

In the Buru example, a neutral coordinator is used as a sequential relation marker; it has several other functions in the language, including that of an additive coordinator (Grimes 1991). In the Big Nambas example, \text{ti} is a general complementizer that is used here to introduce a reason clause, but in addition may introduce clauses functioning as sentential objects to matrix verbs, and does not specifically have the meaning 'because' (Fox 1979). Compare the Buru and Big Nambas examples above with the two examples below.
240. Buru

\[\textit{du flehe bia lea-lea, tu du hapu une-t}\]

3.PL pound sago RD-day and 3.PL tie cuscus.snare-NZr

'They pounded sage every day, and they tied cuscus snares.'
(Grimes 1991, p 404)

241. Big Nambas

\[\textit{i-vr ti a-kə-v-m'akar pr}\]

3.SG.REAL-say COMP 3.PL.REAL-NEC-PL-work NEC

'He said that they had to work.' (Fox 1979, p 105)

Markers such as these, which are uncommitted to specific semantic relations in one of their most common uses, have also been regarded to be implicit relation markers when semantic relations can be inferred (as in examples (238) and (239) above). It is of course possible to say that Buru \textit{tu} ('and') and Big Nambas \textit{ti} (complementizer) are polysemous between the different types of constructions in which they occur, and thus, that they are explicit in examples (238) and (239), and implicit in examples (240) and (241). This would be in line with the view of polysemy adopted in chapter 5 where no distinction is made between cases of polysemy arising from different senses (inherent meaning) and those arising from different uses. However, when a relation marker has the ability to connect clauses in constructions in which a particular semantic relation cannot be said to be conveyed (e.g. relation markers of the \textit{and}-type), I find it justified to assume that the marker is relationally implicit even when it occurs in constructions in which a semantic relation is intended and easily interpretable, and that such semantic relations arise from contextual inferences alone. When, on the other hand, a relation marker can only occur in constructions in which semantic relations really are conveyed, but then only in constructions with different relational meaning (cf. purposive and resultative \textit{so that} in English), I find it justified to assume that, although context and/or propositional content may play a part in the interpretation of the semantic relation, it does so in combination with the basic sense of the relation marker. This is why I have regarded relation markers of the former type as non-polysemous and implicit in all their uses, while I have regarded relation markers of the latter type as polysemous (cf. chapter 5).\textsuperscript{70}

\textsuperscript{70} Note also that relation markers in implicit clause combining constructions were excluded from the polysemy analysis in chapter 5 for the very reason that I do not regard them as potentially polysemous.
One theoretical difficulty that must be acknowledged here is how to determine that a certain clause combining construction does not convey any semantic relation and that a relation marker connecting the clauses in such a construction is therefore implicit. Although it would be a bit of a stretch to assume that Big Nambas *ti* in example (241) has a basic meaning of reason, there are cases in which the line between implicit and explicit is rather thin. For example, the Tagalog gerund prefix *pag-* is often used with verbs to create nominalized adverbial co-occurrence clauses (242a), although it may also function as a general nominalizer (242b).

242. Tagalog

a. *pag-*dating naming doon
   NZR-arrive 1.PL.EX.GEN there
   in-iwan namin don ang bangka'
   PF.PFTV-abandon 1.PL.EX.GEN there TOP boat
   'When we arrived there, we abandoned the boat.'
   (Himmelman 2005b, p 373)

b. ang *pag-*bi-bili ng mangga
   TOP NZR-RD71-buy GEN mango
   'the selling of mango' (Llamzon 1976, p 88)

Situations similar to that exemplified by Tagalog in (242a) may serve as a way for non-relational morphemes to acquire relational meaning, and this is what appears to have happened in Labuk Kadazan with the prefix *pog-*, which is cognate with Tagalog *pag-*.

243. Labuk Kadazan

*pog-*tingaa-Ø-ku na[antang]an-ku runah-sakit ranau
   IMM-look.up-PF-1.SG.ACTR RF[see]-1.SG.ACTR house-sick NAME
   'As soon as I looked up I saw "Ranau Hospital".' (Hurlbut 1988, p 73)

---

71 Llamzon (1976) makes a distinction between reduplicated *pag-*prefixed verbs and non-reduplicated ones, and claims that the former result in nominal gerund forms, while the latter result in adverbial absolute forms (although in some cases, apparently, the gerund form lacks reduplication; cf. Llamzon 1976, p 107). Other authors, however, do not regard this distinction as quite so systematic, and Himmelmann (2005b), for instance, is content in saying that in some cases, reduplication is part of the process of using the gerund prefix *pag-* and makes no distinction between the nominal and adverbial uses of *pag-*forms in his examples. In any case, though interesting, this distinction, if real, is not immediately relevant for our present purposes.
Since Tagalog \textit{pag-} has non-relational uses other than marking co-occurrence relations in clause combining, constructions such as that in (242a) above have been classified as implicit relationally. The corresponding Labuk Kadazan construction, on the other hand, has been classified as relationally explicit, since \textit{pog-} seems to be restricted to a converb function that indicates immediate anteriority. This is not an obvious decision, since one may well argue that Labuk Kadazan \textit{pog-} is basically implicit and nominalizing, though limited in distribution to contexts in which it takes on the meaning of immediate anteriority. However, this is the method I have used to distinguish between implicit and explicit relation markers, and I have applied it to the data to the best of my ability. On the whole, I feel quite confident that the relation markers classified as relationally implicit (mostly additive coordinators, complementizers and nominalizers) really are significantly less explicit than those classified as relationally explicit.

6.2 Morpho-syntactic types

The present section will discuss the types emerging from the combination of the parameter values presented in the previous section. These types will provide the basis of the distribution analyses to be made in chapter 7. Distributions will be mapped out both across geographic areas and across semantic relations. In the latter case, we are interested in investigating whether there is a tightness correlation between morpho-syntactic type and semantic relation. A structural tightness scale will therefore be suggested and the featured types discussed in section 6.2.1 below. In section 6.2.2, an interesting structural feature prevalent in Austronesian clause combining will be presented. It involves the fronting of a subordinate structure that is then joined to the rest of the clause by a coordinator. I have called the phenomenon 'asymmetric coordination', and since it cuts across several of the types, it will be discussed separately.

6.2.1 The tightness scale

The combination of parameter values described above produces eight logically possible types, as listed below:

- deviating, coordinate, implicit
- deviating, coordinate, explicit
- deviating, subordinate, implicit
- deviating, subordinate, explicit
non-deviating, coordinate, implicit
non-deviating, coordinate, explicit
non-deviating, subordinate, implicit
non-deviating, subordinate, explicit

These form the basis of the structural tightness scale that will be presented here. Structural tightness, as thought of in this study, manifests itself primarily by strong interclausal dependencies (for example, that one clause has a structure that cannot occur on its own, or that one clause depends on the other for the interpretation of grammatical features and/or the identity of actants), and a compact morpho-syntax as compared to a corresponding simplex clause (fewer morphemes/lexemes and or more bound morphology as a result of being part of a clause combining construction).

When arranging the types into a tightness scale, I chose to group all the deviating types together in one group. The reason for this is that constructions vary quite a bit with regard to the number and types of deviating features they display, so that some deviating coordinate constructions, for instance, appear tighter than some deviating subordinate constructions, and vice versa for other constructions. Therefore, it is very difficult to consistently order the deviating types in terms of structural tightness with regard to the parameters of syntactic relation (coordinate vs. subordinate) and explicitness of semantic relation (explicit vs. implicit). The types featuring non-deviating morpho-syntax, on the other hand, comprise constructions for which the structural potential is the same in both of the clauses involved. The parameters of syntactic relation and explicitness of the semantic relation thus become more important as differentiating factors, while, at the same time, delineate groups that are more homogenous internally.

Collapsing all the morpho-syntactically deviating constructions into one type leaves us with five types of constructions. The types have been ordered structurally from tight to loose along a tightness scale, which is presented in Table 22 below.

---

72 One could, of course, consider using other parameters to better capture internal tightness differences among the deviating constructions. But I have decided against this, because adding parameters would multiply the number of resulting types, which may be even more difficult to arrange consistently along a tightness scale, and because it might also in some cases be impossible to determine parameter values from the available sources if the parameters are very numerous.
Table 22. Tightness scale (from tight to loose)

<table>
<thead>
<tr>
<th>TYPE</th>
<th>PARAMETER VALUES</th>
<th>REPRESENTATIVE CONSTRUCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>deviating</td>
<td>echo-subject constructions, nominalized clauses, converb forms</td>
</tr>
<tr>
<td>2</td>
<td>intact, subordinate, implicit</td>
<td>implicit adverbial (&quot;complement&quot;) clauses</td>
</tr>
<tr>
<td>3</td>
<td>intact, subordinate, explicit</td>
<td>explicit adverbial clauses</td>
</tr>
<tr>
<td>4</td>
<td>intact, coordinate, explicit</td>
<td>explicit coordination / coherence adverbs</td>
</tr>
<tr>
<td>5</td>
<td>intact, coordinate, implicit</td>
<td>implicit coordination</td>
</tr>
</tbody>
</table>

The motivation for the ranking of the types will be presented in subsection 6.2.1.6, but before that, the types will be presented one by one, and some issues and particulars pertaining to each will be discussed.

6.2.1.1 Type 1: Deviating morpho-syntact

As mentioned above, the structural tightness of the deviating construction types is somewhat difficult to correlate with the parameters of syntactic relation and explicitness. The difficulties are illustrated here by a number of examples. We may set out by considering the Erromangan construction below, which is deviating, implicit and coordinate.

244. Erromangan

\[
\begin{align*}
\text{menuc} & \quad y-ocep & \quad m-tasi & \quad ra & \quad ndogo-n & \quad nei \\
\text{bird} & \quad 3.SG.DISTPST-fly & \quad SS.SG-alight & \quad \text{LOC} & \quad \text{branch-CONST} & \quad \text{tree}
\end{align*}
\]

'The bird flew and then alighted on the branch of the tree.'

(Crowley 1998, p 247)

The second clause in the construction features a same-subject marker, which is ungrammatical in simplex clauses. The verb cannot be marked with a normal person-TMA subject marker (as the verb in the first clause) as long as the identity of the subject referent is the same as that in the first clause. If it is different, however, the subject marker has to be of the normal person-TMA kind. This is a switch reference system reminiscent of many clause-chaining constructions in New Guinea (although a mirror image of them
since they have the same-subject markers on the non-final verb in the chain, owing to their SOV basic word order). Switch reference clause chaining constructions are normally considered to be very tightly connected, which also seems to be the case in Erromangan for the same-subject constructions of the type displayed above. The second clause has a compact appearance. The form of the verb is severely restricted, and an NP subject is not allowed. The second clause verb also depends on the first clause for the interpretation of some of its features, i.e. person and TMA categories. However, the second clause in the Erromangan construction cannot be considered to be embedded, as confirmed by various subordination tests. The same-subject marker, often called an echo-subject marker in the South Vanuatu languages, etymologically stems from a coordinating conjunction meaning 'and' (Lynch 1978).

Now, let us have a look at a similar construction, which is, however, relationally explicit. While the sequential reading of the Erromangan example is purely inferential, Lenakel, another South Vanuatu language, has echo-subject constructions that can be used with a specific sequential marker, to single out the semantic relation, as in the example below.

245. Lenakel

\[
\begin{array}{ll}
k\text{-}im\text{-}a\text{-}ini & \text{petimw} \\
3\text{.NSG-PST-PL-say} & \text{all} \\
m\text{-}ep\text{-}a\text{-}lis & \text{io} \\
\text{SS-SEQ-PL-take} & 1\text{.SG} \\
\end{array}
\]

'They told me all their names, and then they took me away.'

(Lynch 1978, p 50)

The only feature that sets the Lenakel construction structurally apart from the Erromangan construction above is the presence of the sequential TMA-marker \textit{ep-}, explicitly indicating the state of affairs of the second clause to be posterior in time to that of the first. So, the explicitness parameter would be splitting two constructions very similar in structural tightness if enforced here.

However, explicit echo-subject constructions also include examples in which verbs have grammaticalized into relation markers. They are less tightly knit since they feature a relation marker separating the clauses, but they still show the same parameter values as, for instance, the Lenakel con-

---

73 The term 'echo subject marker' (e.g. Lynch 1978, 1983, 2001; Crowley 1998, 2002a) derives, of course, from the fact that it occurs in the second clause, echoing the properties of the first clause subject, unlike the same-subject markers in the New Guinean languages, which only occur in non-final clauses.
struction (deviating, coordinate, explicit). An example is provided from Erromangan.

246. Erromangan

\textit{kaml-omonki}
\begin{align*}
1.\text{PL.EX.DISTPST-drink}
\end{align*}

\textit{m-ompi ml-etu-taru-hai armai}
\begin{align*}
\text{SS.SG-make SS.PL-NEG-think-NEG.EMPH properly}
\end{align*}

'We drank, so we didn't think properly at all.' (Crowley 1998, p 259)

Note that the echo subject marker on the verb-gone-relation marker is locked to its singular form here and does not copy the subject properties over from the preceding clause. The echo subject marker on the verb of the result clause, however, is used productively (for instance, it distinguishes between singular and plural forms) and copies the subject properties over from the clause preceding \textit{mompi}. This indicates that \textit{mompi} is not seen as a verb intervening between the clauses but rather as a relation marker. In this case, the explicitness parameter would lump together the constructions of examples (245) and (246) into one type while separating the constructions of examples (244) and (245), although the latter two are more similar to each other than the former two.

Turning now to deviating but subordinate constructions, these also seem to range from very tight to somewhat less tight in the sample languages. On the one hand, we have nominalized clauses and converb clauses (with or without an accompanying explicit relation marker), which are certainly both dependent and normally relatively compact in their morpho-syntax. An example of a nominalized clause is taken from Yakan below.

247. Yakan

\textit{ka-matey-ne-in / mag-iskul ne anak-ne-in}
\begin{align*}
\text{NZR-die-3.SG GEN-def INTR-school TMA child-3.SG GEN-def}
\end{align*}

'At the time of her death, her child was in school.'
(Brainard & Behrens 2002, p 51)

In the nominalized clause, the actant is represented by a bound possessive pronominal marker. In non-nominalized Yakan clauses, however, there are no bound argument markers. The state of affairs in the first clause of the example is represented in a more compact way than a corresponding declarative clause would be represented. In some Austronesian languages, however, we have deviating subordinate constructions that are less compact than the
Yakan nominalization. An example is Tagalog clauses with so-called basic form verbs (sometimes also called infinitive verbs).

248. Tagalog

\[
\{\text{oras na}\} \ \text{ma-kita} \ \text{ko} \ \text{siya} \\
\text{as.soon.as PF-see 1.SG.GEN 3.SG.TOP}
\]

\[
i-ta-tapon \ \text{ko} \ \text{ito} \ \text{sa mukha niya} \\
\text{PF-RD:FUT-throw 1.SG.GEN this.TOP to face 3.SG.GEN}
\]

'The minute I see him, I'll throw this in his face.'

(Schachter & Otanes 1972, p 471)

The first clause is subordinate to the second clause, and while it is morphosyntactically deviating, the only feature formally distinguishing the sequence \text{makita ko siya} from a corresponding simplex clause is the absence of a TMA category. Focus categories are indicated productively on basic form verbs, and there is nothing peculiar about the form of the arguments. So, if the deviating constructions involving subordination were considered as being tighter than those involving coordination, the Tagalog construction in (248) would be ranked as being tighter than the Lenakel echo-subject construction in (245). If, on the other hand, the deviating constructions involving coordination were considered as being tighter than those involving subordination, the Yakan construction in (247) would be ranked as being tighter than Erromangan construction in (246). Neither of these procedures, however, would produce satisfactory rankings. This shows some of the difficulties involved in ranking the deviating types for structural tightness with respect to the syntactic relation and explicitness parameters. Consequently, I have assigned all of these constructions to one type.

6.2.1.2 Type 2: Non-deviating, subordinate, implicit

Typical type 2 constructions – non-deviating, syntactically subordinate and relationally implicit – can be illustrated in Indonesian with constructions featuring the relation marker \text{agar}, which can introduce both complement clauses and purpose clauses depending on the semantics of the superordinate clause verb (Sneddon 1996). Compare the two examples below.
249. Indonesian

a. *kami ber-angkat pagi-pagi*
   1.PL.EX ACTV-depart morning-morning
   
   *agar kami tidak ter-lambat*
   COMP 1.PL.EX NEG STAT-be.late
   
   'We left early in the morning so that we wouldn't be late.'
   (Sneddon 1996, p 344)

b. *dia ber-citacita agar anak-nya*
   3.SG ACTV-desire COMP offspring-3.SG.POSS
   
   *di-lamar oleh orang kaya saja*
   PASS-propose by person wealthy just
   
   'He desires that his daughter be proposed to by no one but a rich man.'
   (Sneddon 1996, p 297)

The English complementizer *that* can of course be used in a similar way for both complement and purpose clauses, although purpose clauses made up this way seem to a large extent to have gone out of use, sounding old-fashioned in the current day language. Austronesian type 2 constructions may also consist of juxtaposed clauses, one of which can be identified as subordinate to the other.

250. Samoan

   *sa toli i lalo e pula le niū*
   PST fetch to down ERG NAME the coconut
   
   *e fe-inu ai Ø tama*
   GNR PL-drink ANA PL boy
   
   'Pula brought down the coconut in order for the boys to drink.' (AJ)

A CSC test (see chapter 2) suggests that the second clause is subordinate to the first, although structurally, its clausal pattern is intact and it could well be used as a simplex clause.74

74 The particle *ai* is anaphoric and refers back to an earlier context. The particle is not confined to clause combining.
6.2.1.3 Type 3: Non-deviating, subordinate, explicit

Type 3 constructions – non-deviating, syntactically subordinate, and relationally explicit – are relatively diverse but primarily involve adverbial clauses of various sorts, with an explicit marker for the relevant semantic relation. Consider the three examples below.

251. Big Nambas

>naran i-le-i i-dedrən
when 3.SG.REAL-see-3.SG 3.SG.REAL-be.afraid

'When he saw him, he was afraid.' (Fox 1979, p 45)

252. Samoan

>se'iloga e sau le tamāloa fa'atoā moe lona to'alua
unless GNR come the man just sleep his wife

'Unless the man comes (back), his wife will just go to sleep.' (AJ)

253. Kusaiean

>tuku-n el orekma upac ah, el mas-ack
back-POSS 3.SG work hard DET 3.SG sick-become

'After he worked hard, he became sick.' (Good 1989, p 173)

The two last examples quite clearly illustrate two different grammaticalization paths for adverbial subordinators in Austronesian languages, those coming from verbal expressions and those coming from nominal expressions. The Kusaiean anteriority relation marker tukun is derived from the noun tohkoh, 'back', suffixed by -n (the third person singular construct state suffix), which indexes a certain set of nominal possessors (Lee 1975, pp 103-4). The word tukun can also be used lexically but in its use as an anteriority relation marker, it is clearly generalized in meaning – semantically bleached (Sweetser 1988) – which suggests it has undergone a grammaticalization process. Another nominal expression commonly used as a relation marker is the word for 'time' in denoting co-occurrence relations. Note the examples below.

75 The suffixation causes some phonological changes to the stem so that the resulting form is tukun.
254. Labuk Kadazan

\[ \text{tontok do ko-pon-giup} \]
\[ \text{time} \quad \text{COMP} \quad \text{COOC-AF.INTEN.DRAM-drink} \]

\[ \text{tokou do kinomol, nga siri iri rogon} \]
\[ 1.{\text{PL.IN.TOP}} \quad \text{NTOP} \quad \text{kassava.beer then there that demon} \]

'…when we drink the cassava beer, then the demon is there'
(Hurlbut 1990, p 109)

255. Indonesian

\[ waktu 76 \text{ amir be-kerja di kebun} \]
\[ \text{time} \quad \text{NAME} \quad \text{ACTV-work at garden} \]

\[ \text{siti be-kerja di rumah} \]
\[ \text{NAME} \quad \text{ACTV-work at house} \]

'While Amir worked in the garden, Siti worked at home.'
(Sterner et al 1976, p 96)

256. Acehnese

\[ \text{ayah matê bak masa nyang lôn duek di bireuen} \]
\[ \text{father die at time REL 1.HON dwell in Bireuen} \]

'My father died at the time I was living in Bireuen.'
(Durie 1985, p 238)

Among the sample languages, the word for 'time' being used as a co-occurrence relation marker has been attested especially in the western Austronesian area. Other recent studies seem to confirm that it is common also among the Oceanic Austronesian languages (e.g. Bril (2010b) for Sonei in West Papua, and François (2010) for Hiw and Lo-Toga in Vanuatu). Example (256) shows that type 3 construction relation markers can be fairly elaborate. The more complex relation markers of this group tend to show signs of grammaticalization into more compact units. Samoan \text{se'iloga}, 'unless', from example (252) above, for instance, derives from \text{se'i}, an irrealis TMA particle, and \text{iloga}, 'be obvious'. The two morphemes have merged, and as the TMA particle can no longer change in clause combining, it indicates that it is

\[ ^{76} \text{The word waktu, 'time', is a loan from Arabic found in Austronesian languages spoken in islamic parts of Southeast Asia. Among the sample languages, it has been attested as a co-occurrence relation marker in Indonesian, Ma'anyan and Muna, and as a noun in Buru.} \]
perceived of as one unit – an adverbial subordinator. Used lexically, however, *iloqa* may freely take different TMA markers.

Some relation markers drawn from verbal expression have grammaticalized even further, leaving their lexical origin obscure. The Palauan conditional relation marker *lsêkum* is an example.

257. Palauan

\[tē mo ɢr a che\]
3.PL go.PRES to PHR fishing

\[a \text{ *lsêkum } ng \text{ ungil } a \text{ chei}\]
PHR if 3.SG be.good PHR tide

'They'll go fishing if the tide is good.' (Josephs 1975, p 387)

The *l-* of *lsêkum* looks suspiciously like the third person pronominal prefix of the hypothetical set. These morphemes are prefixed to Palauan verbs denoting irrealis states of affairs. Further, the phrase marker *a* is normally used with lexical phrases not beginning with a functional word (such as a subordinator). Its use with *lsêkum* indicates that the word used to have a more concrete lexical meaning. However, no other person markers may substitute for (once) third person *l-*, and the form *-*sêkum does not have an independent use (it occurs only as *lsêkum* and only in conditional clauses). Thus, this relation marker apparently constitutes a frozen form and currently functions exclusively as a conditional subordinator.

Other relation markers seem to be in a flux between keeping and losing inflections in their relation marker function. The Muna verb stem *rato*, 'arrive', can be used as a co-occurrence adverbial subordinator, either with third person pronominal inflections or without any inflections (Berg 1989). Below, its inflected verbal use, its inflected relation marker use, and its uninflected relation marker use are illustrated, respectively.

258. Muna

a. \[o \text{ kadadi-hi no-rato-mo}\]
ART aminal-PL 3.SG.REAL-arrive-PFTV

'The animals have arrived.' (Berg 1989, p 48)

---

77 In Muna, plural animals as subjects can either invoke a singular pronominal prefix on the verb (as here) or a plural pronominal prefix (Berg 1989), being located between humans and inanimates on the animacy hierarchy.
b. no-rato sabhangka-hi-no nagha
   3.SG.REAL-arrive friend-PL-3.SG.POSS DEM

do-wule do-po-kalalambu...
   3.PL.REAL-tired 3.PL.REAL-PLAY-game

'When her friends were tired of playing games…'
(Berg 1989, p 250)

c. rato no-suli ne-tulatula-mo bhe waiwai
   when 3.SG.REAL-return 3.SG-chat-PFTV with NAME

'When he came home, he chatted with Wai-Wai.'
(Berg 1989, p 250)

A verbal expression that often grammaticalizes into a relation marker in
Austronesian languages is that for 'finish'. The interesting thing here is that it
leads to relation markers for both posteriority relations (when functioning as
an adverb or a clause medial coordinator) and anteriority relations (when
functioning as a subordinator, taking the following clause as its comple-
ment). Only in the latter case, the constructions are classified as type 3 con-
structions, since the other do not involve subordination. Compare the exam-
pies below.

259. Tetun

nia n-anasa n-odi tatidin-án;
   3.SG 3.SG-laugh 3.SG-use jump-squat

hotu, nia-kan kidu-n tama bá rai
   finish 3.SG-POSS bottom-GEN enter go earth

'He laughed, jumping up and down, then his bottom went into the
ground.' (Klinken 1999, p 327)

260. Ma'anyan

luput m-asiq, parey na-tuyuk hawuwang panduk
   finish TR-harvest paddy PASS-heap inside hut

'Aafter we had done the harvesting, the paddy was put in the hut.'
(Gudai 1985, p 276)

In both these examples, the relation marker is a verb meaning 'finish'. In the
first example, it is analyzed as part of the second clause (non-subordinate),
marking the state of affairs as posterior, and in the second example, it is clearly part of the first clause (subordinate), marking the state of affairs as anterior.\textsuperscript{78}

\subsection*{6.2.1.4 Type 4: Non-deviating, coordinate, explicit}

The most common relation markers in type 4 constructions are adverbs (or adverbial sequences), in one or both of the clauses. The clauses are either juxtaposed or joined by means of semantically neutral coordinators. The Tetun example (259) above provides an illustration of the former, while the Muna example (261) below is an illustration of the latter.

261. Muna

\begin{verbatim}
amári  ísi  tágo  i-raʔe-be
sun    yet  NEG  3.SG.REAL-go.up-and
\end{verbatim}

\begin{verbatim}
alúlu  anúa  i-péreʔ-i
messenger village 3.SG.REAL-leave-3.SG
\end{verbatim}

'BBefore the sun rose, the messenger left the village.'

(Lichtenberk 1983, p 519)

Another thing illustrated by the Muna example above is something fairly common among the construction within type 4, i.e. to express posteriority by means of the equivalents of the sequence 'not yet' (or some similar negative expression). Conversely, it is perhaps even more common within this group to express anteriority by means of the equivalents of the word 'already', as illustrated by example (262) below.

262. Tetun

\begin{verbatim}
fó  sia  r-emu  ti’a,  /  nia  n-ák  á...
give  3.PL  3.PL-drink  already  3.SG  3.SG-say  HES
\end{verbatim}

'AAfter (I) had given to them (and) they had drunk, he said...'

(Klinken 1999, p 305)

\textsuperscript{78} Actually, the Tetun verb \textit{hotu}, 'finish', can also be used as an anteriority relation marker taking the following clause as a complement, often in head-tail linkage style (e.g. 'They sang their song. (Having) finished singing, they ate their food.') (Klinken 1999, pp 304-5). This use is thus identical to the use of Ma'ananyan \textit{luput} ('finish'). Consequently, Tetun \textit{hotu} has been classified both as an anteriority and as a posteriority relation marker in the data, and the constructions in which they are used as belonging to type 3 and type 4, respectively.
Another quite large type 4 group makes use of semantically explicit coordinators.

263. Hoava

\[
\begin{array}{llllllll}
\text{mate} & \text{paho} & \text{ria} & \text{sogi} & \text{hoqa} & \text{taloa} & \text{ria} \\
\text{be.dead} & \text{INTEN} & 3.\text{PL} & \text{so} & \text{flee} & \text{leave} & 3.\text{PL}
\end{array}
\]

'They were all dead so they (the others) ran away.'
(Davis 2003, p 263)

Sometimes the difference between an explicit coordinator and a coherence adverb is difficult to make, but it need not concern us here. A small number of type 4 constructions also indicate the semantic relation by other means, such as TMA markers (bound or free) or other grammatical elements.

6.2.1.5 Type 5: Non-deviating, coordinate, implicit

Type 5 constructions, finally, are either connected by a neutral coordinator or simply juxtaposed.

264. Yabem

\[
\begin{array}{llllll}
\text{gê-bêc} & \text{mu} & \text{gê-buc} & \text{ŋaténa}, \\
3.\text{SG.REAL-be.dusk} & \text{wind} & 3.\text{SG.REAL-pull} & \text{much}
\end{array}
\]

\[
\begin{array}{llllll}
\text{ma} & \text{kom} & \text{gê-jac} & \text{gê-mêŋ} \\
\text{and} & \text{rain} & 3.\text{SG.REAL-hit} & 3.\text{SG.REAL-come}
\end{array}
\]

'Last night the wind blew violently and then the rain came.'
(Dempwolff 2005 [1939], p 104)

265. Mekeo

\[
\begin{array}{llllll}
\text{matsi} & \text{iu-ŋa} & \text{fo-peni-au,} & \text{fama-ani-a} \\
\text{wallaby} & \text{tail-3.SG.POSS} & \text{2.SG.OBLG-give-1.SG} & 1.\text{PL.OBLG-eat-3.SG}
\end{array}
\]

'Give me the tail of the wallaby, so as we may eat it.'
(Jones 1998, p 477)

\[79\] This relation marker possibly consists of a loan from English via Bislama of the result relation marker \textit{so}, which has merged with Hoava \textit{gi}, 'and', to form a result coordinator (Davis 2003).
Relational meanings in these cases are entirely, or at least primarily, derived from context and semantic contents. Conventional usage may sometimes delineate a preferred set of relational meanings.

6.2.1.6 The ranking of types

The ranking of the different types from type 1 (tightest) to type 5 (loosest) will be motivated in this section.

Type 1 constructions are considered to occupy the tightest end of the scale because, whatever the other properties, deviating patterns in the formal makeup of one of the clauses (or clause-like structures) related to the fact that it is a part of a clause combining construction by necessity entail inter-clausal dependencies. If a structure that does not occur in simplex clauses occurs in one of the clauses in clause combining, the clause in which it occurs becomes dependent on the other clause in the construction for its existence. From this follows that all constructions involving deviating morphosyntax are structurally tighter than all constructions involving non-deviating constructions.

The presence of subordination in types 2 and 3 entail greater dependency for these types than for types 4 and 5. Subordinate clauses depend on their matrix clause for assertive power, i.e. the matrix clause contributes with the assertive power for the construction, while the subordinate clause has none. Coordinate clauses, on the other hand, both have assertive power. This gives us types 2 and 3 as structurally tighter than types 4 and 5.

With regard to the explicitness parameter, I have found it to coincide with structural tightness in contrasting ways in non-deviating subordinate clauses and non-deviating coordinate clauses. That is, implicitly related coordinate clauses are more loosely tied together than their explicitly related counterparts, while implicitly related subordinate clauses are more tightly tied to their matrix clause than their explicitly related counterparts. Consider the following Austronesian examples.

---

80 Note, however, that there are constructions containing clauses with a subordinate format that nonetheless have their own assertive power (e.g. disjunct adverbial clauses, assertive complement clauses and restrictive relative clauses), as indicated by the fact that they fail one or more of the subordination tests (see section 2.1.2). The extent to which such constructions occur in the Austronesian languages has not been possible to test in the present study but remains an interesting topic for future research.
6. MORPHO-SYNTAX

266. Indonesian – SUBORDINATE, IMPLICIT:

\[
\begin{align*}
kami & \quad \text{ber-angkat} & \quad \text{pagi-pagi} \\
1.\text{PL}.\text{EX} & \quad \text{ACTV-depart} & \quad \text{morning-morning} \\
\end{align*}
\]

\[
\begin{align*}
\text{agar} & \quad \text{kami} & \quad \text{tidak} & \quad \text{ter-lambat} \\
\text{COMP} & \quad 1.\text{PL}.\text{EX} & \quad \text{NEG} & \quad \text{STAT-be.late} \\
\end{align*}
\]

'We left early in the morning (so) that we wouldn't be late.'
(Sneddon 1996, p. 344)

267. Coastal Konjo – SUBORDINATE, EXPLICIT:

\[
\begin{align*}
coba & \quad \text{ku-hokka-i} & \quad \text{kondo-a} & \quad \text{sikarie'} \\
\text{if} & \quad 1.\text{ERG-chase.off-3.ABS} & \quad \text{egret-DET} & \quad \text{yesterday} \\
a'-'ribba'-i & \quad \text{INTR-fly-3.ABS} \\
\end{align*}
\]

'If I had chased off the egret yesterday, it would have flown.'
(Timothy Friberg, p.c. 2004)

268. Hoava – COORDINATE, EXPLICIT:

\[
\begin{align*}
[\ldots] & \quad \text{ada} & \quad \text{ria} & \quad \text{tige} & \quad \text{buki-a} & \quad \text{ria} & \quad \text{sa} & \quad \text{buki} \\
\text{wake} & \quad 3.\text{PL} & \quad \text{then} & \quad \text{blow-3.SG} & \quad 3.\text{PL} & \quad \text{ART.SG} & \quad \text{conchshell} \\
\end{align*}
\]

'...they woke up, then they blew the conchshell.' (Davis 2003, p 262)

269. Araki – COORDINATE, IMPLICIT:

\[
\begin{align*}
\text{nam} & \quad \text{lito-vi-a} & \quad \text{mo} & \quad \text{le} & \quad \text{mle} \\
1.\text{SG}.\text{REAL} & \quad \text{spit-TR-3.SG} & \quad 3.\text{REAL} & \quad \text{again} & \quad \text{go.back} \\
\end{align*}
\]

'I insulted him, (so) he went away.' (François 2002, p 172)

That the types are lined up like this is no coincidence. Implicitly related coordinate constructions (juxtaposed clauses and those related by means of a neutral coordinator, 'and') are simply coupled in discourse and connected in a very uncommitted way. Their relation is primarily of a discourse-pragmatic nature. Relational meaning is derived from context. Implicitly related subordinate constructions, on the other hand, are clearly integrated in the syntactic structure, as the subordinate clause assumes a distinct syntactic function in its matrix clause; its relation to the matrix clause is primarily of a syntactic nature. Relational meaning is normally derived from the construction itself.
In between, we have the explicitly related constructions, types 3 and 4, which I consider to be tighter than type 5 constructions (implicit coordination) because they have a designated relational meaning and are less susceptible to pragmatic inferences, and looser than type 2 constructions (implicit subordination) because they are not dependent on the construction format in the same sense. Underlying this reasoning is the assumption that grammatical dependencies, and thus tightness, increase in relation to how much of the meaning is interpreted via (i) pragmatic cues, (ii) semantic items, and (iii) syntactic structures. Note, however, that this assumption is independent of specific semantic relations. That is, whether the semantic relation is understood from a particular syntactic construction, from semantically explicit relation markers or from the discourse-pragmatic context does not mean that the semantic relations necessarily is of a specific kind. At least, we have no a priori reasons to assume so.

Taken together, we find ourselves with the structural ordering of the types as given in Table 22; from tight to loose: type 1, type 2, type 3, type 4, type 5.

### 6.2.2 Asymmetric coordination

Before discussing the distribution of the above types across the semantic relations outlined in chapter 4, and across geographic areas, we will have a look at the phenomenon briefly mentioned above that I have called asymmetric coordination. It cuts across several types and seems to be fairly common among Austronesian languages.

There is a small risk of confusion in using the term 'asymmetric coordination', since it has previously been used differently from how I use the term in this study. It has been used, for instance, to refer to certain word order asymmetries, primarily in German (e.g. Höhle 1990; Frank 2002), in which sometimes only the first of two coordinated consecutive subordinate clauses follow the general word order rule for German, whereby the finite verb occurs last in a subordinate clause.

270. German

> wenn Du in ein Kaufhaus kommst
> if you into a shop come
> und (Du) hast kein Geld, kannst Du nichts kaufen.
> and you have no money can you nothing buy

'If you enter a shop and (you) don't have any money, you can't buy anything.' (Frank 2002, p 176)
Another example is when asymmetric coordination is used for various kinds of coordinate structures conveying semantically asymmetric relations (e.g. Schmerling 1975; Na & Huck 1992), as in the following examples.

271. a. I left the door open and the cat got in.
    b. I went to the store and bought some whiskey.

This is not, however, what I intend by the term 'asymmetric coordination'. In spite of the fact that the term has been used for other phenomena, I have decided to stick with 'asymmetric coordination' for the Austronesian constructions to be described, partly because there seems to be no standard usage of the term – it has previously been used for several different phenomena – but mostly because I think it precisely captures the properties of the phenomenon at hand.

The way I use the term, then, is to refer to a pattern in which an initial "subordinate clause" (normally adverbial in nature) is connected to the following "main clause" by means of a coordinator. An English literal equivalent would be to turn the sentence "I'll hit him if he is mean" into "If he is mean, and I'll hit him". The coordinator seems to function as a dividing line between topical information in the initial clause and the predication made in the following clause. It is not quite so easy to determine whether this is subordination or coordination, and if it is subordination, which clause is subordinate to which – especially, as we will see examples of in what follows, when the coordinator is homophonous with a complementizer. I have not been able to find references to the construction in any general linguistic literature, although Bril (2010b) identifies coordinators as markers commonly developing into topic and focus markers in Austronesian languages, resulting in the constructions I refer to as asymmetric coordination.

6.2.2.1 Austronesian examples

Let us have a look at some Austronesian examples. Consider first the sentence below from Labuk Kadazan (Borneo), showing the unmarked occurrence of a postposed reason clause, which cannot be connected to the previous clause by a coordinating morpheme.

272. Labuk Kadazan

\[
pang-\text{akan-no} \ d\text{i} \ \text{iolo}, \ [\ sabap \ do \ au \ iolo
\]
\[
\text{VOL.AF-eat-?} \ \text{EMPH} \ 3.\text{PL.TOP} \ \text{because} \ \text{COMP} \ \text{not} \ 3.\text{PL.TOP} \\
\[
\text{minang-arap} \ d\text{i} \ \text{nibilin} \ do \ \text{kinoroingan} \ ] \\
\text{AF.CMPL-trust} \ \text{GEN} \ \text{command} \ \text{GEN.SPEC} \ \text{god}
\]

'They ate it because they did not trust in the command of god.'

(Hope Hurlbut p.c. 2007)
The second clause is definable as subordinate, both language internally, on grammatical grounds by the use of the complementizer do, and cross-linguistically, in denoting non-asserted information. The latter can be checked by the CSC test (coordinate structure constraint) saying that extraction of a nominal is not allowed from either of two coordinate clauses (cf. section 6.1.2 and section 2.1.2 above).

273. Labuk Kadazan

\[\text{onu \ dii \ in-akan \ diolo \ sabap \ do \ au \ iolo}\]
what EMPH PF.CMPL-eat 3.PL.GEN because COMP not 3.PL.TOP

\[\text{minang-arap \ di \ nibilin \ do \ kinoroingan?}\]
AF.CMPL-trust GEN command GEN.SPEC god

'What did they eat because they did not trust in the command of god?'
(Hope Hurlbut p.c. 2007)

Since extraction of a question word with scope of the entire sentence does not result in ungrammaticality in the Labuk Kadazan example, we can be quite confident that subordination is really involved here. In example (274) below, however, the reason clause from example (272) is in initial position instead, in which case a coordinator, om ('and'), is optionally inserted between the clauses. This is a typical example of what I call 'asymmetric coordination'.

274. Labuk Kadazan

\[\text{[ sabap \ do \ au \ iolo \ minang-arap \ di \ nibilin}\]
because COMP not 3.PL.TOP AF.CMPL-trust GEN command

\[\text{do \ kinoroingan} \ ] (om) \ \text{pang-akan-no \ dii \ iolo}\]
GEN.SPEC god and VOL.AF-eat-? EMPH 3.PL.TOP

'Because they did not trust in the command of God, they ate it.'
(Hope Hurlbut p.c. 2007)

The Labuk Kadazan word om is normally used in plain coordination of constituents, at various levels of syntax as shown below.

275. Labuk Kadazan

\[\text{ison \ karaja \ om \ isoon \ do \ mongoi \ mandam \ do \ kawan}\]
one work and one COMP go visit GEN firend

'There is one time for work and one time for visiting friends.'
(Hurlbut 1990, p 116)
Similar examples can be found in all parts of the Austronesian area. Note, for instance, the Karo Batak (Sumatra), Palauan (Micronesia), Samoan (Polynesia) and Cèmuhî (Melanesia) examples below. The (a) examples show an adverbial clause in final position, while in the (b) examples the same (or a similar) clause is fronted, generating asymmetric coordination. The (c) examples show regular coordination for comparison (for Palauan, I have provided some additional examples; see comments in connection with them).

276. Karo Batak

a. meriah kal ku-akap i jënda,
   happy very 1.SG-feel at here
   [ sebap ia per-jagar-jagar ]
   because 3.SG CAUS-jovial-RD
   'I feel very much at home here, because he is a jovial fellow.'
   (Woollams 1996, p 333)

b. [ perbahanken aku la nggit man ba-na ]
   because 1.SG NEG desire towards at-3.SG
   è maka ndelis ia
   and so hang.REFL 3.SG
   'Because I didn't love him, he hanged himself.'
   (Woollams 1996, p 333)

c. rempet ia guling è maka mis mate
   suddenly 3.SG fall.down and then directly die
   'Suddenly he fell down and died.' (Woollams 1996, p 335)

Palauan has two different coordinators: e, which has a sequential connotation to it, and mȩ, which is neutral (compare the (c) and (e) examples below). As evident from examples (b) and (d), both can be used in asymmetric coordination with different types of conditional clauses.

277. Palauan

a. tȩ mo ěr a che
   3.PL go to PHR fishing
   [ a lṣekum ng ungil a chei ]
   PHR if 3.SG good PHR tide
   'They will go fishing if the tide is good.' (Josephs 1975, p 389)
b.  [ a  lsękum  ng  ungil  a  chei ]
    PHR  if  3.SG good PHR tide

    e  tê  mo  ěr  a  che
    and  3.SG FUT be.not PHR class

    'If the tide is good, then they'll go fishing.' (Josephs 1975, p 388)

c.  ng  mirrael  a  droteo  e  ak  mlo  męchiuaiu
    3.SG leave.PST PHR Droteo and 1.SG go.PST sleep

    'Droteo left and I went to sleep.' (Josephs 1975, p 487)

d.  [ ulékum  a  sensei  ěr  kêmam  a  mo  ěr  a  guam ]
    if.only PHR teacher to 1.PL.EX PHR go to PHR Guam

    mé  ng  mo  diak  a  klas
    and  3.SG FUT be.not PHR class

    'If only our teacher would go to Guam, then we wouldn't have any
    class.' (Josephs 1975, p 389)

e.  a  bile-k  a  bęcheleliu  mé
    PHR shirt-1.SG PHR be.white and

    a  bibe-l  a  droteo  a  bękerekared
    PHR shirt-3.SG PHR Droteo PHR be.red

    'My shirt is white and Droteo's shirt is red.' (Josephs 1975, p 484)

Most languages use a neutral coordinator (or less often, one with sequential
meaning) in asymmetric coordination, but Samoan makes use of the adversa-
tive coordinator ' ae ('but') when a concessive or concessive conditional
clause is put in initial position – obviously to emphasize the contrast that is
present between the states of affairs (cf. also section 4.3.7).

278. Samoan

a.  e  lë  mafia  lava  ona  fo'i  atu  le  tama
    TMA NEG be.possible indeed COMP return there the boy

    [ tīgā  ona  tagi  mai  lava  iā  ila ]
    although COMP cry here indeed ABS NAME

    'The boy could not return, although Ila cried a lot.' (AJ)
b. \[ \text{tīgā ona tagi mai lava iā ila } \]'ae \\
\text{although} \text{ cry} \text{ here} \text{ indeed} \text{ ABS NAME but} \\
\text{e lē mafia lava ona fo'i atu le tama} \\
\text{TMA NEG be.possible indeed} \text{ COMP return there} \text{ the} \text{ boy} \\

'Although Ila cried a lot, the boy could not return.' (AJ)

c. \text{'}ua ala lo'\text{u tuagane 'ae olo'o moe lo'\text{u tama}} \\
\text{TMA awake} \text{ my} \text{ brother} \text{ but} \text{ TMA sleep} \text{ my} \text{ father} \\

'My brother was awake but my father was asleep.' (AJ)

Usually, the coordinating morpheme is optional between the initial clause and the following clause in Austronesian asymmetric coordination. However, in Cèmuhî (New Caledonia), the use of a coordinator is obligatory when a subordinate co-occurrence clause or conditional clause is used initially in clause combining (Lynch 2002).

279. Cèmuhî (Southern Oceanic)

a. \text{kō néjɛ ħɛmɛ nyē kō enī} \\
\text{real} \text{ tasty} \text{ when} \text{ 1.PL.IN real} \text{ eat} \\

'It is real tasty when one eats it.' (Lynch 2002, p 763)

b. \text{mēpīgō mwó tē ŭcāa mwo} \\
\text{if} \text{ 2.SG transitory real} \text{ wait} \text{ again} \\
\text{kā ē bō túiɛ} \\
\text{and} \text{ 3.SG INDEF.FUT arrive} \\

'If you just wait on, he will probably come.' (Lynch 2002, p 763)

c. \text{ni ū kā ni wāéō kā ni ūjā...} \\
\text{PL yam and} \text{ PL taro and} \text{ PL sugarcane} \\

'yams and taro and sugarcane...' (Lynch 2002, p 757)

Asymmetric coordination, then, is a recurring pattern throughout the Austronesian family. There are no apparent genealogical or geographic concentrations in its distribution within Austronesian. Of all the Austronesian languages I have checked for asymmetric coordination – the sample languages as well as a fair number of additional Austronesian languages – the
phenomenon is found in roughly half. It is attested in the following 17 sample languages (languages indicated by E are in the extended sample only):

<table>
<thead>
<tr>
<th>Language</th>
<th>Language</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amis</td>
<td>Coastal Konjo</td>
<td>Hoava (E)</td>
</tr>
<tr>
<td>Palauan</td>
<td>Labuk Kadazan</td>
<td>Longgu</td>
</tr>
<tr>
<td>Central Cagayan Agta</td>
<td>Karo Batak</td>
<td>Erromgan</td>
</tr>
<tr>
<td>Tagalog</td>
<td>Kambera</td>
<td>Samoan</td>
</tr>
<tr>
<td>Western Subanon</td>
<td>Buru</td>
<td>Hawaiian (E)</td>
</tr>
<tr>
<td>Tboli</td>
<td>Yabem</td>
<td></td>
</tr>
</tbody>
</table>

This corresponds to 40% of the 43 extended sample languages and to 50% of the 30 core sample languages. It does not seem far-fetched to assume that asymmetric coordination is even more widely distributed in Austronesia, since the phenomenon may well have been overlooked in some grammatical descriptions not attending to the details of fronted adverbial clauses.

### 6.2.2.2 Analysis

A way of understanding asymmetric coordination is in terms of discourse-pragmatic coherence. It is widely recognized that large constituents tend to occur late in the sentence for ease of processing. For this reason, adverbial clauses – being comparatively quite large constituents as a rule – tend to occur sentence finally. However, for various other reasons – syntactic, iconic and/or pragmatic – they sometimes occur sentence initially. In these cases, the need arises to create cohesion between a cumbersome initial constituent and the rest of the sentence. The larger the initial constituent, the greater the need for a cohesive device. Speaker can do this in different ways. In English, for instance, resumptive adverbs are often used with this pragmatic function (cf. 'if…then'), the entire adverbial clause being condensed, so to speak, into an adverb that is more easily related to the following sentence structurally. Note these examples (some perhaps rather colloquial in style, but others entirely integrated in idiomatic language use):

---

81 The use of certain cohesive devices may of course be conventionalized and spread to contexts in which the initial constituent is not necessarily very long, but have some other properties in common with initial adverbial clauses, for instance, in providing topical information. There are indications that such instances of asymmetric coordination (initial adverb + coordinator) may be the source of some items that have grammaticalized into relation markers in Austronesian languages (see Table 23 in the next section for possible examples). Whether or not such a grammaticalization process can actually be established, however, has not been investigated more closely.
If you give me the bag, then I'll tell you what happened.

When we reached the cabin, then we started to organize our group.

Because we didn't know what had happened, so we just kept on going.

When we had reached the end of the track, at that point, we were all exhausted.

In many Austronesian languages, however, it is the role of the coordinator to create cohesion under these circumstances. This illuminates a fundamental difference in syntactic and discourse-pragmatic structure between European and Austronesian languages in this respect. From a syntactic point of view, **resumptive adverbs** bring the initial clause into the structure of the following clause; the entire contents of the initial clause being represented by the resumptive adverb. **Asymmetric coordinators**, on the other hand, explicitly exclude the initial clause from the structure of the following one, forcing an independent clause interpretation on both clauses. From a pragmatic point of view, **resumptive adverbs** stress the **non-asserted** status of the initial adverbial clause, formally including both the adverbial clause and the resumptive adverb in the scope of the assertion made in the following main clause. Conversely, **asymmetric coordinators** render the initial clause an assertion of its own, formally excluding it from the scope of the assertion made in the following clause.

It may seem to come natural to choose a coordinator as a cohesive device if the relation marker of the fronted clause originates from a verb, or some other item that is often used predicatively. In such cases, the relation marker is easily rendered to have assertive power in itself. As discussed in chapters 4 and 5, a fair number of Austronesian relation markers have verbal/predicative origins. Two clear examples are the Samoan concessive conditional and the Tagalog anteriority constructions shown below.

Samoan

\[
\begin{align*}
\text{tusa} & \quad \text{lava} & \quad \text{fo'i} & \quad \text{pē} & \quad \text{lē} & \quad \text{lalelei} & \quad \text{ma} & \quad \text{aulelei}, \\
\text{be.same} & \quad \text{indeed} & \quad \text{also} & \quad \text{Q} & \quad \text{NEG} & \quad \text{beautiful} & \quad \text{and} & \quad \text{handsome} \\
\text{'ae} & \quad \text{'o} & \quad \text{le} & \quad \text{mea} & \quad \text{e} & \quad \text{aupito} & \quad \text{sili} \\
\text{but} & \quad \text{PRS} & \quad \text{the} & \quad \text{thing} & \quad \text{GNR} & \quad \text{be.most} & \quad \text{highest} \\
\text{ona} & \quad \text{mana'o-mia} & \quad \text{e} & \quad \text{le} & \quad \text{tagata}, & \quad \text{'o} & \quad \text{le} & \quad \text{soifua} & \quad \text{malolo-in} \\
\text{COMP} & \quad \text{want-AUG} & \quad \text{ERG} & \quad \text{the} & \quad \text{person} & \quad \text{PRS} & \quad \text{the} & \quad \text{health} & \quad \text{rest-AUG} \\
\end{align*}
\]

'Even if [someone] is beautiful and handsome, what he or she wants most is a healthy life.' (Mosel and Hovdhaugen 1992, p 664)

Alternatively: 'It's the same if [someone] is beautiful and handsome, but what he or she wants most is a healthy life.'
285. Tagalog

\[\text{pag-katapos} \ niya-ng \ \text{mag-talumpati},\]
NZR-finish 3.SG.GEN-LIG AF-hold.a.speech

\[\text{at} \ \text{saka} \ \text{k[um]anta} \ \text{kami} \ \text{ng} \ \text{ilang} \ \text{awit}\]
and then [AF.PFTV-sing] 1.PL.TOP GEN some song

'After he spoke, we sang a few songs.' (MJ)\textsuperscript{82}

Alternatively: 'He finished speaking, and then we sang a few songs.'

The tendency to use verbal relation markers in the initial clause in constructions such as the above may have led the way to using a coordinator as a cohesive device between the clauses. When the verbal relation marker is reinterpreted as a subordinator, a possible medial coordinator may sometimes be retained.

Both strategies – the use of resumptive adverbs and asymmetric coordinators – probably ease the burden of cognitive processing. In the case of resumptive adverbs, the initial clause is "zipped" down – to use a computer term – to a more manageable unit, and in the case of asymmetric coordinators, the initial clause is processed independently. In the latter case, it is easier to proceed to the next clause when the initial clause is left behind as an independent statement. To use further computer terminology, it could be said that speakers "empty cache memory" of any structural information.

6.2.2.3 Points of interest for further study

An interesting fact pertaining to asymmetrically coordinate constructions in some Austronesian languages is that the coordinator used is homonymous with a preposition or complementizer. This potentially opens up for the analysis of the fronted adverbial clause as the main clause and the following clause as its complement. Consider, for instance, the Hawaiian examples below with ā, used both as coordinator and as an allative preposition stressing distance.

\textsuperscript{82} Although my informant does not reject the example, she prefers a version without the medial coordinator \textit{at}, 'and', which indicates that \textit{pagkatapos} is generally perceived of as a subordinator in modern Tagalog. However, Ramos & Cena (1990, p 140) state that the pattern with a medial \textit{at} is allowed in Tagalog, although they do not provide an example.
286. Hawaiian

ASYMMETRIC COORDINATION

a. inā 'oe e 'ae ana e kau pū
   when 2.SG IPFV agree IPFV INTEN place together

   ku'u mau 'ope'ope me a'u
   1.SG.POSS PL bundle with 1.SG

   ā 83 laila holo pū kāua
   and then sail together 1.DU.IN

   'If you agree to place my bundles with me, then we'll sail together.' (Hawkins 1982, p 138)
   Alternatively: '(It is) when you agree to put my bundles with me (that) we (go) as far as sailing together.'

COORDINATOR

b. hānai punahele 'ia 'o kūapaka'a
   raise favorite PASS PROP NAME

   ā 84 ua a'o aku-la 'oia
   and PFTV learn away-there 3.SG

   i nā mele no keawe
   ACC PL chant for NAME

   'Kūapaka'a was raised as a favorite and he learnt the chants for Keawe.' (Hawkins 1982, p 136)

PREPOSITION

c. hele ā ke kuahiwi
   go far.to the mountain

   'going (all the way) to the mountain' (Elbert & Pukui 1979, p 136)

83 In the original example, Hawkins (1982) uses a (without a superposed macron) for the coordinator ā, but Elbert & Pukui (1979) as well as Pukui & Elbert (1986) clearly distinguish between the coordinator/preposition ā ('and', 'to', 'as far as'), on the one hand, and the possessive preposition a ('of') on the other.

84 See previous footnote.
A similar situation is found in Coastal Konjo, in which the relation marker \textit{na} functions as both as a complementizer and coordinator according to Friberg (1996).

287. Coastal Konjo

**ASYMMETRIC COORDINATION**

a. \textit{sangge-na} \textit{rie’ pulisi na kung-a’-cari’carita}  
until-3.POSS exist police and 1.PERI-INTR-tell.stories  

'Until the police arrived, I told stories.'  
(Timothy Friberg, p.c. 2004) 
Alternatively: '(It was) until the police arrived that I told stories.'

**COORDINATOR**

b. \textit{ku-pa-kanre-i na ku-pa-inung-i}  
1.ERG-CAUS-eat-3.ABS and 1.ERG-CAUS-drink-3.ABS  

'I fed him and I gave him a drink.' (Friberg 1996, ex 103)

**COMPLEMENTIZER**

c. \textit{angngura-i na nu-isse’i ang-kua ia-mi}  
why-3.ABS COMP 2.ERG-know-3ABS TR-say 3-COMP  

\textit{inni tau toa-ku}  
this person old-1.POSS  

'How is it that you know that these are my parents.'  
(Friberg 1996, ex 95)

However, while the Hawaiian (c) example above clearly displays \textit{ā} as a preposition taking a nominal complement, the Coastal Konjo (c) example does not necessarily identify \textit{na} as a complementizer. It could well be analyzed as a coordinator here, in which case (c) becomes another example of asymmetric coordination. If so, the use of a cohesive coordinator (\textit{na}) seems to have spread from instances of quite long initial constituents to shorter ones that have some other properties in common with preposed adverbial clauses. In fact, some Austronesian relation markers seem to have grammaticalized

\[85\] References to Friberg (1996) are given with example numbers rather than page numbers, as my source is a Word-version of the article published in Steinhauer (1996).
from constructions of asymmetric coordination with shorter initial constituents. At least three potential cases have been found among the sample languages, as listed in the table below.

Table 23.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>REL. MARKER MEANING</th>
<th>LITERAL MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaiian</td>
<td>aia (no) ā</td>
<td>when</td>
<td>aia there</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>no indeed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ā and</td>
</tr>
<tr>
<td>Tagalog</td>
<td>kung kaya't</td>
<td>so</td>
<td>kung if</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>kaya-at so-and</td>
</tr>
<tr>
<td>Central Cagayan Agta</td>
<td>awá</td>
<td>because</td>
<td>awe NEG and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(?)</td>
</tr>
</tbody>
</table>

Regular asymmetric coordination with an initial "adverbial clause" coordinated to the following clause is also attested in all three languages. An example of the use of the Tagalog result relation marker kung kaya't follows below.

288. Tagalog

'umulu
RD[FX]:AF.IPFW-rain

kung kaya't p[um]asok kami sa sine
if so-and enter[AF.PFTV] 1.PL.EX.TOP to cinema

'It was raining, so we went to the movies.'
(Schachter & Otanes 1972, p 543)

In order to determine how to best analyze asymmetric coordinate constructions with medial relation markers that are ambivalent between coordinator and preposition or complementizer, much more data should be compiled and systematic comparison made of the morpho-syntactic behavior and distribution of the ambivalent elements in the different languages. Unfortunately, that is outside the scope of the present study.

Interestingly, however, analyzing the "main clause" as a complement to the initial "adverbial clause" parallels the view put forth in a number of papers by Henry Yungli Chang (e.g. Chang 2006a, 2009) that certain adverbs in Formosan languages that are formally marked as verbs and occur sentence initially are in fact matrix verbs, taking the following "main" verb as complement. The adverbial verb – traditionally analyzed as dependent on the
following verb – would then, according to this view, actually be the head of the construction. An example is given from Paiwan.

289. Paiwan (Formosan)

\[ ku-g[\text{in}]alu \quad a \quad k[\text{em}]im \quad a \quad hung \]

1.SG.GEN-be.slow[PF.PERF] LNK search[AF] NOM book

'I searched the book slowly.' (Chang 2006b)
Alternatively: 'I was slow to search the book.'

Also in some of these constructions, an intervening linker is possible (or perhaps sometimes even obligatory) between the initial adverbial verb and the following clause, as in the Paiwan example above. The status of the linker as a coordinator or complementizer in these cases is not quite clear. This further illuminates the problems involved in analyzing these constructions.

Interesting topics for further study, then, are the questions of whether there are several subtypes of asymmetric coordination, whether some of them are better viewed as cases of complementation (in which the "main" clause is better viewed as the complement), how complex relation markers apparently stemming from the grammaticalization of asymmetric coordination emerge and how wide-spread they are, and how the phenomenon of asymmetric coordination (or possible subtypes of it) relates to Formosan adverbial verbs. A closer look at either of these phenomena may shed considerable light on the other.
In this chapter, the distribution of the morpho-syntactic types identified in the previous chapter will be mapped out across the semantic relations discussed in chapter 4 and across three Austronesian geographic areas.

7.1 Distribution of semantic relations across structural types

We shall begin by investigating whether there is a systematic correlation between structural types and semantic relations in Austronesian clause combining. Cross-linguistically, there are studies supporting such a correlation, but there are also studies denying that it exists. Scholars claiming that the correlation is real often explain it in terms of iconic motivations: the idea that the representation of something tends to bear a resemblance to that which is represented. Iconic motivations have been applied to explain several form-meaning correlations in language, and the underlying principle is that signs are easier to process when they resemble their denotees, or as Givón (1985) puts it: "a coded experience is easier to store, retrieve and communicate if the code is maximally isomorphic to the experience" (p 189). An iconic pattern in clause combining, then, would not only presuppose a systematic form-meaning correlation, but also that it could be argued that semantic relations with tighter morpho-syntactic representations are also tighter conceptually in some respect.

Making use of the tightness hierarchy of clause combining constructions established in chapter 6, we are in a position to investigate whether there is a form-meaning correlation in Austronesian TC clause combining or not and to confirm or reject iconic motivations as valid explanations in this case.

7.1.1 Paradigmatic iconicity

There are several types of iconicity. In his typology of signs, Peirce (1932) distinguishes between three types of icons: metaphors, images and diagrams. While metaphors and images would seem to be opposing extremes in the
sense that the former achieve their effect via creative association and the latter are mere copies of their referents (or some trait in them, cf. onomatopoeia), diagrammatic iconicity ends up in between, being about isomorphic patterns between parts of the representation and parts of the concept. This type of iconicity is usually intended when various grammatical phenomena are explained in terms of iconic motivations, including clause combining constructions. In modern linguistics, several different labels have been used for various subtypes of diagrammatic iconicity. The following are most prominently featured:

- **Iconicity of independence**
  Greater independence between objects or states of affairs is reflected by greater independence between the linguistic forms representing them (*He knew that Joanne had left* vs. *He wanted Joanne to leave*).

- **Iconicity of distance** (aka proximity or cohesion)
  Greater conceptual distance is represented by greater formal distance in terms of number and type of morphemes (A'ara: *kmanya*, 'my father', vs. *nogu mola iara*, 'my canoe').

- **Iconicity of complexity** (aka quantity86)
  Greater complexity in meaning is reflected by greater complexity in linguistic form (Balinese: *luh*, 'female', vs. *luhlhuh*, 'females').

- **Iconicity of sequence** (aka order, or tense iconicity)
  Temporal order between states of affairs is represented by linear order in linguistic form (*John left the party and we turned up the music* vs. *We turned up the music and John left the party*).

I would like to separate these into what I will call paradigmatic iconicity, encompassing the first three, and linear iconicity, encompassing the last one. The difference between the two is obvious. Paradigmatic iconicity has to do with morpho-syntactic patterns mimicking the configuration of conceptual units, resulting in more or less elaborate morpho-syntax. It can only be established by comparing relations between expressions – paradigms in the structuralist sense – with relations between their meaning. Linear iconicity is of a different nature: it has to do with linguistic strings mimicking temporal order. That is, all other things being equal, implicitly related states of affairs

---

86 Haspelmath (2008), however, separates iconicity of complexity from iconicity of quantity, giving causatives vs non-causatives as an example output of the former and comparative inflection as an example output of the latter.
7. DISTRIBUTIONS

tend to be interpreted to occur in the order they are reported. Although linear iconicity is clearly relevant for some types of clause combining constructions, applying iconic explanations to correlations between structural makeup and semantic relations in clause combining must make appeals to paradigmatic iconicity.

Givón (1980), for instance, finds that the tightness of a complement clause (from full clauses via subjunctives and infinitives to nominalizations) correlates systematically cross-linguistically with the semantics of its matrix verb. This is what he calls the binding hierarchy for complements. In Givón (1985), iconic principles are suggested to be an important motivation for this correlation (as well as for other morpho-syntactic structures in language). The RRG school of theoretical linguistics (e.g. Foley & Van Valin 1984; Van Valin 2005; and various other RRG work) has as one of its major components the iconic relation between a semantic hierarchy of interclausal relations and a syntactic hierarchy of clausal relations (refer to section 2.2.2 above). The RRG approach is of a stipulative nature, based on theoretical claims, though richly backed up in various studies by empirical data. Cristofaro (2005), on the other hand, takes an altogether empirical approach by investigating the morpho-syntax of subordinate clauses, but she also finds that various semantic relations correlate systematically with the presence or absence of various morpho-syntactic features. As a result, she establishes several hierarchies of semantic relations based on different morpho-syntactic traits. The hierarchies are explained partly in terms of iconicity (refer to section 2.2.3). Cristofaro's hierarchies are largely in agreement with the hierarchies of clausal relations in RRG.

On the other hand, some studies deny a form-meaning correlation in clause combining, which would mean that iconic motivations have no effect in this area, or at least that they are overshadowed by other factors in shaping morpho-syntactic structures. Lehmann (1988), for instance, defines six scalar morpho-syntactic parameters, along which clause combining constructions may assume different values from compact to elaborate (refer to section 2.2.1). But he claims that semantic relations play no important role in the classification of clause combining constructions along these parameters. He states: "It rather appears that the grammatical types that will emerge on the basis of the above six parameters cut across the semantically different clause

87 It should be acknowledged, of course, that this is only the case if the inferential framework does not imply differently. The nature of the states of affairs may make some temporal orders more likely than other according to the human experience, regardless of linear order in syntax. In John hurt his knee, he fell off his bike the falling off the bike is naturally interpreted as the reason for the injury, and therefore, as occurring first in real time, even though it is reported last in the utterance. In John left the party and we turned up the music, on the other hand, both states of affairs are equally likely to be the reason or the result in the relation, and only then does the linear order become important, and the reason is interpreted as being reported before the result.
linkage relations" (Lehmann 1988, p 183). Harris (1988) adopts a similar yet slightly more cautious position and argues that what is the unmarked clause combining construction for a specific semantic relation varies diachronically, and that speakers may depart from unmarked constructions for pragmatic, stylistic or syntactic reasons. Harris (1989) makes two claims related to the present discussion: (1) that syntactic structures and discourse-pragmatic structures are independent of one another, so that two independent propositions may well be represented by a main-subordinate clause format in syntax, and (2) that conceptual categories such as "condition" and "concession" are mere focal points in the domain of semantic relations and cannot be discretely delimited. Both of these claims would seem to distort to some extent any form-content correlations.

In the following subsections, we will have a closer look at the situation from an Austronesian perspective. Using the Austronesian sample as an arbitrator, which of the two sides above is favored?

### 7.1.2 Distributional patterns

In this section, we shall see how structural types of combined clauses are distributed in Austronesian languages over semantic relations.

The table below displays the number of languages from the extended sample in which a specific structural type has been attested to be used for a specific semantic relation. (The relative distributions do not differ significantly between the core sample of 30 languages and the extended sample of 43 languages, so only the results from the extended sample will be accounted for here.)

<table>
<thead>
<tr>
<th>RELATIONS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>co-occurrence</td>
<td>15</td>
<td>5</td>
<td>39</td>
<td>13</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>posteriority</td>
<td>7</td>
<td>1</td>
<td>16</td>
<td>39</td>
<td>31</td>
<td>43</td>
</tr>
<tr>
<td>anteriority</td>
<td>15</td>
<td>4</td>
<td>20</td>
<td>10</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>term. boundary</td>
<td>6</td>
<td>1</td>
<td>25</td>
<td>5</td>
<td>3</td>
<td>31</td>
</tr>
<tr>
<td>initial boundary</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>condition</td>
<td>6</td>
<td>5</td>
<td>43</td>
<td>8</td>
<td>14</td>
<td>43</td>
</tr>
<tr>
<td>concession</td>
<td>3</td>
<td>1</td>
<td>22</td>
<td>8</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>purpose</td>
<td>15</td>
<td>10</td>
<td>32</td>
<td>9</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td>reason</td>
<td>11</td>
<td>3</td>
<td>38</td>
<td>9</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>result</td>
<td>6</td>
<td>3</td>
<td>19</td>
<td>30</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td><strong>ALL</strong></td>
<td>32</td>
<td>17</td>
<td>43</td>
<td>42</td>
<td>34</td>
<td>43</td>
</tr>
</tbody>
</table>
As illustrated, structural type 3 (involving non-deviating, subordinate, explicit constructions) is represented in all languages, and type 4 (non-deviating, coordinate, explicit constructions) is represented in all but one language. Type 1 (deviating constructions) and 5 (non-deviating, coordinate, explicit constructions) are also fairly well represented in the sample languages. Type 2, on the other hand, occurs in less than half the sample. The semantic relations are all well attested in the sample, with the exception of initial boundary, which is attested in less than one fifth of the sample.

We can already see in this table that some relations are better represented towards the loose end of the type-scale (types 4 and 5) than others. However, the results must be calibrated in order to give a more truthful picture. The following table specifies the number of languages in which both a specific type (regardless of its meaning) and a specific semantic relation (regardless of its form) are attested for each of the type-relation combinations.

**Table 25. Number of extended sample languages for which a specific type (for any relation) and a specific relation (by any type) are attested**

<table>
<thead>
<tr>
<th>TYPES &gt;</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>deviating?</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>subordinate?</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>explicit?</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RELATIONS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>co-occurrence</td>
<td>32</td>
<td>17</td>
<td>42</td>
<td>41</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>posteriority</td>
<td>32</td>
<td>17</td>
<td>43</td>
<td>42</td>
<td>34</td>
<td>43</td>
</tr>
<tr>
<td>anteriority</td>
<td>24</td>
<td>14</td>
<td>31</td>
<td>31</td>
<td>24</td>
<td>31</td>
</tr>
<tr>
<td>term. boundary</td>
<td>22</td>
<td>12</td>
<td>31</td>
<td>30</td>
<td>25</td>
<td>31</td>
</tr>
<tr>
<td>initial boundary</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>condition</td>
<td>32</td>
<td>17</td>
<td>43</td>
<td>42</td>
<td>34</td>
<td>43</td>
</tr>
<tr>
<td>concession</td>
<td>20</td>
<td>9</td>
<td>26</td>
<td>25</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>purpose</td>
<td>32</td>
<td>17</td>
<td>42</td>
<td>41</td>
<td>34</td>
<td>42</td>
</tr>
<tr>
<td>reason</td>
<td>32</td>
<td>17</td>
<td>43</td>
<td>42</td>
<td>34</td>
<td>43</td>
</tr>
<tr>
<td>result</td>
<td>26</td>
<td>16</td>
<td>36</td>
<td>36</td>
<td>28</td>
<td>36</td>
</tr>
<tr>
<td>ALL</td>
<td>32</td>
<td>17</td>
<td>43</td>
<td>42</td>
<td>34</td>
<td>43</td>
</tr>
</tbody>
</table>

The table shows, for instance, that 24 languages (marked cell) have type 1 constructions (for any relation), while they, at the same time, are attested for anteriority relations (by any type). Now, we are in a position to compare the two tables and determine the share the numbers in Table 24 make up of the numbers in Table 25. This will provide the calibrated frequencies of occurrence in the sample for each type-relation combination. The figures are displayed in Table 26.
Table 26. Calibrated percentage of languages using specific types for specific relations

<table>
<thead>
<tr>
<th>TYPES &gt;</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>deviating?</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>subordinate?</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>explicit?</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RELATIONS</th>
<th>100%</th>
<th>100%</th>
<th>100%</th>
<th>100%</th>
<th>100%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>co-occurrence</td>
<td>47%</td>
<td>29%</td>
<td>93%</td>
<td>32%</td>
<td>64%</td>
<td>100%</td>
</tr>
<tr>
<td>posterity</td>
<td>22%</td>
<td>6%</td>
<td>37%</td>
<td>93%</td>
<td>91%</td>
<td>100%</td>
</tr>
<tr>
<td>anteriority</td>
<td>63%</td>
<td>29%</td>
<td>65%</td>
<td>32%</td>
<td>4%</td>
<td>100%</td>
</tr>
<tr>
<td>term. boundary</td>
<td>27%</td>
<td>8%</td>
<td>81%</td>
<td>17%</td>
<td>12%</td>
<td>100%</td>
</tr>
<tr>
<td>initial boundary</td>
<td>40%</td>
<td>0%</td>
<td>75%</td>
<td>13%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>condition</td>
<td>19%</td>
<td>29%</td>
<td>100%</td>
<td>19%</td>
<td>41%</td>
<td>100%</td>
</tr>
<tr>
<td>concession</td>
<td>15%</td>
<td>11%</td>
<td>85%</td>
<td>32%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>purpose</td>
<td>47%</td>
<td>59%</td>
<td>76%</td>
<td>22%</td>
<td>24%</td>
<td>100%</td>
</tr>
<tr>
<td>reason</td>
<td>34%</td>
<td>18%</td>
<td>88%</td>
<td>21%</td>
<td>15%</td>
<td>100%</td>
</tr>
<tr>
<td>result</td>
<td>23%</td>
<td>19%</td>
<td>53%</td>
<td>83%</td>
<td>46%</td>
<td>100%</td>
</tr>
</tbody>
</table>

In this table we can see, for instance, that of those languages in which both type 1 constructions are attested (for whatever relation) and anteriority relations are attested (by whatever type), 63% (marked cell) use type 1 constructions for anteriority.

If there was a perfectly systematic correlation between meaning and form, we should find the peak percentage for the different semantic relations to be evenly distributed across the different types, so that certain relations would correlate with tighter morpho-syntax and others with looser morpho-syntax. Representing the semantic relations by lines in a coordinate system with the different constructions on the x-axis and number of languages on the y-axis, a perfectly systematic correlation would look something like the hypothetical idealized graph in Figure 8 (for a selection of semantic relations).
This way of illustrating distributional patterns can be used to establish hierarchical relationships between semantic relations based on where they peak along the x-axis. At the same time, it can easily show more complex patterns, such as possible multiple peaks for the same relation, or irregularities revealing the lack of hierarchical relationships between semantic relations.

As is plainly shown in Table 26, we do not find the symmetrical distribution of Figure 8 in the 43-language Austronesian sample. The results are nonetheless interesting. Based on the data in Table 26, the actual distributions are illustrated in Figure 9 and Figure 10. I have separated the temporal relations and the co-varying relations into two figures, partly because it would be difficult to discern each line if all relations were represented in the same graph, and partly because some interesting similarities and differences between the two groups of relations become apparent this way.
Three points are especially noteworthy. First, the contour of many of the curves are strikingly similar between the two graphs. There are several pairwise similarities between specific temporal relations and specific co-varying relations:

- The distribution of posteriority and result relations are nearly identical, although posteriority are stronger with type 5 constructions.
Co-occurrence relations and conditional relations also display near identical curves for types 2-5. The relatively strong position of co-occurrence relations with type 1 constructions is also interesting and will be commented on in the next subsection.

Although initial boundary is only attested in eight of the sample languages, the distribution of this relation parallels that of reason relations quite well.

These three pairs of relations are teased out of Figure 9 and Figure 10 and put together pairwise in the figures below, making the similarities between the curves obvious.

**Figure 11. Distribution of semantic relations across structural types (extended language sample): result and posteriority.**

**Figure 12. Distribution of semantic relations across structural types (extended language sample): condition and co-occurrence.**
It is no coincidence that these relations pair up in these ways. Recall the discussion in chapter 5 on polysemic patterns identifying co-occurrence-condition and posteriority-result as the two most common relation pairs to employ the same constructions for their representation in the Austronesian sample. The curves above are a reflection of this and are yet another way of illuminating the close connection between these pairs of relations. The similarities between the initial boundary and reason curves, however, is not as strongly underpinned as the others, since the initial boundary relation is poorly attested among the sample languages. However, as the development of reason relations from initial boundary relations is identified in some studies as a crosslinguistic trend (e.g. Genetti 1986; Ohori 1996; Kortmann 1997), it would be interesting to find out if the curve in Figure 13 can be maintained and reinforced with further Austronesian data.

The second point has to do with terminal boundary and purpose relations. The distribution of terminal boundary and purpose is displayed pairwise in Figure 14 below.
Here, we can see that the distributional pattern for terminal boundary and purpose, which are reported in several studies to be closely connected, do not show the same kind of similarity, at least not in the left section of the figure. Instead, while purpose relations dominate with types 1 and 2 constructions, terminal boundary relations are very weak in this region, especially with type 2 constructions. From type 3 constructions onwards, however, the distribution of purpose and terminal boundary are very similar. Nonetheless, this shows that at the constructional level, purpose and terminal boundary tend to employ somewhat different types of constructions for their realization in Austronesian languages.

Third, nearly all relations peak on type 3 constructions (non-deviating, subordinate, explicit). This is a very salient tendency, and we can be confident that this type of constructions is the most common in Austronesian languages for TC relations. However, while all other relations peak on type 3 constructions, posteriority relations and result relations peak further to the right; on type 4 constructions. Note that the single difference between types 3 and 4 is that type 3 involves subordination and type 4 involves coordination. Posteriority and result relations thus tend to be expressed by coordination to a greater extent. Both these relations are also very weak on types 1 and 2 and comparatively weak on type 3. This evidence leads us to conclude that posteriority and result relations are normally realized by looser constructions than the other relations under study in the Austronesian languages. Furthermore, if these two relations stand out as employing looser morpho-syntax, it is also quite clear the purpose relations tend to employ somewhat tighter morpho-syntax than the other relations. While peaking with type 3 constructions – as most relations do – its type 3 value is not
higher than its type 2 value to the same extent that it is for all other relations. And both its type 1 and, especially, its type 2 values are very strong (co-occurrence relations actually parallel the type 1 value of purpose relations, and anterioirity relations surpass it, but both of these relations have much weaker type 2 values). Taken together, this suggests the semantic relations hierarchy in example (290) below.

290. purpose < co-occurrence < anteriority < posteriority

This hierarchy is not implicational. It does not tell us that if a language uses a specific construction for a specific relation, it must use the same or a looser construction for any relation to the right in the hiearachy, which is a common reading of typological hierarchies. It does tell us, however, that when an Austronesian language represents TC relations by means of constructions of varying tightness, it is more common than not that (i) purpose relations can be represented by tighter constructions than most other relations, and/or that (ii) posteriority and result relations can be represented by looser constructions than most other relations. With only two cutoff points, the hierarchy in (290) is not very stratified, and its resolution of detail is far from that claimed by the RRG semantic relations hierarchy, for instance. But the data does not support a more stratified hierarchy with the approach used here.

Some of the relations have quite strong secondary peaks in various places along the type axis. Strong secondary peaks, of course, are difficult to translate into one dimentional hierarchies, since they tend to blur cutoff points. In this respect, representations in a coordinate system, as in Figure 9 and Figure 10 above, are superior and provide a better overview of the true picture. Some of the secondary peaks will be briefly discussed and exemplified below.

It is interesting to note the high type 1 values of co-occurrence and anteriority. Many of the type 1 constructions representing co-occurrence or anteriority relations are found in the Western Malayo-Polynesian languages where a deviating subordinate clause – often a nominalization or a converb form – represents the co-occurring or anterior state of affairs in these relations. When this type of construction is used for both co-occurrence and anteriority in the same language, verbal morphology usually distinguish between the two readings. Consider the Tagalog examples below where
different gerund prefixes convey different temporal settings for the state of affairs represented by the verb.

291. Tagalog

a. **panging-isda ni juan,**
   NZR-fish GEN NAME

   s[um]a-sama ang kapatid niya
   RD[FX]:AF.IPFV.join TOP brother 3.SG.GEN

   'When Juan goes fishing, his brother goes along.'
   (Schachter & Otanes 1972, p 446)

b. **hu-hugas-an ng katulong ang pinggan**
   RD:IRR-wash-PF GEN maid TOP dishes

   pagka-walis niya ng sahig
   NZR.PFTV.sweep 3.SG.GEN GEN floor

   'The maid will wash the dishes when [/after] she has swept the floor.' (Schachter & Otanes 1972, p 446)

Co-occurrence and/or anteriority constructions of type 1 are not exclusive to the Western Malayo-Polynesian languages, although they seem to be more common in languages in this group. In the core sample (30 languages), they were attested in 9 Western Malayo-Polynesian languages, 2 Central-Malayo-Polynesian languages, and 1 Eastern Malayo-Polynesian language. In the extended sample (43 languages), they were attested in 2 additional Western Malayo-Polynesian languages (11 total), and 3 additional Eastern Malayo-Polynesian languages (4 total).

Reason and initial boundary relations also have quite salient local peaks with type 1 constructions. In the case of initial boundary, this could be due to the relation being rarely attested among the sample languages in general (although further data might reinforce this peak). The reason relations represented by type 1 constructions mostly consist of nominalized clauses combined with some kind of reason marker – for instance, a preposition or adverbial subordinator. Maori provides an example.
292. Maori

\[
i \text{toomuri mātou i te puni-nga ai}
\]
PST be.late 2.PL.IN from the block-NOM ANA

\[
o \text{te haurahi i nga horonga}
\]
GEN the road from the.PL slip

'We were late because the slips blocked the road.'
lit. 'We were late because of the blocking of the road by the slips.'
(Bauer 1993, p 68)

These constructions are particularly widespread in Polynesia (all Central-East Malayo-Polynesian examples were from Polynesian sample languages) but are fairly frequent also in the western part of Austronesian.

Finally, both co-occurrence and conditional relations have a distinctive upward bend with type 5 constructions (non-deviating, coordination, implicit). Together with posteriority and result, co-occurrence and condition are the most common semantic relations that may be represented by two implicitly coordinate intact clauses. In the case of posteriority and result, their high type 5 values are a natural continuation of their being represented by looser morpho-syntax in general. Indeed, a linear iconic interpretation of implicitly coordinate clauses representing a temporal sequence (classified as posteriority in this study) or a result relation seems perfectly natural, as long as context do not indicate otherwise. However, given the appropriate context, two implicitly coordinate intact clauses may also be interpreted as representing a conditional or co-occurrence relation in quite a few Austronesian languages.

Contextual cues are not the only factors determining the interpretation of implicit coordination, although they seem to be more important in some languages than others. In actual fact, there is often an intricate interplay between context and several other factors, primarily prosody, grammatical marking and encoding of participants. In Loniu (Hamel 1994), for instance, intonation first indicates if two clauses should be interpreted as clause combining or not. If the answer is yes and the verbs of both clauses are in the realis form and the subjects are identical, the construction commonly (but not necessarily) represents temporal succession. If the verbs are in the realis form and the subjects are different, the construction commonly (but not necessarily) represents co-occurrence. If the verb of the first clause is in the potential mood form, on the other hand, presence or absence of the intentional particle \textit{ma} is decisive: if present, the construction commonly (but not necessarily) represents co-occurrence, and if absent, the construction commonly (but not necessarily) represents a conditional relation. In general,
conditional constructions are often associated with some kind of irrealis marking (recall the discussion in 5.2.1). Thus, in some Austronesian languages, conditional interpretations of type 5 constructions are distinguished from other possible interpretations by the use of irrealis morphology in one or both of the clauses. However, the patterns are complex and often involve the distinction between open and counterfactual conditionals as well. Furthermore, as in several other languages in the world, the distinction between future time co-occurrence and open conditional relations is one that some Austronesian languages do not make. Note the Erromangan example below.

293. Erromangan

\textit{kokomle-sentvi nacave}

1.PL.IN.FUT-wipe kava

\textit{kokl-amprog-i ovon nevyarep}

1.PL.IN.FUT-call-CONST PL youth

'When we wipe the kava, we will call the youths.' or

'If we wipe the kava, we will call the youths.' (Crowley 1998, p 268)

The difference in meaning between the two interpretations – if the states of affairs are anticipated to take place (co-occurrence) or if the likelihood of realization is not known (condition) – is slight in most practical situations. This is, of course, the reason for the distinction being absent in some languages (see chapter 5 for statistics).

7.1.3 Range of variation

In this section, we will discuss a different approach to measuring the morpho-syntactic tightness with which various semantic relations are represented in Austronesian languages. As illustrated in the previous section, there is great variation in how different relations are distributed across structural types in Austronesian languages, with almost all relations being attested in at least one language for each morpho-syntactic type. This raises the question of how great the average structural variation is for each relation. And further, what is the difference in structural tightness between the relation represented by the tightest and loosest constructions in average? If the cline is very steep, we have a strong case for a systematic correlation between form and meaning in the Austronesian sample. On the other hand, if the cline is quite level, and especially if the range of structural variation for the relations tends to be larger than the difference between the tightest and
loosest relations, then there is very little support for a form-meaning
correlation in Austronesian.

Let us begin by calculating the average tightness of the tightest and
loosest constructions attested for each language and each relation in the
extended language sample. By way of illustration, the tightest and loosest
constructions in each language for reason relations, together with the
resulting average values, are displayed in Table 27 below.

Table 27. Tightest and loosest constructions for reason relations (bracketed lan-
guages are in the extended sample only).

<table>
<thead>
<tr>
<th>ID</th>
<th>Language</th>
<th>Group</th>
<th>Area</th>
<th>tightest</th>
<th>loosest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amis</td>
<td>Form</td>
<td>Taiwan</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>[Seediq]</td>
<td>Form</td>
<td>Taiwan</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Palauan</td>
<td>WMP</td>
<td>Micronesia</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>C. Cagayan Agta</td>
<td>WMP</td>
<td>Philippines</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Tagalog</td>
<td>WMP</td>
<td>Philippines</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Western Subanon</td>
<td>WMP</td>
<td>Philippines</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>[Sarangani Manobo]</td>
<td>WMP</td>
<td>Philippines</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Tboli</td>
<td>WMP</td>
<td>Philippines</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Yakan</td>
<td>WMP</td>
<td>Philippines</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Sangir</td>
<td>WMP</td>
<td>Sulawesi</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>[Tondano]</td>
<td>WMP</td>
<td>Sulawesi</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>Muna</td>
<td>WMP</td>
<td>Sulawesi</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>Coastal Konjo</td>
<td>WMP</td>
<td>Sulawesi</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Ma’anyan</td>
<td>WMP</td>
<td>Borneo, Indonesia</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>Labuk Kadazan</td>
<td>WMP</td>
<td>Borneo, Malaysia</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>Karo Batak</td>
<td>WMP</td>
<td>Sumatra</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>[Acehnese]</td>
<td>WMP</td>
<td>Sumatra</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>Eastern Cham</td>
<td>WMP</td>
<td>Vietnam</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>[Indonesian]</td>
<td>WMP</td>
<td>Indonesia</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>Kambera</td>
<td>CMP</td>
<td>Nusa Tenggara, Indon.</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>Tetun</td>
<td>CMP</td>
<td>Nusa Tenggara, Indon.</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>Leti</td>
<td>CMP</td>
<td>Maluku</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>23</td>
<td>Buru</td>
<td>CMP</td>
<td>Maluku</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>24</td>
<td>Taba</td>
<td>SHWNG</td>
<td>Maluku</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>25</td>
<td>Loniu</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>26</td>
<td>Manam</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>27</td>
<td>Mbula</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>Yabem</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>[Iwal]</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>Mekeo</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>31</td>
<td>Nakanai</td>
<td>Oceanic</td>
<td>PNG</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>32</td>
<td>[Hoava]</td>
<td>Oceanic</td>
<td>Solomon Islands</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>33</td>
<td>Longgu</td>
<td>Oceanic</td>
<td>Solomon Islands</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>34</td>
<td>Erromanga</td>
<td>Oceanic</td>
<td>Vanuatu</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>35</td>
<td>[Lenakel]</td>
<td>Oceanic</td>
<td>Vanuatu</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>36</td>
<td>Araki</td>
<td>Oceanic</td>
<td>Vanuatu</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>37</td>
<td>[Big Nambas]</td>
<td>Oceanic</td>
<td>Vanuatu</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 28 below shows the average tightness for each of the relations. Also indicated in the table are the mean value between the average tightest and average loosest constructions, as well as the difference between the average tightest and average loosest constructions, for each relation.

Table 28. Average tightness (construction type), extended language sample

<table>
<thead>
<tr>
<th>Relations</th>
<th>tightest</th>
<th>mean</th>
<th>loosest</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anteriority</td>
<td>2.10</td>
<td>2.55</td>
<td>3.00</td>
<td>0.90</td>
</tr>
<tr>
<td>Initial boundary</td>
<td>2.63</td>
<td>2.75</td>
<td>2.88</td>
<td>0.25</td>
</tr>
<tr>
<td>Purpose</td>
<td>2.31</td>
<td>2.83</td>
<td>3.36</td>
<td>1.05</td>
</tr>
<tr>
<td>Reason</td>
<td>2.51</td>
<td>2.94</td>
<td>3.37</td>
<td>0.86</td>
</tr>
<tr>
<td>Term. boundary</td>
<td>2.74</td>
<td>2.98</td>
<td>3.23</td>
<td>0.48</td>
</tr>
<tr>
<td>Concession</td>
<td>2.81</td>
<td>2.98</td>
<td>3.15</td>
<td>0.35</td>
</tr>
<tr>
<td>Co-occurrence</td>
<td>2.26</td>
<td>3.17</td>
<td>4.07</td>
<td>1.81</td>
</tr>
<tr>
<td>Condition</td>
<td>2.65</td>
<td>3.20</td>
<td>3.74</td>
<td>1.09</td>
</tr>
<tr>
<td>Result</td>
<td>2.94</td>
<td>3.58</td>
<td>4.22</td>
<td>1.28</td>
</tr>
<tr>
<td>Posteriority</td>
<td>3.30</td>
<td>4.00</td>
<td>4.70</td>
<td>1.40</td>
</tr>
</tbody>
</table>

As shown in the table, the difference between the tightest mean value (2.55 for anteriority) and the loosest mean value (4.00 for posteriority) is 1.45, while the difference for each relation – i.e. the range of variation – varies between 0.25 (for initial boundary) and 1.81 (for co-occurrence). The average difference for the whole set of relations is 0.95. These two numbers (1.45 and 0.95) are not hugely different – which would not seem to lend strong support for a correlation between form and meaning in Austronesian TC clause combining. However, some interesting facts can be teased out. The graph in Figure 15 illustrates clearly some of the points to be made below.
The area between the upper line (average loosest constructions) and the lower line (average tightest constructions) illustrate the fact that, for most relations, the range of structural variation is quite comparable to the total cline of the middle line (mean values) from tightest to loosest construction types. Apart from the upward bend for result and posteriority relations, the line is quite flat, and especially around the middle, the relative order for some of the relations are probably due to chance. An indication of this is that the rank of relations are different depending on whether they are ranked by the average tightest or average loosest constructions types representing them – as shown by the jagged nature of upper and lower lines in Figure 15 – or whether they are ranked by the mean values. The varying range of variation for the different relations might in fact be more interesting than the cline of the mean values, since it indicates the constructions used for some relations are more varied than those used for other relations. This line of research has not been pursued further here, but it is an interesting topic for future research.

Another indication of the weak correlation between meaning and form in Austronesian clause combining comes from comparing the results from the 30 languages of the core sample with the results from the 43 languages of the extended sample above. The graph for the core sample is essentially the
same as that in Figure 15 above, and the values differ only slightly from those of the extended sample: the tightest mean value is 2.55 in the extended sample and 2.54 in the core sample (both for anteriority); the loosest mean value is 4.00 and 3.97, respectively (both for posteriority); and the average range of variation is 0.95 and 1.02, respectively. However, around the middle of the curve, where it is the flattest, the rank between relations differs a little, since reason relations has a somewhat looser relative rank between concession and co-occurrence in the core sample.

Posteriority and result, however, still stand out quite clearly as being represented by looser constructions than the other relations. This is in agreement with the findings of the previous subsection. Behind these results is the tendency to use type 4 constructions for both of these relations (typically tense iconic then and so constructions, respectively). And between posteriority relations and result relations, the difference between their mean values is even greater than that between result and condition. It seem that posteriority relations are represented by looser constructions by average than any other TC relations in Austronesian. On the tighter side of the curve, anteriority seem to take the mean value down a bit compared to the rest of the relations. However, although type 1 constructions are comparatively numerous with this relation, anteriority is still most commonly represented by type 3 constructions, and the difference in average tightness from most other relations is so small that it is difficult to draw any far reaching conclusions with regards to this relation.

Representing the results of the average tightness analysis presented in this subsection in a hierarchy, we would have to group most relations together on the left-hand side.

```
anteriority
  purpose
  terminal boundary
  reason
  initial boundary
  concession
  co-occurrence
  condition
```

294. result < posteriority

This hierarchy says that posteriority relations are by average represented by looser constructions than the rest of the relations under study in the sample languages, followed by result relations, and then the rest of the relations.

However, even though the difference in mean value between most neighboring semantic relations in Figure 15 is small, the difference in mean value between the tightest and loosest relations in the group to the left in hierarchy (294) is large in relative terms (0.65 between anteriority and
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

condition). The table below displays the difference in mean value between all the semantic relations individually in this study.

Table 29. Difference in mean value, extended language sample

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ant.</td>
<td>0.20</td>
<td>0.28</td>
<td>0.39</td>
<td>0.43</td>
<td>0.43</td>
<td>0.62</td>
<td>0.65</td>
<td>1.03</td>
<td>1.45</td>
</tr>
<tr>
<td>Initial b.</td>
<td>0.08</td>
<td>0.19</td>
<td>0.23</td>
<td>0.23</td>
<td>0.42</td>
<td>0.45</td>
<td>0.83</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>Purp.</td>
<td>0.11</td>
<td>0.15</td>
<td>0.15</td>
<td>0.34</td>
<td>0.37</td>
<td>0.75</td>
<td>1.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason</td>
<td>0.04</td>
<td>0.04</td>
<td>0.23</td>
<td>0.26</td>
<td>0.64</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Term. b.</td>
<td>0.00</td>
<td></td>
<td>0.19</td>
<td>0.22</td>
<td>0.60</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conc.</td>
<td>0.19</td>
<td>0.22</td>
<td>0.60</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-oc.</td>
<td>0.03</td>
<td>0.41</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond.</td>
<td></td>
<td></td>
<td>0.38</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td></td>
<td></td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since the cutoff points in hierarchy (294) are defined by a difference in mean value of at least 0.38 (result-condition), and since the largest difference in mean value between any two neighboring relations ranked in the same group to the left in the hierarchy is 0.20 (initial boundary-anteriority), we may use any number between these two as a threshold, and define cutoff points between individual relations that are comparable to the two cutoff points in hierarchy (294). The number 0.30 is convenient, being as it is 25% of the difference between the tightest (anteriority) and loosest (posteriority) relations in the sample. The table below lists the relations between which the difference in mean value is larger than 0.30, from the viewpoint of each of the semantic relations to the left in hierarchy (294).

Table 30. Cutoff points between individual relations (based on extended sample)

**Anteriority**
- anteriority < reason
- anteriority < terminal boundary
- anteriority < concession
- anteriority < co-occurrence
- anteriority < condition
- anteriority < result
- anteriority < posteriority

**Initial boundary**
- initial boundary < co-occurrence
- initial boundary < condition
- initial boundary < result
- initial boundary < posteriority
7. DISTRIBUTIONS

Purpose
purpose < co-occurrence
purpose < condition
purpose < result
purpose < posteriority

Reason
anteriority < reason
reason < result
reason < posteriority

Terminal boundary
anteriority < terminal boundary
terminal boundary < result
terminal boundary < posteriority

Concession
anteriority < concession
concession < result
concession < posteriority

Co-occurrence
anteriority < co-occurrence
initial boundary < co-occurrence
purpose < co-occurrence
co-occurrence < result
co-occurrence < posteriority

Condition
anteriority < condition
initial boundary < condition
purpose < condition
condition < result
condition < posteriority

7.1.4 Discussion

The results from the two previous subsections are complex. On the one hand, it is clear that result relations and, even more so, posteriority relations are generally selected by Austronesian languages for looser constructions, being represented by coordination and non-deviating morpho-syntax, as well as posteriority quite often by implicitly related clauses. Many languages also have constructions for purpose relations that are manifested by tighter co-constructions. However, when it comes to average tightness, anteriority relations are represented by even tighter morpho-syntax, though differences are slight. This suggests that with respect to at least some of the semantic relations, there are systematic correlations between meaning and form. And since the relations that stand out as looser or tighter than the rest largely correspond to those hierarchies for which paradigmatic iconicity is suggested as a molding force in other studies, it appears that Austronesian languages lend some support to the notion of paradigmatic iconicity in morpho-syntax, although the Austronesian hierarchies found in this study are far from as detailed as that suggested by RRG.

Having said that, one must acknowledge that the results are not clear-cut. The distribution patterns for some semantic relations are quite complex with one or more secondary peaks (see Figure 9 and Figure 10), and the average tightness cline is fairly flat for most relations in relation to possible structural variation within each semantic relation (see Figure 15). The range
of variation within a semantic relation is sometimes even greater than the average variation between the relations. Thus, the Austronesian support for paradigmatic iconicity in clause combining constructions is somewhat weak, although it can certainly be argued to be present in the case of posteriority and result relations.

A comparison between the hierarchies of earlier studies and the hierarchies established in this study is in place. The RRG semantic hierarchy of interclausal relations, adapted to include only the relevant semantic relations, as well as Cristofaro's (2005) hierarchies for adverbial subordination, are presented below, followed by the hierarchies of the present study.

295. RRG semantic hierarchy of interclausal relations (adapted)

TIGHT
- purposive
- circumstance (= after-relations, when-relations)
- reason
- conditional
- concession
- simultaneous (= while-relations)
- sequence (= before-relations, then-relations)

LOOSE

296. Hierarchy of adverbial subordination, verb form (Cristofaro 2005)

TIGHT
- purpose
- before, after, when
- reason, condition

LOOSE

297. Hierarchy of adverbial subordination, arguments (Cristofaro 2005)

TIGHT
- purpose
- before, after, when, reason, condition

LOOSE
298. Austronesian hierarchy, distribution of semantic relations

**TIGHT**
- purpose
- anteriority, terminal boundary, reason, initial boundary, concession, co-occurrence, condition
- result, posteriority

**LOOSE**

299. Austronesian hierarchy, average tightness

**TIGHT**
- anteriority, purpose, terminal boundary, reason, initial boundary, concession, co-occurrence, condition
- result
- posteriority

**LOOSE**

As illustrated, all hierarchies are basically congruent, although the RRG hierarchy is more fine-grained than the others. The main point of difference is that posteriority relations (before-relations) are featured as tighter in Cristofaro's hierarchies than in all other hierarchies. The obvious reason is that she is only concerned with subordinate constructions and does not, for instance, include relations of temporal succession (mostly coordinate in structure) in her posteriority category. Naturally, this results in increased tightness for the category. Note also that the RRG hierarchy features relations of circumstance (which include anteriority (after-)relations) near the top of the hierarchy, which is very much in agreement with the results from Austronesian languages. The inclusion of when-relations here is also interesting, since Austronesian languages (especially WMP) have similar constructions for (certain types of) co-occurrence relations and anteriority relations (recall the discussion in subsection 7.1.2 above; illustrative examples are given in (291)). The wide range of co-occurrence relations in Austronesian languages may also partly be reflected in the RRG hierarchy by the tightness of when-relations and looseness of while-relations (see section 4.3.1 above for some discussion about these two types of co-occurrence relations).
7.2 Distribution across areas

In order to determine if any of the structural types of clause combining constructions are concentrated in any specific Austronesian region, the types will be mapped out across three Austronesian macro-regions in this section. The groups are somewhat unevenly divided with regards to the number of Austronesian languages spoken in each, but they were chosen for specific reasons.

The focus construction languages (including the Austronesian languages in Taiwan and in the Philippine area) were grouped in one group, which was thus chosen for a specific structural feature. Then the Austronesian languages of Malaysia, Indonesia and mainland Southeast Asia were grouped in another group, because these languages are generally spoken on large land areas, and as the speakers of Austronesian languages spread into these areas, they came into contact with speech communities speaking very different languages from their own. Finally, the remainder of the Austronesian languages consists of the languages of the vast Pacific Ocean. These represent speech communities widely separated by large amounts of water, and most of them did not encounter speakers of foreign languages when they first arrived to their islands. In other words, the two last groups were chosen because of the very different social and environmental contexts in which they evolved. The geographic areas are more precisely delimited below:

**PT – The Philippines-Taiwan area**
(193 languages (Gordon 2005))
The Austronesian languages spoken in Taiwan and the Philippines, with the addition of the Philippine type languages spoken on the northern tip of Sulawesi (e.g. Sangir) and Palauan.

**MI – The Malaysia-Indonesia area**
(577 languages (Gordon 2005))
The Austronesian languages spoken in Malaysia, Indonesia, Brunei and on the Southeast Asian mainland (including coastal islands), with the exception of the Philippine-type languages spoken on the northern tip of Sulawesi and the Oceanic languages of West Papua.

**OC – The Pacific Ocean area**
(498 languages (Gordon 2005))
The Oceanic languages.
For convenience, I have allowed the Pacific Ocean area to coincide with the Oceanic subgroup. The sample languages within each region are listed in the table below.

Table 31. Geographic areas (languages marked by E are in extended sample only)

<table>
<thead>
<tr>
<th>Philippines-Taiwan area</th>
<th>Malaysian-Indonesian area</th>
<th>Oceanic area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Amis</td>
<td>Tondano (E)</td>
<td>Loniu</td>
</tr>
<tr>
<td>2. Seediq (E)</td>
<td>Muna</td>
<td>Manam</td>
</tr>
<tr>
<td>3. Palauan</td>
<td>Coastal Konjo</td>
<td>Mbula</td>
</tr>
<tr>
<td>4. Central Cagayan Agta</td>
<td>Ma’anyan</td>
<td>Yabem</td>
</tr>
<tr>
<td>5. Tagalog</td>
<td>Labuk Kadazan (E)</td>
<td>Iwal (E)</td>
</tr>
<tr>
<td>6. Sarangani Manobo (E)</td>
<td>Karo Batak</td>
<td>Mekeo</td>
</tr>
<tr>
<td>7. Western Subanon</td>
<td>Acehnese (E)</td>
<td>Nakanai</td>
</tr>
<tr>
<td>8. Tboli</td>
<td>Eastern Cham</td>
<td>Hoava (E)</td>
</tr>
<tr>
<td>9. Yakan</td>
<td>Indonesian (E)</td>
<td>Longgu</td>
</tr>
<tr>
<td>10. Sangir</td>
<td>Kambera</td>
<td>Erromangan</td>
</tr>
<tr>
<td>11. Tetun</td>
<td>Lenakel (E)</td>
<td></td>
</tr>
<tr>
<td>12. Leti (E)</td>
<td>Araki</td>
<td></td>
</tr>
<tr>
<td>13. Buru</td>
<td>Big Nambas (E)</td>
<td></td>
</tr>
<tr>
<td>14. Taba</td>
<td>Tinrin</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I have chosen to investigate the distribution of morpho-syntactic types across geographical regions rather than across genealogical language groups within Austronesian. This was borne out of necessity as the Austronesian languages at almost all levels split into one or more rather small branches and one very large branch (see section 3.1.1). It is quite impossible to find genealogically related groups that are comparable in size if the whole Austronesian family is to be covered. Instead, the distribution was mapped out across the above three geographically defined groups, although one of them coincides with a genealogically defined language group: Oceanic.

It must also be stressed that since only 43 languages in the extended sample and 30 languages in the core sample were investigated, the number of languages from each group is very low, and the results must be seen as mere indications of distributional patterns and as a starting point for further study rather than firm evidence of distributional patterns. The number of
sample languages from the Philippines-Taiwan area is 10 and 8 for the two samples, respectively; from the Malaysia-Indonesia area 14 and 10, respectively; and from the Pacific Ocean area 19 and 12, respectively. Thus, the percentages will not stand up to any rigorous statistical analysis. This is a flaw that many traditional typological investigations suffer from, and one that can be countered to some extent by a well balanced sample. Nonetheless, the results in our case allow us to detect some interesting tendencies. The distribution of construction types are displayed below for each geographical area (black represents the core sample; gray represents the extended sample).

*Table 32. Distribution of construction types in the PT area*

<table>
<thead>
<tr>
<th></th>
<th>Type 1 (N)</th>
<th>Type 1 (%)</th>
<th>Type 2 (N)</th>
<th>Type 2 (%)</th>
<th>Type 3 (N)</th>
<th>Type 3 (%)</th>
<th>Type 4 (N)</th>
<th>Type 4 (%)</th>
<th>Type 5 (N)</th>
<th>Type 5 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core (C)</td>
<td>8</td>
<td>100.0</td>
<td>3</td>
<td>37.5</td>
<td>8</td>
<td>100.0</td>
<td>8</td>
<td>100.0</td>
<td>5</td>
<td>62.5</td>
</tr>
<tr>
<td>Ext. (E)</td>
<td>10</td>
<td>100.0</td>
<td>3</td>
<td>30.0</td>
<td>10</td>
<td>100.0</td>
<td>10</td>
<td>100.0</td>
<td>6</td>
<td>60.0</td>
</tr>
</tbody>
</table>

*Figure 16. Distribution of construction types in the PT area*

*Table 33. Distribution of construction types in the MI area*

<table>
<thead>
<tr>
<th></th>
<th>Type 1 (N)</th>
<th>Type 1 (%)</th>
<th>Type 2 (N)</th>
<th>Type 2 (%)</th>
<th>Type 3 (N)</th>
<th>Type 3 (%)</th>
<th>Type 4 (N)</th>
<th>Type 4 (%)</th>
<th>Type 5 (N)</th>
<th>Type 5 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core (C)</td>
<td>7</td>
<td>70.0</td>
<td>4</td>
<td>40.0</td>
<td>10</td>
<td>100.0</td>
<td>10</td>
<td>100.0</td>
<td>9</td>
<td>90.0</td>
</tr>
<tr>
<td>Ext. (E)</td>
<td>9</td>
<td>64.3</td>
<td>6</td>
<td>42.9</td>
<td>14</td>
<td>100.0</td>
<td>13</td>
<td>92.9</td>
<td>11</td>
<td>78.6</td>
</tr>
</tbody>
</table>
7. DISTRIBUTIONS

Figure 17. Distribution of construction types in the MI area

Table 34. Distribution of construction types in the OC area

<table>
<thead>
<tr>
<th></th>
<th>Type 1</th>
<th>Type 2</th>
<th>Type 3</th>
<th>Type 4</th>
<th>Type 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N)</td>
<td>(%)</td>
<td>(N)</td>
<td>(%)</td>
<td>(N)</td>
</tr>
<tr>
<td>Core (C)</td>
<td>7</td>
<td>53.3</td>
<td>4</td>
<td>33.3</td>
<td>11</td>
</tr>
<tr>
<td>Ext. (E)</td>
<td>13</td>
<td>68.4</td>
<td>8</td>
<td>42.1</td>
<td>18</td>
</tr>
</tbody>
</table>

Figure 18. Distribution of construction types in the OC area

The first thing to observe is that the distribution of types 2 to 4 is nearly identical in all regions. For type 3 (non-deviating, subordinate, explicit) and type 4 (non-deviating, coordinate, explicit), this is obviously because these constructions occur in almost every sample language. For type 2 (non-deviating, subordinate, implicit), we may conclude that it is not only relatively less frequent in Austronesian languages as a whole, but also in each area; it has no regional centers.

The second thing to observe, then, is that while type 1 constructions are universal in the Philippine-Taiwan area of Austronesian (100.0% in both
samples), they are comparatively less frequent in the Malaysia-Indonesia area (70.0% in the core sample), and even less so in the Pacific Ocean area (53.3% in the core sample). At the same time, type 5 constructions seem to be quite a bit more frequent in the Malaysia-Indonesia and Pacific Ocean areas (91.7% and 90.0%, respectively, in the core sample) than they are in the Philippine-Taiwan area (62.5% in the core sample). It seems that the Oceanic languages have a propensity to employ looser morpho-syntax in clause combining in general (more of type 5, less of type 1) than do the Philippine-Taiwan area languages. The languages of the Malaysia-Indonesia area end up somewhere in between, though closer to the Oceanic languages in this respect. This is very interesting since it coincides with what is generally taken to be the direction of spread of Austronesian languages in history: from their origin in Taiwan, via the Philippines and east and west in the Indonesian archipelago, including New Guinea, then finally to the Pacific Ocean islands. It seems that as the Austronesian languages spread into new areas, they increasingly dispensed with the more compact clause combining constructions with deviating morpho-syntax (type 1), in favor of implicit coordination of clauses (type 5). This may have implications for historical studies on the ways in which the Austronesian languages have spread and for studies of language-contact situations in the Austronesian area. These interesting topics, however, must be left to future research.
This section summarizes the main findings in this study and wraps up some discussions. In the introduction chapter, a number of aims were presented that are listed again below for convenience. They fall under three main domains of investigation.

- **Semantically, the study aims**
  (i) to provide an overview of the semantic space of TC relations in the Austronesian languages. What relations tend to be present in the sample languages, and what nuances of meaning are expressed? Which distinctions are common in Austronesian, and which are rare?
  (ii) to investigate the polysemy of relation markers. To what extent can one relation marker be used in constructions expressing different semantic relations? Are polysemic patterns in Austronesian compatible with diachronic tendencies of semantic shift reported for relation markers cross-linguistically?

- **Morpho-syntactically, the study aims**
  (iii) to develop a typology of Austronesian clause combining based on three parameters related to features common to clause combining constructions.
  (iv) to review some Austronesian constructions of special interest, in particular, the Austronesian preference for deriving relation markers for some relations from verbs, and the tendency to use coordinators to link initial subordinate clauses to the rest of the sentence.

- **Distributionally, the study aims**
  (v) to map out the distribution of the construction types across semantic relational categories, such as co-occurrence, condition, and concession, etc.
(vi) to evaluate the notion of paradigmatic iconicity in clause combining, i.e. the extent to which the degree of integration between states of affairs for different semantic relations correspond to the degree of tightness for morpho-syntactic types.
(vii) to shed light on the extent to which structural types correlate with geographic areas within the Austronesian region.

In the following sections, these aims are revisited, and a short summary of the investigations and results are presented.

8.1 Semantics of Austronesian clause combining

The 10 semantic relations selected for investigation were not all attested in all languages. While co-occurrence, posteriority, conditional, purpose and reason relations were attested in all or nearly all sample languages, initial boundary relations were attested in a minority of the languages. The other relations were attested in more than half the sample languages. The overall picture is outlined in Figure 19 below.

![Figure 19. Percentage of sample languages for which the relations were attested](image)

8.1.1 Relation internal nuances of meaning

Some relation internal semantic differentiation was found, such that some languages have specialized relation markers to express some subcomponent
of meaning for a particular relation or some aspect of meaning in addition to the relational meaning. There are also cases where the sample languages use only one relation marker for different relation internal meanings, where different relation markers might have been motivated. The findings are summarized below.

**CO-OCCURRENCE:**

**continuous durative vs. punctual momentary**
This distinction is quite common among the sample languages and is similar to that between *while* and *when* in English, though sometimes with language specific idiosyncrasies. Example: Samoan *'a'o* (continuous) vs. *ina 'ua* (momentary).

**individual vs. indefinite time**
Another common distinction is that between two individual co-occurring states of affairs (~*when*), on the one hand, and habitually co-occurring states of affairs with indefinite time specifications (~*whenever*), on the other. Example: Lenakel *nian* ('*when*', lit. 'time') vs. *nian miin* ('*whenever*', lit. 'times').

**simultaneity vs. same temporal setting**
This distinction is not represented by different relation markers in the sample languages. On the contrary, it seems possible in most Austronesian languages to use a general co-occurrence relation marker for both simultaneity (e.g. *He ate his sandwiches when I packed his bags*) and sequences co-occurring within the same temporal setting but not actually simultaneously (e.g. *When the ship berthed, the passengers went ashore*).

**rare co-occurrence distinctions**
Some rarer co-occurrence distinctions found in one or a few languages only include markers indicating (i) continuity vs. discontinuity with prior discourse context, as in Buru *bama* (~'and when… then…') vs. *mama* (~'but when [P takes place], then [Q takes place]'); (ii) general co-occurrence vs. co-occurrence with an initial stage of a state of affairs, as in Muna *rato* ('*when*') vs. *paka-* ('*just when*'); and (iii) general co-occurrence vs. near materialization co-occurrence, as in Muna *rato* ('*when*') vs. *kirakira* ('*when about to*').
CLAUSE COMBINING IN AUSTRONESIAN LANGUAGES

POSTERIORITY

**general vs. immediate posteriority**
In a handful of languages, a special posteriority marker is used to indicate that the interval of time interval between the states of affairs being very short or non-existent. Example: Taba *malai* ('then') vs *turus* ('then straight away').

ANTERIORITY

**general vs. immediate anteriority**
Mirroring the previous distinction, some languages use special anteriority markers to indicate that the time interval between the states of affairs is very short or non-existent. Example: Acehnese *ban* ('just', 'as soon as').

TERMINAL BOUNDARY

**general vs. distant terminal boundary**
This is a rare distinction attested only in a few languages. These languages have special relation markers to indicate that the temporal distance to the terminal boundary state of affairs is fairly long. Example: Rapanui *ata* ('until', general) vs. *'ahará* ('until', distant).

INITIAL BOUNDARY

There were no relation internal distinctions found for initial boundary. This might be due to the fact that initial boundary is the least attested relation among the sample languages.

CONDITION

**open vs. counter fact condition**
This distinction is made in about half the sample languages. Some employ the same relation markers for both and make the distinction by other grammatical means, while other languages have specialized counter fact conditional relation markers. Example: Samoan *'āfai* (open condition) vs. *'ana* (counter fact condition).

**regular vs. concessive condition**
The distinction between regular condition (*if*) and concessive condition (*even if*) is made in slightly less than half the sample languages. Concessive conditionals may be open or counter factual,
just as regular conditionals, but this four-way distinction is rarely attested (probably because it is seldom dealt with in descriptive grammars). Example: Coastal Konjo pung- (‘if’, open cond.) vs. coba- (‘if’, counter fact cond.) vs. mang-...-ji (‘even if’, open conc. cond.) vs. mang-...kedde’ (‘even if’, counter fact conc. cond.).

affirmative vs. negative condition
Some of the sample languages have a separate relation marker for negative conditional relations (unless). And in at least two of these languages (Muna and Karo Batak), the marker was derived from an adverb meaning ‘only’. Example: Muna ane (‘if’) vs. tabea (‘unless’, ‘only’).

rare conditional distinctions
Some rarer conditional distinctions found in one or a few languages only include markers indicating (i) regular condition vs. exclusive condition (only if), as in Mbula so (‘if’, ‘supposedly’) vs. bela (‘only if’); (ii) regular condition vs. desired condition, as in Palauan use of hypothetical mood forms (~‘if’) vs. ułekum (‘if only’); (iii) general concessive condition vs. diminutive concessive condition, as in Tagalog sukdan (‘even if’) vs. huwag hindi (‘however little [P takes place], still [Q takes place]’); and (iv) continuity vs. discontinuity with prior discourse context, as in Buru bama (~‘and if [P takes place], then [Q takes place]’) vs. mama (~‘but if [P takes place], then [Q takes place]’).

CONCESSION

general vs. unfortunate concession
A rare concessive distinction, attested only in a single language, is made by using a specific relation marker to express a situation that is perceived as negative for the actor of the concessive clause (e.g. However much I begged her, she still didn’t give it to me). Example: Samoan ui (general concession) vs. ūgā (unfortunate concession).

general vs. disapproved concession
Another rare concessive distinction, also attested only in a single language, is made by using a specific relation marker to express disapproval on the part of the speaker (e.g. They say they won’t do the work even though you paid them). Example: Tagalog bagaman (general concession) vs. gayong (disapproved concession).
PURPOSE

affirmative vs. negative purpose
About one-third of the sample languages have a separate relation marker for negative purpose relations. Example: (West) Mekeo *ema* (‘in order that’) vs. *poka* (‘lest’).

general vs. fulfilled purpose
Typologically rare but attested in three of the Austronesian sample languages (Buru, Longgu, Manam) is the distinction between general and fulfilled purpose. Example: Longgu *ni* (‘[P took place] in order that [Q would take place, whether it did or not]’) vs. *ania* (‘[P took place] in order that [Q would take place, and it did]’).

general vs. anticipatory purpose
The distinction between general and anticipatory purpose is found in only two sample languages (Muna and Central Cagayan Agta). Anticipatory purpose denotes a purpose that is desired but far from certain. Example: Muna *so* (‘in order that’) vs. *bhahi* (‘so that perhaps’).

REASON

reason vs. cause
Reason relations, as intended in the present study, includes both relations involving external motivation (causes) and those involving internal motivations (in the mind of an agent). Similar to most of the languages in the world, Austronesian languages seems to use the same relation marker for both.

rare reason distinctions
The following reason distinctions were found in only one language each: (i) general reason vs. reason with doubted truth value, as in Central Cagayan Agta *te* (‘because’) vs. *awá* (‘because [P took place, I doubt that Q took place]’); (ii) general reason vs. unfortunate reason, as in Samoan *'ona* (‘because’) vs. *leaga* (‘because unfortunately’); and (iii) general reason vs. exclusive reason, as in Indonesian *sebab* (‘because’) vs. *mentang-mentang* (‘just because’).

RESULT

general vs. graded result
In three of the sample languages (Indonesian, Kambera, Tboli), a special relation marker is used for a graded result. Example: Indo-
nesian kesimpulannya (‘so’, ‘conclusion’) vs. sehubungan (‘to the extent that’).

### 8.1.2 Polysemy

The investigation of polysemic patterns with Austronesian relation markers carried out in chapter 5 showed that the majority of the sample languages – 70% in the core sample – have relation markers that can be used for both condition and co-occurrence. This is by far the highest number for any pair of relations. Most languages, however, have more than one relation marker for the co-occurrence/conditional domain, and it turns out that even when languages have a general co-occurrence/conditional relation marker, they frequently also have separate relation markers for co-occurrence (or possibly only past co-occurrence) and for condition (or possibly only counter fact condition). The four most common pairs of relations to feature polysemic relation markers in the Austronesian sample are the following:

1. co-occurrence/condition (core sample freq.: 70%)
2. posteriority/result (core sample freq.: 27%)
3. condition/concession (core sample freq.: 23%)
4. purpose/reason (core sample freq.: 20%)

It was also investigated to what extent Austronesian polysemic patterns were compatible with diachronic shifts in meaning, which have been reported to be common for relation markers cross-linguistically. Three salient paths of semantic change are often described (see Genetti 1986; Ohori 1996; Kortmann 1997): (I) Markers of spatial co-occurrence (at) tend to develop into markers of temporal co-occurrence (when, while), which tend to develop into conditional markers (if). (II) Markers of spatial goal (to) tend to develop into markers of (temporal) terminal boundary (until), which tend to develop into markers of purpose ((in order) to). And (III), markers of spatial source (from) tend to develop into markers of (temporal) initial boundary (since), which tend to develop into markers of reason (because). It can be shown that Austronesian languages frequently have polysemic relation markers compatible with all three paths, although the clearest patterns concur with path (I). Another two domains of polysemy attested in the Austronesian sample languages also suggest the grammaticalization of some lexical items into relation markers: (i) from verbs for 'say' or 'speak' to conditional relation markers, and (ii) from verbs of directed motion to relation markers for terminal boundary or purpose. An overview of the findings is presented in Table 35.
Table 35. Diachronic development suggested by polysemic patterns

<table>
<thead>
<tr>
<th>Polysemy domain</th>
<th>Attested in Austronesian sample?</th>
<th>Suggested change</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sameness</td>
<td>Well attested</td>
<td>space &gt; co-oc</td>
<td>Hawaiian īā 'at, when'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>time &gt; co-oc, cond</td>
<td>Erromangan nempong 'time, when, if'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>time &gt; condition</td>
<td>Tagalog sakali 'moment, in case'</td>
</tr>
<tr>
<td>Goal</td>
<td>Sporadically attested</td>
<td>space &gt; term b &gt; purp</td>
<td>Loniu la 'go, into, until, in order to'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>space &gt; purpose</td>
<td>Muna so 'for, in order to'</td>
</tr>
<tr>
<td>Source</td>
<td>Few examples</td>
<td>space &gt; init b &gt; reas.</td>
<td>Tinrin ghe(gi) 'from, since, because'</td>
</tr>
<tr>
<td>'say', 'speak'</td>
<td>Several examples</td>
<td>say &gt; condition</td>
<td>Araki co de 'it would say, if'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>speak &gt; condition</td>
<td>Indonesian bicara 'speak, if'</td>
</tr>
<tr>
<td>Motion</td>
<td>Several examples</td>
<td>go &gt; term. boundary</td>
<td>Big Nambas dav'a 'go, until'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bring &gt; purpose</td>
<td>Tetun hodi 'bring, in order to'</td>
</tr>
</tbody>
</table>

8.2 Morpho-syntax of Austronesian clause combining

8.2.1 Structural typology

The clause combining constructions in the Austronesian sample were classified along three binary parameters. The parameters are (with values in parentheses): (A) deviating morpho-syntax in relation to a simplex clause (deviating, non-deviating), (B) syntactic relation between clauses (coordination, subordination), and (C) explicitness of semantic relation (explicit, implicit). The parameters were chosen because they relate to the most common features associated with combined clauses. The combination of parameter values results in eight logically possible types. However, for the deviating types, the distinction between explicit and implicit relations as well as that between the subordinate and coordinate relations was ignored. The five remaining types and representative constructions are listed below:
Type 1 (deviating):
- echo-subject constructions, nominalized clauses, converb forms

Type 2 (non-deviating, subordinate, implicit):
- implicit adverbial clauses

Type 3 (non-deviating, subordinate, explicit):
- explicit adverbial clauses

Type 4 (non-deviating, coordinating, explicit):
- explicit coordination, coherence adverb

Type 5 (non-deviating, coordinating, implicit):
- implicit coordination

Types 2 and 3 were attested in nearly all sample languages, and types 1 and 5 were also quite well attested, but type 2 was only attested in slightly more than one-third of both samples. The overall picture is outlined in Figure 20 below.

![Figure 20. Percentage of sample languages for which the types were attested](image)

Based on notions of morpho-syntactic compactness (fewer items, more bound morphology) and dependency (reliance on external information for the interpretation of various features), the five types were ranked in structural tightness from tight to loose in their numeric order, with type 1 as the tightest and type 5 as the loosest. This was done in order to determine if specific semantic relations were consistently represented by tighter or looser morpho-syntax.
8.1.2 Constructions of special interest

Several Austronesian languages have relation markers that apparently have grammaticalized from lexical items. These were discussed mainly in chapters 4 and 5 (focusing on their semantics). However, as the grammaticalization of lexical items lies at the interface between semantics and morphosyntax, some morpho-syntactic aspects were also dealt with there. Some of these issues will be summarized here.

One interesting finding pertains to the grammaticalization of a verb meaning 'finish' into a relation marker. This has taken place in some of the sample languages, which in some languages has resulted in a posteriority marker and in other languages an anteriority marker. At first sight, this may seem odd since the two relations are opposite in meaning to each other. However, depending on whether the verb grammaticalizes into an adverbial subordinator (or similar) or into an adverb (or similar), the relational meaning turns out differently. A series of juxtaposed clauses, such as 'I cleaned the house, it finished, I went for a walk' may be the starting point in both cases. But whether the cleaning or the walking is thought of as finished will imply two opposite temporal orders, as illustrated in (300) below.

300. a. I cleaned the house, then I went for a walk.
   from: I cleaned the house, (and when) it finished, I went for a walk.

   b. I cleaned the house after I went for a walk.
   from: I cleaned the house (when) it finished (that) I went for a walk.

Apparently, both grammaticalization paths have been taken in Austronesian languages (and sometimes even within the same language). Table 36 lists some examples.
In connection with anteriority and posteriority, it was also noted in chapter 4 that some Austronesian languages use a relation marker meaning 'not yet' or similar for posteriority, and oppositely, a relation marker meaning 'already' or similar for anteriority. The logic is, of course, that if something has not yet taken place, then it is posterior in time, while if something has already taken place, then it is anterior in time. One example of each is given below, followed by some examples of relation markers of these kinds from various Austronesian languages in Table 37.

301. Buru

\[
\text{da mata } \text{mo} \text{hede}, \text{ da stori gam naa } [...] \\
3.\text{SG die not.yet 3.\text{SG speak like this}}
\]

'Before he died, this is what he said...' (Grimes 1991, p 421)

302. Tetun

\[
\text{kawen } \text{ti'a}, \text{ tur iha ne'e dei} \\
\text{marry already sit LOC this only}
\]

'After (we) are married, (we) must live here.' (Klinken 1999, p 236)
Table 37.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th>RELATION MARKER</th>
<th>LITERAL MEANING</th>
<th>TC RELATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Konjo</td>
<td>ri maeng-</td>
<td>at already</td>
<td>anteriority</td>
</tr>
<tr>
<td>Indonesian</td>
<td>sesudah</td>
<td>already</td>
<td>anteriority</td>
</tr>
<tr>
<td>Tetun</td>
<td>ti’a</td>
<td>already</td>
<td>anteriority</td>
</tr>
<tr>
<td>Mbula</td>
<td>kek</td>
<td>already</td>
<td>anteriority</td>
</tr>
<tr>
<td>Sangir</td>
<td>bedang ta</td>
<td>still not</td>
<td>posteriority</td>
</tr>
<tr>
<td>Buru</td>
<td>mohede</td>
<td>not yet</td>
<td>posteriority</td>
</tr>
<tr>
<td>Mbula</td>
<td>zen</td>
<td>not yet</td>
<td>posteriority</td>
</tr>
<tr>
<td>Samoan</td>
<td>‘ae le’i</td>
<td>but not yet</td>
<td>posteriority</td>
</tr>
</tbody>
</table>

In chapter 6 (section 6.2.2), a construction that I call asymmetric coordination, characteristic of many Austronesian languages, was taken up for discussion. This construction involves the use of a coordinator to connect a fronted topicalized adverbial clause to the rest of the sentence. It is not obvious whether this is a case of coordination or subordination. And if it is subordination, it is not obvious that the fronted adverbial clause is the subordinate one. In some languages the coordinator is homophonous with a complementizer, which leaves it open for analysis that the fronted adverbial clause is actually the main clause, with the following clause as its complement. Compare the alternative translations of the example below.

303. Coastal Konjo

\[\text{sangge-na } \text{rie’ pulisi na kung-a’-cari’carita} \]
until-3.POSS exist police and 1.PERI-INTR-tell.stories

'Until the police arrived, I told stories.' (Timothy Friberg, p.c. 2004)
Alternatively: '(It was) until the police arrived that I told stories.'

That Coastal Konjo \textit{na} can also be used as a regular coordinator is illustrated by the next example.

304. Coastal Konjo

\[\text{ku-pa-kanre-i } \text{na } \text{ku-pa-inung-i} \]
1.ERG-CAUS-eat-3.ABS and 1.ERG-CAUS-drink-3.ABS

'I fed him and I gave him a drink.' (Friberg 1996, ex 103)
The choice made in the present study is to view these examples as coordinate. The purpose of the coordinator is to facilitate cognitive processing by rendering both clauses as independent. That way, the hearer does not need to bother with trying to fit a cumbersome initial constituent into the structure of the following clause.

8.3 Distributional patterns

Chapter 7 was concerned with the distributions of the types established in chapter 6, both across geographical areas and, more importantly, across the semantic relations discussed in chapter 4.

By measuring the number of languages for which a specific type of construction was used for a specific semantic relation, it could be seen that (somewhat simplified) purpose relations had a higher peak than the other relations for type 2 constructions, while posteriority and result relations had a higher peak than the other relations for type 4 relations. This led to the statistical hierarchy presented below.

```
305.  purpose  <  co-occurrence  <  posteriority
      condition  reason       result
      concession
      reason
```

This indicates that when an Austronesian language represents TC relations by means of constructions of varying tightness, it is more common than not that (i) purpose relations can be represented by tighter constructions than most other relations, and/or that (ii) posteriority and result relations can be represented by looser constructions than most other relations.

Focusing on the range of variation that most sample languages display with respect to the different construction they have at their disposal for expressing one and the same relation, the tightest and loosest types used for each relation were indicated in a graph. This is reproduced below.
This figure shows that, for most relations, the range of structural variation is quite comparable to the total cline of the mean values (middle line). The difference in structural tightness between anteriority to the left and for instance condition near the right end is quite small, and particularly for relations around the middle, the relative order is probably due to chance. Noting that the steepest cline is caused by result and posteriority relations on the right, the following hierarchy was established:

\[
\text{antiority} \quad \text{purpose} \quad \text{terminal boundary} \quad \text{reason} \quad \text{initial boundary} \quad \text{concession} \quad \text{co-occurrence} \quad \text{condition} < \text{result} < \text{posteriority}
\]

Both hierarchies are largely congruent with hierarchies suggested earlier in cross-linguistic studies (Foley & Van Valin 1984; Van Valin & LaPolla 1997; Van Valin 2005; Cristofaro 2005). This means that the results of the
present study lend some support to a universal correlation between structural type and semantic relation, such that some relations are consistently represented by comparatively looser morpho-syntax. Such a correlation may be motivated by paradigmatic iconicity, i.e. the notion that the degree of integration between states of affairs in different semantic relations is reflected by the degree of tightness in the morpho-syntactic types representing them. However, since the difference in tightness between most relations in Austronesian languages was shown to be very small, I would hesitate to make the hierarchies too stratified.

Finally, the distribution of types across three Austronesian macro-regions (Philippines-Taiwan, Malaysia-Indonesia and Pacific Ocean) revealed few differences between the areas, although two tendencies could be detected: the Oceanic languages generally employ slightly looser morpho-syntax for clause combining constructions, while the Austronesian languages of the Philippine-Taiwan area employ slightly tighter morpho-syntax. This coincides with what is generally assumed to be the direction of spread of Austronesian languages in history: from Taiwan via the Philippines to Indonesia and Malaysia, and then from Indonesia, via New Guinea, into the Pacific Ocean. It appears that as the languages spread, they lost some of the tighter constructions and replaced them with new, looser constructions.


REFERENCES


MacDonald, Lorna. 1990. **A grammar of Tauya.** Berlin: Mouton de Gruyter.


REFERENCES


Ross, John Robert. 1986. Infinite syntax! (originally presented in 1967 as the author's PhD thesis at Massachusetts Institute of Technology, Boston, under the title Constraints on variables in syntax). Norwood: ABLEX.


Appendix 1: Native speaker's questionnaire

The questionnaire presented to native speakers of Austronesian languages consisted of 70 simplex clauses and 42 complex clause constructions constructed from the simplex clauses. The informant was asked to translate each of the sentences into his or her language, and if possible, to provide information about the meaning of the words used in the translation. The 70 plus 42 clauses are listed below.

1. The phone rang.
2. The child saw the drawing in my hand.
3. I bought a new record.
4. He got badly hurt.
5. I sat down.
6. The boys played a game of cards.
7. Their mother hugged them hard.
8. They will fire him.
9. She loved someone else.
10. I hit the boy.
11. They played together.
12. She sang happily to herself.
13. They left.
14. Her husband called.
15. He pushed me.
16. Her mother arrived.
17. I went to the record store.
18. The man changed a lot.
19. I didn't hear what she said.
20. They had a nice vacation.
21. I tripped on the wire.
22. They went to bed.
23. They packed their bags.
24. I stood up.
25. We celebrated.
26. We will celebrate.
27. They went to the airport.
28. I saw the animals.
29. She kissed Tony.
30. She will kiss Tony.
31. They went to town.
32. The coconut fell in the man's head.
33. His friend waited outside.
34. The woman got a raise.
35. He sneezed.
36. The girl whistled happily.
37. He opened the door.
38. His friend died.
39. We met at school.
40. He saw me.
41. He laughed.
42. He will laugh.
43. The man spoke on the phone.
44. His mother heard about the accident.
45. They heard the plane.
46. They fired him.
47. His body shook.
48. She listened to the radio.
49. We shared everything.
50. They heard a loud explosion.
51. He went to her office.
52. He looked at his award.
53. He wanted to see her again.
54. They saw the fire.
55. The boys were doing their homework.
56. He got the job.
57. The sun shined on his nose.
58. The bomb exploded.
59. She worked hard.
60. I fell.
61. We left the house.
62. I waved my arms.
63. He did a good job.
64. They all looked up.
65. She brushed her hair.
66. The bird flew away.
67. She spoke quietly.
68. She spoke so quietly.
69. She spoke very quietly.
70. He smiled.
1. The man spoke on the phone while his friend waited outside.
2. The girl whistled happily while she brushed her hair.
3. The boys were doing their homework when the phone rang.
4. He sneezed whenever the sun shined on his nose.
5. He smiled whenever he looked at his award.
6. They had a nice vacation until her mother arrived.
7. They played together until they went to bed.
8. The man had changed a lot since his friend died.
9. We have shared everything ever since we first met at school.
10. They went to the airport after they packed their bags.
11. She sang happily to herself after her husband called.
12. The boys played a game of cards before they went to town.
13. Their mother hugged them hard before they left.
14. I hit the boy as soon as he opened the door.
15. They all looked up immediately when they heard the plane.
16. They heard a loud explosion immediately before they saw the fire.
17. We left the house immediately before the bomb exploded.
18. I fell because he pushed me.
19. He went to her office since he wanted to see her again.
20. The woman got a raise because she worked hard.
21. I didn't hear what she said because she spoke very quietly.
22. The coconut fell in the man's head so that he got badly hurt.
23. I tripped on the wire so that I fell.
24. He laughed so his body shook.
25. She spoke so quietly that I didn't hear what she said.
26. She spoke quietly so I didn't hear what she said.
27. He pushed me so that I would fall.
28. I sat down in order for the child to see the drawing in my hand.
29. I went to the record store to buy a new record.
30. I stood up to see the animals.
31. His mother heard about the accident if she listened to the radio.
32. We will celebrate if he gets the job.
33. He will laugh if he sees me.
34. He would laugh if he saw me.
35. He would have laughed if he had seen me.
36. The bird would have flown away if I had waved my arms.
37. She will kiss Tony even if she loves someone else.
38. She would have kissed Tony even if she had loved someone else.
39. She kissed Tony although she loved someone else.
40. They will fire him even if he does a good job.
41. They would have fired him even if he had done a good job.
42. They fired him although he did a good job.
Appendix 2: Linguist's questionnaire

The following questionnaire was presented, together with a brief description of my dissertation project, to a number of linguists specializing in Austronesian languages.*

QUESTIONNAIRE
Temporal and co-varying clause combining in Austronesian languages
PhD thesis
Niklas Jonsson
Stockholm University

Dear fellow linguist,

I have prepared the present questionnaire with the hope of obtaining valuable data on combined clause constructions in Austronesian from professional linguists specializing in one or more of these languages. Therefore, I would be immensely grateful if you would take the time to fill in the questionnaire. Information from specialists in particular Austronesian languages, such as yourself, is vital for the study and for me to be able to make accurate and detailed descriptions of the linguistic patterns and to make sound and firm generalizations.

The questions in the questionnaire may take a while to answer in detail. However, do not hesitate to answer some or all of the questions sketchily if you cannot find the time to elaborate. A little information is better than no information at all. Your contribution will be greatly valued, and I would like to thank you sincerely beforehand for taking the time to provide answers to my questions.

* The questionnaire was sent in slightly different versions to different linguists, as it was updated over time. The versions do not differ significantly, however, and the questions are the same in all versions. The version presented here represents the latest edition.
The questions are general and apply to each of the interclausal relations I am interested in, which are the ones listed below (see also example sentences at the end of the questionnaire):

A. co-occurrence ("while-, when-relations")
B. posteriority ("before-relations")
C. anteriority ("after-relations")
D. terminal boundary ("until-relations")
E. initial boundary ("since-relations")
F. condition ("if-relations")
G. concession ("although-relations")
H. purpose ("to-, in order that-relations")
I. reason ("because-relations")
J. result ("so that-relations")

QUESTIONS:

I. Specify the language to be analyzed (L2).

II. What constructions are commonly used in L2 to express each of the ten relations (A-J) listed above? Please, illustrate with enumerated examples including interlinear morphemic translations into English.

If the same construction is used to cover several relations, please indicate this. If different constructions (or different individual morphemes) are used to express different shades of meaning within the domain of one relation (such as counter-fact condition vs. open condition, or immediate vs. neutral anteriority or posteriority), please illustrate this as well.

III. For each of the constructions you have identified, please comment on morpho-syntactic restrictions on their use. Compare the two clauses in each example with the corresponding simplex clauses. Relevant questions to be addressed are:

- What tenses, aspects, moods (including subjunctive forms) are possible/impossible/obligatory in each the two clauses (or clause-like structures)?

- Is the use of certain special forms of the verb (such as infinitives, nominalizations, participles, converbs, clause chains medial verbs, etc.) possible/impossible/
obligatory in any of the two clauses (or clause-like structures)? Please specify.

- Are any other grammatical morphemes possible/impossible/obligatory in any of the two clauses (or clause-like structures)? Please specify.

- Does the construction differ, obligatorily or optionally, depending on whether any of the arguments in one clause is identical to any of the arguments in the other clause? If so, in what way?

- Is the reversed order between the two clauses (or clause-like structures) possible? If so, would it change the meaning?

IV. Which of your L2 examples would you say involve subordination of clauses and which ones would you say involve coordination of clauses, and on what grounds? (A characterization in terms of "more or less" subordinated/coordinated rather than "either or" subordinated/coordinated may be useful.)

V. If possible, comment on the historical source of the different structures you have identified.

EXAMPLE SENTENCES:

The following section consists of 42 example sentences covering the relations investigated. The headings make a more fine-grained division of the semantic space covered by temporal and co-varying relations between states of affairs than the list of relations in (A-J) above. No assumption, however, should be made that all or the same distinctions are necessarily made in L2.

Co-occurrence:
1. The boys were doing their homework when the phone rang.
2. The man spoke on the phone while his friend waited outside.
3. The girl whistled happily while she brushed her hair.

Indefinite time:
4. He sneezed whenever the sun shined on his nose.
5. He smiled whenever he looked at his award.
Posteriority:
6 The boys played a game of cards before they went to town.
7 Their mother hugged them hard before they left.

Immediate posteriority:
8 They heard a loud explosion immediately before they saw the fire.
9 We left the house immediately before the bomb exploded.

Anteriority:
10 They went to the airport after they packed their bags.
11 She sang happily to herself after her husband called.

Immediate anteriority:
12 I hit the boy as soon as he opened the door.
13 They all looked up immediately when they heard the plane.

Terminal boundary:
14 They had a nice vacation until her mother arrived.
15 They played together until they went to bed.

Initial boundary:
16 The man had changed a lot since his friend died.
17 We have shared everything ever since we first met at school.

Open condition:
18 His mother heard about the accident if she listened to the radio.
19 We will celebrate if he gets the job.
20 He will laugh if he sees me.
21 He would laugh if he saw me.

Counter-fact condition:
22 He would have laughed if he had seen me.
23 The bird would have flown away if I had waved my arms.

Open concessive condition:
24 She will kiss Tony even if she loves someone else.
25 They will fire him even if he does a good job.

Counter-fact concessive condition:
26 She would have kissed Tony even if she had loved someone else.
27 They would have fired him even if he had done a good job.
Concession:
28 She kissed Tony although she loved someone else.
29 They fired him although he did a good job.

Purpose:
30 He pushed me so that I would fall.
31 I sat down in order for the child to see the drawing in my hand.
32 I went to the record store to buy a new record.
33 I stood up to see the animals.

Reason:
34 I fell because he pushed me.
35 He went to her office since he wanted to see her again.
36 The woman got a raise because she worked hard.
37 I didn't hear what she said because she spoke very quietly.

Result:
38 The coconut fell in the man's head so that he got badly hurt.
39 I tripped on the wire so that I fell.
40 He laughed so his body shook.
41 She spoke so quietly that I didn't hear what she said.
42 She spoke quietly so I didn't hear what she said.
Appendix 3: Principal source material

The grammars and other language descriptions used for each of the sample languages are listed below.

**Acehnese**
Durie 1985

**Amis**
Liu 2003a, 2003b; Wu 1995

**Araki**
Francois 2002

**Big Nambas**
Fox 1979

**Buru**
Grimes 1991

**Central Cagayan Agta**
Healey 1960; Mayfield 1972

**Coastal Konjo**
Friberg 1992; Friberg 2003

**Eastern Cham**
Blood 1977; Thurgood 2005

**Eastern Kadazan**
Hurlbut 1988; Hurlbut 1990

**Erromangan**
Crowley 1998; Crowley 2002b

**Hawaiian**
Elbert & Pukui 1979; Hawkins 1982; Pukui & Elbert 1986
Hoava
Davis 2003

Indonesian
Englebretson 2003; Ewing 2005; Macdonald 1976; Northern Illinois University (Indonesian-English internet dictionary); Sneddon 1996; Sterner et al 1976

Iwal
Bradshaw 2001

Kambera
Klamer 1998

Karo Batak
Woollams 1996

Kusaiean
Good 1989; Lee 1975

Lenakel
Lynch 1978; Lynch 1983; Lynch 2001

Leti
Engelenhoven 1995

Longgu
Hill 1992, 2002

Loniu
Hamel 1994

Ma'anyan
Gudai 1985

Manam
Lichtenberk 1983

Maori
Bauer 1993; Biggs 1969

Mbula
Bugenhagen 1995
Mekeo
Jones 1998

Muna
Berg 1989

Nakanai
Johnston 1980

Palauan
Josephs 1975

Rapanui
Du Feu 1996

Samoan
Mosel & Hovdhaugen 1992

Sangir
Maryott 1979

Sarangani Manobo
Dubois 1976

Seediq (Paran dialect)
Holmer 1996

Taba
Bowden 1997, 2000, 2001

Tagalog
Himmelmann 2005b; Llamzon 1976; Northern Illinois University (internet Tagalog dictionary); Northern Illinois University (internet Tagalog grammar); Ramos & Cena 1990; Schachter & Otanes 1972

Tboli
Forsberg 1992; Porter 1977

Tetun (Fehan dialect)
Klinken 1999

Tinrin
Osumi 1995
Tondano
Sneddon 1975

Western Subanon
Hall 1973

Yabem
Dempwolff 2005; Ross 2002

Yakan
Brainard & Behrens 2002
Appendix 4: Language samples

The tables here list the sample languages for both samples and the Austronesian subgroups they represent. The A-column shows the total number of languages for each group according to Gordon (2005). The B-column shows the actual number of languages in the samples.

Table 38. Core sample

<table>
<thead>
<tr>
<th>Mutually exclusive subgroups (non-genealogical classifications in brackets), sample languages in bold face</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austronesian</td>
<td>1268</td>
<td>30</td>
</tr>
<tr>
<td>[Formosan]</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>East Formosan: <strong>Amis</strong> (Taiwan)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>MP</td>
<td>1248</td>
<td>29</td>
</tr>
<tr>
<td>[Unclassified]</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>[WMP]</td>
<td>536</td>
<td>13</td>
</tr>
<tr>
<td>Northern Philippine: <strong>Central Cagayan Agta</strong> (Philippines)</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>Meso-Philippine: <strong>Tagalog</strong> (Philippines)</td>
<td>61</td>
<td>1</td>
</tr>
<tr>
<td>Southern Philippine: <strong>Western Subanon</strong> (Philippines)</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>South Mindanao: <strong>Tboli</strong> (Philippines)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Sama-Bajaw: <strong>Yakan</strong> (Philippines)</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Palauic: <strong>Palauan</strong> (Micronesia)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sangiric: <strong>Sangir</strong> (Sulawesi)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Muna-Buton: <strong>Muna</strong> (Sulawesi)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>South Sulawesi: <strong>Coastal Konjo</strong> (Sulawesi)</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Barito: <strong>Ma’anyan</strong> (Borneo)</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>Northwest Borneo: <strong>Labuk Kadazan</strong> (Borneo)</td>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td>Sumatra: <strong>Karo Batak</strong> (Sumatra)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Malayic: <strong>Eastern Cham</strong> (Vietnam)</td>
<td>70</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>124</td>
<td>0</td>
</tr>
<tr>
<td>CEMP</td>
<td>708</td>
<td>16</td>
</tr>
<tr>
<td>[Unclassified]</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CMP</td>
<td>168</td>
<td>3</td>
</tr>
<tr>
<td>Bima-Sumba: <strong>Kambera</strong> (Nusa Tenggara)</td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>Timor: <strong>Tetun</strong> (Nusa Tenggara)</td>
<td>48</td>
<td>1</td>
</tr>
<tr>
<td>Central Maluku: <strong>Buru</strong> (Maluku)</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>EMP</td>
<td>539</td>
<td>13</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td>SHWNG: <strong>Taba</strong> (Maluku)</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>Oceanic</td>
<td>498</td>
<td>12</td>
</tr>
<tr>
<td>Admiralty Islands: <strong>Loniu</strong> (PNG)</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Schouten: <strong>Manam</strong> (PNG)</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Ngero-Vitiaz: <strong>Mbula</strong> (PNG)</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>Huon Gulf: <strong>Yabem</strong> (PNG)</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>Papuan Tip: <strong>Mekeo</strong> (PNG)</td>
<td>62</td>
<td>1</td>
</tr>
<tr>
<td>Meso-Melanesian: <strong>Nakanai</strong> (PNG)</td>
<td>66</td>
<td>1</td>
</tr>
<tr>
<td>Southeast Solomonic: <strong>Longgu</strong> (Solomon Is)</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>North Vanuatu: <strong>Araki</strong> (Vanuatu)</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>Nuclear Southern Oceanic: <strong>Erromanga</strong> (Vanuatu)</td>
<td>88</td>
<td>2</td>
</tr>
<tr>
<td>Nuclear Southern Oceanic: <strong>Tinrin</strong> (New Caledonia)</td>
<td>(88)</td>
<td>(2)</td>
</tr>
<tr>
<td>Micronesian: <strong>Kusaian</strong> (Micronesia)</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Central Pacific: <strong>Samoan</strong> (Polynesia)</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>22</td>
<td>0</td>
</tr>
</tbody>
</table>

Represented languages: 1069 (84.3 perc.)
Unrepresented languages: 199 (15.7 perc.)

Note: The Nuclear Southern Oceanic subgroup is represented by two languages, Erromanga and Tinrin, rather than just one. This is because the selection of languages was made based on the genealogy in Grimes (2000), where these languages belonged to two different subgroups, while the genealogy represented in the table here is based on more recent research (Lynch et al 2002; Gordon 2005). However, since the Nuclear Southern Oceanic subgroup is quite large and diverse, and the languages in question are spoken in two different regions, this has no real consequences for the representativeness of the sample.
Table 39. Extended sample

<table>
<thead>
<tr>
<th>Mutually exclusive subgroups (non-genealogical classifications in brackets), sample languages in bold face, italicized languages in extended sample only</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austronesian</td>
<td>1268</td>
<td>43</td>
</tr>
<tr>
<td>[Formosan]</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>East Formosan: Amis (Taiwan)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Atayalic: Seediq (Taiwan)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>MP</td>
<td>1248</td>
<td>41</td>
</tr>
<tr>
<td>[Unclassified]</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>[WMP]</td>
<td>536</td>
<td>17</td>
</tr>
<tr>
<td>Northern Philippine: Central Cagayan Agta (Philippines)</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>Meso-Philippine: Tagalog (Philippines)</td>
<td>61</td>
<td>1</td>
</tr>
<tr>
<td>Southern Philippine</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Subanun: Western Subanon (Philippines)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Manobo: Sarangani Manobo (Philippines)</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>South Mindanao: Tboli (Philippines)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Sama-Bajaw: Yakan (Philippines)</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Palauic: Palauan (Micronesia)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sangiric: Sangir (Sulawesi)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Minahasan: Tondano (Sulawesi)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Muna-Buton: Muna (Sulawesi)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>South Sulawesi: Coastal Konjo (Sulawesi)</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Barito: Ma'anany (Borneo)</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>Northwest Borneo: Labuk Kadazan (Borneo)</td>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td>Sumatra: Karo Batak (Sumatra)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Malayic</td>
<td>70</td>
<td>3</td>
</tr>
<tr>
<td>Acehnese: Acenese (Sumatra)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chamic: Eastern Cham (Vietnam)</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Malayan: Indonesian (Java)</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>135</td>
<td>0</td>
</tr>
<tr>
<td>CEMP</td>
<td>708</td>
<td>24</td>
</tr>
<tr>
<td>[Unclassified]</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CMP</td>
<td>168</td>
<td>4</td>
</tr>
<tr>
<td>Bima-Sumba: Kambera (Nusa Tenggara)</td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>Timor</td>
<td>48</td>
<td>2</td>
</tr>
<tr>
<td>Nuclear Timor: Tetun (Timor)</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Southwest Maluku: Leti (Maluku)</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Central Maluku: Buru (Maluku)</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>52</td>
<td>0</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>EMP</th>
<th>539</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHWNG: <strong>Taba</strong> (Maluku)</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>Oceanic</td>
<td>498</td>
<td>19</td>
</tr>
<tr>
<td>Admiralty Islands: <strong>Loniu</strong> (PNG)</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Schouten: <strong>Manam</strong> (PNG)</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Ngero-Vitiq: <strong>Mbula</strong> (PNG)</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>Huon Gulf</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>North Huon Gulf: <strong>Yabem</strong> (PNG)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>South Huon Gulf: <strong>Iwal</strong> (PNG)</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Papuan Tip: <strong>Mekoe</strong> (PNG)</td>
<td>62</td>
<td>1</td>
</tr>
<tr>
<td>Meso-Melanesian</td>
<td>66</td>
<td>2</td>
</tr>
<tr>
<td>Villaumez: <strong>Nakanai</strong> (PNG)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>New Ireland: <strong>Hoava</strong> (PNG)</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>Southeast Solomonic: <strong>Longgu</strong> (Solomon Is)</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>North Vanuatu: <strong>Araki</strong> (Vanuatu)</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>Nuclear Southern Oceanic</td>
<td>88</td>
<td>4</td>
</tr>
<tr>
<td>Central Vanuatu: <strong>Big Nambas</strong> (Vanuatu)</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>Erromanga: <strong>Erromanga</strong> (Vanuatu)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Tanna: <strong>Lenakel</strong> (Vanuatu)</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>New Caledonia: <strong>Tinrin</strong> (New Caledonia)</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>Micronesia: <strong>Kusaiean</strong> (Micronesia)</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Central Pacific</td>
<td>45</td>
<td>4</td>
</tr>
<tr>
<td>Samoic-Outlier: <strong>Samoan</strong> (Polynesia)</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>Rapanui: <strong>Rapanui</strong> (Polynesia)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Tahitian: <strong>Maori</strong> (Polynesia)</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Marquesian: <strong>Hawaiian</strong> (Polynesia)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>[Not represented]</td>
<td>50</td>
<td>0</td>
</tr>
</tbody>
</table>

Represented languages: 1018 (80.3 perc.)
Unrepresented languages: 250 (19.7 perc.)